Diagnostic rather than prognostic markers-relationship between EpCAM overexpression and lung cancer: a meta-analysis

Tingting Zhu^, Xiaolin Peng^, Ziwei Cheng, Dongwei Xing, Minguang Zhang^

Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China

Contributions: (I) Conception and design: T Zhu, M Zhang; (II) Administrative support: D Xing, M Zhang; (III) Provision of study materials or patients: T Zhu, D Xing; (IV) Collection and assembly of data: T Zhu, X Peng; (V) Data analysis and interpretation: T Zhu, X Peng, Z Cheng; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Dongwei Xing; Minguang Zhang. Shanghai Municipal Hospital of Traditional Chinese Medicine, Shanghai University of Traditional Chinese Medicine, Shanghai, China. Email: xdw021@163.com; mgzhang200909@163.com.

Background: Epithelial cell adhesion molecule (EpCAM) is one of the most commonly used markers of cancer stem cells (CSCs). However, the diagnostic and prognostic significance of EpCAM in lung cancer remains largely undetermined. In the present study, we systematically summarized and elucidated the correlation between EpCAM overexpression and lung cancer through a meta-analysis.

Methods: Six databases (PubMed, Web of Science, Cochrane Library, and Embase, CnKI and Wanfang Database) were systematically searched. Quality Assessment of Diagnostic Accuracy Studies-2 (QUADAS-2) criteria were adopted to assess the qualities of the included studies. Relevant data were extracted for meta-analysis using the Stata12.0 software. Unadjusted mixed odds ratios (ORs) or hazard ratios (HRs) with 95% confidence interval (95% CI) were estimated to evaluate the correlation between EpCAM overexpression and lung cancer. The sensitivity and specificity of the included studies were used to construct the summary receiver operator characteristic (SROC) curve and calculate the area under the SROC curve (AUC).

Results: A total of 14 studies consisting of 2,658 lung cancer patients were included following the PICOS principle. We found that the EpCAM expression was significantly higher in lung cancer patients compared with normal controls, including patients with benign pulmonary diseases (OR =63.71, 95% CI, 14.59–278.21, P=0.003) and healthy individuals (OR =520.08, 95% CI, 16.38–16,510.80, P=0.002), and its overexpression was negatively associated with the TNM stage (III + IV) (OR =0.41, 95% CI, 0.21–0.82, P=0.073. The combined sensitivity and specificity of EpCAM overexpression in the diagnosis of lung cancer were 0.79 (95% CI, 0.59–0.90) and 0.98 (95% CI, 0.95–0.99), respectively, and the SROC-AUC was 0.98 (95% CI, 0.97–0.99). Multivariate analysis of 322 lung cancer patients showed that there was no significant correlation between the EpCAM overexpression and prognosis of lung cancer (HR =2.28, 95% CI, 0.80–6.51, P=0.002). Deeks’ funnel plot analysis showed the existence of publication bias (P=0.000).

Conclusions: Our present findings suggested that EpCAM overexpression was not sensitive enough to predict the prognosis of lung cancer. Moreover, it was also a potential diagnostic indicator for lung cancer and correlated with TNM staging of lung cancer.

Keywords: Epithelial adhesion molecule (EpCAM); lung cancer; biomarker; meta-analysis

doi: 10.21037/apm-20-2013
View this article at: http://dx.doi.org/10.21037/apm-20-2013

^ ORCID: Tingting Zhu, 0000-0002-6110-7982; Xiaolin Peng, 0000-0003-3505-0289; Minguang Zhang, 0000-0002-1554-9472.
Introduction

Lung cancer ranks the second most common cancer worldwide, leading to 1.8 million deaths each year. Lung cancer patients have a low survival rate (4–17%) and poor prognosis, seriously threatening human health (1-3). Although rapid advances have been made in biology, early diagnosis, treatment, and mechanisms of drug resistance, lung cancer patients are usually diagnosed at late stages with a poor prognosis (4,5). The postoperative survival rate of patients is significantly associated with the staging of lung cancer, and the 5-year survival rate for lung cancer patients with stage IA and II is 90% and 60%, respectively (6). However, once the chance of surgical resection is lost, patients with advanced cancer are associated with a poor prognosis (7). Therefore, early screening, diagnosis, and treatment are of great importance to improve the survival rate and prolong the life span of patients with lung cancer. However, only very few lung cancer biomarkers with high sensitivity and specificity have been identified. Therefore, it is significantly necessary to determine more specific biomarkers, which can benefit the whole diagnosis, treatment, and prognosis of lung cancer patients.

With the deepening of research and the progress of detection technology, lung cancer autoantibodies, exosomes, circulating tumor cells (CTC), circulating tumor DNA (ctDNA), DNA methylation and other biomarkers have been identified (8). Epithelial cell adhesion molecule (EpCAM) is a common marker for CTC enrichment. EpCAM, also known as TACSTD1, EGP-2, and KS1/4, is a transdermal glycoprotein-coding gene located on chromosome 2P21 with a relative molecular weight of 40,000 (9). EpCAM is first identified in colon cancer and expressed in healthy human epithelial cells. EpCAM has many biological functions, which can accelerate cell cycle, and promote cell proliferation, differentiation, migration, and immune escape (10). Recent studies have shown that EpCAM can be used as a screening marker for cancer stem cells (CSCs) due to its unique characteristics, such as self-renewal, strong tumorigenicity, and genetic stability (11). The failure of traditional cancer therapies can largely be attributed to the ability of CSCs to evade cancer drugs and develop drug resistance. As EpCAM is highly expressed on the surface of CSCs but not in mature tumor cells, it is expected to be a new target for early diagnosis of liver cancer and novel anticancer therapy (12). Besides, studies have shown that patients with more than two circulating EpCAM-positive tumor cells in their peripheral blood have a significantly increased risk of postoperative tumor recurrence (13,14). Moreover, previous reports have also indicated that EpCAM is highly expressed in breast cancer, ovarian cancer, head and neck squamous cell carcinoma, and other cancers (15-17). Therefore, it is necessary to further study the function of EpCAM as a promising tumor marker (18).

However, it remains largely unclear whether EpCAM can be used as a marker of lung cancer. For example, Zhu et al. (19) and Zhou et al. (20) have suggested that EpCAM plays an important role in both diagnosis and prognosis of lung cancer, while Mao et al. (21) have shown that EpCAM only has diagnostic value for lung cancer. There are several unclear questions. Does EpCAM have different effects on different developmental stages of diverse types of tumors? Do the different prognoses of the same tumor patients indicate that the biological role of EpCAM is two-sided? Is EpCAM specific and sensitive in the diagnosis of lung cancer? In the present study, we evaluated whether EpCAM overexpression was related to the clinicopathological features of lung cancer. Moreover, we also evaluated the diagnostic and prognostic value of EpCAM in lung cancer patients by meta-analysis.

We present the following article in accordance with the PRISMA reporting checklist (available at http://dx.doi.org/10.21037/apm-20-2013) (22).

Methods

Literature and search strategy

Studies on the correlation between EpCAM and lung cancer published as of December 31, 2020 were searched from four English databases (PubMed, Web of Science Cochrane Library, Cochrane Library, and EMBASE) and two Chinese databases (CnKI and Wanfang Database). PICOS principles (23) of the meta-analysis were strictly followed for retrieval: P: pathologically diagnosed lung cancer patients; I: EpCAM positive; C: healthy individuals or patients with benign lung diseases; O: studies exploring the diagnostic and prognostic value of EpCAM overexpression for lung cancer patients; and S: clinical experimental diagnostic study and cohort study. The keywords searched were “EpCAM”, “CD326”, “lung”, “pulmonary”, “cancer”, “tumor”, “carcinoma”, and “neoplasm”. To obtain further potential studies, the references listed in the articles were also screened.

Eligibility criteria

The inclusion criteria in this meta-analysis were set as
follows: (I) clinical experimental diagnostic study or cohort studies investigating the correlation between EpCAM expression detected by immunohistochemistry (IHC) or reverse-transcription polymerase chain reaction (RT-PCR) and pathologically confirmed lung cancer patients; and (II) the odds ratios (ORs) and 95% confidence intervals (CIs) were reported.

The exclusion criteria were set as follows: (I) the original literature data were incomplete, HR and 95% CI data could not be obtained through direct or indirect methods, and the relevant pathological characteristics of lung cancer could not be obtained; (II) similar or identical repeated reports, old literature, and incomplete literature; for those repeatedly published studies, the literature with the most integrated materials and the latest data should be taken; (III) studies with glaring errors in the writing; and (IV) studies on only cell lines or animal models.

Data extraction and quality assessment

The data were extracted from the included studies by two independent reviewers (Zhu and Peng), and any dispute was resolved by consulting the third reviewer (Zhang). To obtain more detailed information, the demographic data, necessary clinical features, and clinical results of the patients included in the literature were fully extracted. The quality of the included studies was assessed using the Diagnostic Accuracy Study Quality Assessment-2 (QUADAS-2) (24) standard, which mainly evaluated four bias risks and three clinical applications, including case selection, implementation of diagnostic tests, application of the gold standard, and case processing and progression. A judgment of “low risk” (low bias or good applicability), “high risk” (high bias or poor applicability), or “unclear” (lack of relevant information or uncertainty of bias) was made accordingly for each entry.

Data synthesis and statistical analysis

STATA (version 12.0, Stata Corporation, USA) was adopted to analyze data. Firstly, the Spearman correlation coefficient was used to assess the threshold effect, and P<0.05 was considered statistically significant. The non-threshold effect was evaluated by Cochran’s Q test and I² test, and the significant difference was set as P<0.01 or I²>50% (25). When there was significant heterogeneity between studies, the random-effects model was used to combine the statistics. Otherwise, the fixed-effects model was used to carry out the weighted combination of effects (26). The sources of heterogeneity were investigated by sensitivity analysis. Subsequently, the pooled sensitivity and specificity, as well as the summary receiver operator characteristic (SROC) curve (AUC) values, were calculated and constructed according to the bivariate analysis to evaluate the performance of the diagnostic capacity of EpCAM expression for lung cancer in this study. Meanwhile, the relationship between EpCAM overexpression and lung cancer prognosis was calculated by HRs (unadjusted) with the corresponding 95% CIs. Clinicopathological conditions and overall survival (OS) rate of lung cancer was used to estimate 95% CIs and ORs. Finally, the publication bias was assessed by Deek’s funnel plot using the effect quantity DOR as the x-coordinate, and the effective sample content as the y-coordinate. P<0.05 indicated a publication bias.

Results

Study selection and features

Finally, we preliminarily retrieved 2,294 records from the above-mentioned databases. Then 535 duplicate studies were excluded, and the remaining 1,759 studies were used for further evaluation. After reading the title and abstract, 75 references met the inclusion criteria. After full-text evaluation, 14 studies (19-21,27-37) consisting of 2,658 lung cancer patients and published between 2006 and 2020 were included in the final analysis (Figure 1). Only three studies reported the OS data or disease-free survival (DFS) information for lung cancer patients. Of the 14 studies, there were two from South Korea, one from Japan, one from Poland, and one from Switzerland, and the remaining nine studies were all from China. Six studies were published in Chinese, and the others were published in English (Table 1). We found that most of the included studies in this meta-analysis had four or more of the seven items corresponding to “Yes” in QUADAS-2, indicating that the overall quality of the included studies was generally high (Figures 2,3).

Heterogeneity analysis

As shown in the Bivariate boxplot (Figure 4), although most of the studies fell in the middle region, there are still 3 studies that fell outside the region, suggesting that there may be heterogeneity among studies. In addition, the results of specificity and sensitivity analysis (Figure 5)
Reports identified through PubMed [326], Cochrane Library [13], Web of Science [943], EMBASE database [853], Chinese CNKI [48] and the Chinese Wan Fang database [111]
N=2,294

Additional reports identified through other sources
N=0

Reports after duplicates removed
N=1,759

Primary selection through browsing the retrieved titles and abstracts

Studies excluded non-gastric related studies; non-human related studies; non-original articles (review, letter)
N=1,684

Researches retrieved for detailed evaluation. Secondary selection through reading the full texts of potentially eligible articles
N=75

Studies excluded without sufficient data
N=61

Literature included in this systematic review
N=14
(8 in English and 6 in Chinese)

Figure 1 Flow diagram of study selection in this meta-analysis.

Table 1 Basic characteristics of the publications included in this study

<table>
<thead>
<tr>
<th>Studies</th>
<th>Year</th>
<th>Country</th>
<th>Cases (n)</th>
<th>Sex (M/F)</th>
<th>Age</th>
<th>Sample</th>
<th>EpCAM overexpression rate (%)</th>
<th>Control groups</th>
<th>OS, DFS</th>
<th>Reference number</th>
</tr>
</thead>
<tbody>
<tr>
<td>Went</td>
<td>2006</td>
<td>Switzerland</td>
<td>1,287</td>
<td>826/201</td>
<td>62 [29–87]</td>
<td>Tissue</td>
<td>823 (64.0%)</td>
<td>NS</td>
<td>NS</td>
<td>27</td>
</tr>
<tr>
<td>Kim</td>
<td>2009</td>
<td>Korea</td>
<td>234</td>
<td>127/107</td>
<td>61 [13–82]</td>
<td>Tissue</td>
<td>64 (27.3%)</td>
<td>NS</td>
<td>NS</td>
<td>28</td>
</tr>
<tr>
<td>Kobayashi</td>
<td>2010</td>
<td>Japan</td>
<td>130</td>
<td>62/68</td>
<td>60 [38–82]</td>
<td>Tissue</td>
<td>87 (67.0%)</td>
<td>NS</td>
<td>OS</td>
<td>29</td>
</tr>
<tr>
<td>Pak</td>
<td>2012</td>
<td>Korea</td>
<td>164</td>
<td>113/51</td>
<td>63 [42–81]</td>
<td>Tissue</td>
<td>94 (57.3%)</td>
<td>NS</td>
<td>NS</td>
<td>30</td>
</tr>
<tr>
<td>Zhu</td>
<td>2013</td>
<td>China</td>
<td>45</td>
<td>28/17</td>
<td>63 [42–88]</td>
<td>Blood</td>
<td>22 (48.9%)</td>
<td>Yes</td>
<td>NS</td>
<td>31</td>
</tr>
<tr>
<td>Sun</td>
<td>2014</td>
<td>China</td>
<td>51</td>
<td>30/21</td>
<td>65 [47–84]</td>
<td>Pleural Effusion</td>
<td>43 (84.3%)</td>
<td>Yes</td>
<td>NS</td>
<td>32</td>
</tr>
<tr>
<td>Skirecki</td>
<td>2014</td>
<td>Poland</td>
<td>41</td>
<td>27/14</td>
<td>67 [49–82]</td>
<td>Blood</td>
<td>25 (60.0%)</td>
<td>Yes</td>
<td>NS</td>
<td>33</td>
</tr>
<tr>
<td>Zhu</td>
<td>2014</td>
<td>China</td>
<td>74</td>
<td>49/25</td>
<td>63 [41–79]</td>
<td>Blood</td>
<td>34 (45.9%)</td>
<td>Yes</td>
<td>OS, DFS</td>
<td>19</td>
</tr>
<tr>
<td>Zhen</td>
<td>2015</td>
<td>China</td>
<td>200</td>
<td>112/88</td>
<td>63 [29–70]</td>
<td>Blood</td>
<td>195 (97.5%)</td>
<td>Yes</td>
<td>NS</td>
<td>34</td>
</tr>
<tr>
<td>Mao</td>
<td>2016</td>
<td>China</td>
<td>64</td>
<td>44/20</td>
<td>NS</td>
<td>Tissue</td>
<td>38 (59.4%)</td>
<td>NS</td>
<td>NS</td>
<td>21</td>
</tr>
<tr>
<td>Zhou</td>
<td>2016</td>
<td>China</td>
<td>118</td>
<td>90/28</td>
<td>NS</td>
<td>Tissue</td>
<td>67 (56.8%)</td>
<td>Yes</td>
<td>OS</td>
<td>20</td>
</tr>
<tr>
<td>Cheng</td>
<td>2018</td>
<td>China</td>
<td>100</td>
<td>55/45</td>
<td>59 [39–75]</td>
<td>Blood</td>
<td>84 (84.0%)</td>
<td>Yes</td>
<td>NS</td>
<td>35</td>
</tr>
<tr>
<td>Wu</td>
<td>2019</td>
<td>China</td>
<td>126</td>
<td>74/52</td>
<td>53 [32–73]</td>
<td>Tissue</td>
<td>105 (83.3%)</td>
<td>NS</td>
<td>NS</td>
<td>36</td>
</tr>
<tr>
<td>He</td>
<td>2020</td>
<td>China</td>
<td>24</td>
<td>14/10</td>
<td>Mean 61.4</td>
<td>Blood</td>
<td>15 (62.5%)</td>
<td>Yes</td>
<td>NS</td>
<td>37</td>
</tr>
</tbody>
</table>

EpCAM, epithelial cell adhesion molecule; NS, not specified; OS, overall survival; DFS, disease-free survival.
also showed significant heterogeneity among studies. The source of heterogeneity was analyzed. First, the ROC curve (Figure 6) was drawn to determine that the study was not "shoulder-arm" distribution. Meanwhile, the results of the heterogeneity test showed that the Spearman correlation coefficient of the overall combined effect was $P=0.079$ of the overall combined effect, suggesting that there was no threshold effect among studies, and heterogeneity among studies mainly caused by the non-threshold effect. Heterogeneity may be caused by the publication time of the literature, research groups, individual characteristics of the research objects and other factors. Therefore, the random effects model was used for statistical analysis.

**Figure 2** Diagram of the risk of bias and applicability of the included literature.

**Figure 3** Summary diagram of the risk of bias and applicability of the included literature.

**Figure 4** Bivariate boxplot of heterogeneity analysis.

**Correlation between EpCAM overexpression and lung cancer**

Cut-off value of EpCAM positive expression is 4 (immunohistochemical positive intensity score) or 0.34 (relative mRNA expression level). Overall OR of 588 lung cancer patients and 185 patients with benign pulmonary lesions in six studies showed that the frequency of EpCAM overexpression in benign pulmonary lesions was significantly lower compared with the lung cancer (OR $=63.71$, 95% CI, 14.59–278.21, $P=0.003$) (Figure 7).
Three studies consisting of 265 lung cancer patients and 487 healthy controls showed that the frequency of EpCAM overexpression in healthy controls was significantly lower compared with the lung cancer group (OR =520.08, 95% CI, 16.38–16,510.80, P=0.002) (Figure 7).

**Diagnostic capacity of EpCAM overexpression in lung cancer**

The combined sensitivity and specificity of EpCAM overexpression in the diagnosis of lung cancer were 0.79 (95% CI, 0.59–0.90), and 0.98 (95% CI, 0.95–0.99), respectively. The SROC-AUC was calculated as 0.98 (95% CI, 0.97–0.99), indicating that EpCAM was oversensitive and specific for the diagnosis of lung cancer (Figure 6).

**Prognostic impact of EpCAM overexpression on lung cancer patients**

Only three studies reported EpCAM overexpression-related OS and (or) DFS in lung cancer patients. Multivariate analysis of 322 lung cancer patients showed that there was no significant correlation between EpCAM overexpression and prognosis of lung cancer (HR =2.28, 95% CI, 0.80–6.51, P=0.002) (Figure 8).

**Correlation between EpCAM overexpression and clinical variables for lung cancer**

We estimated the correlation between EpCAM overexpression and clinical variables of lung cancer in our
Figure 7 Forest plot showing the relationship between EpCAM overexpression and lung cancer in cancer vs. controls, lung cancer vs. normal controls. OR, odds ratio; CI, confidence interval.

Figure 8 Forest plot showing the prognostic impact of EpCAM overexpression in lung cancer patients.

meta-analysis. EpCAM overexpression was only significantly related to TNM stage (III+IV versus I+II: OR = 0.41, 95% CI, 0.21–0.82, P = 0.073 (Figure 9). However, EpCAM overexpression had no relationship with gender (man versus women), age of patients (<60 group versus ≥60 group), tumor size (≤3 cm versus >3 cm), T stage (T3-T4 versus T1-T2), differentiation (poor versus well and moderate differentiation), lymph node metastasis (NO versus YES), and distant metastasis (NO versus YES) (Figure 9).

Publication bias and sensitivity analysis

In the present study, the reciprocal of the standard error of diagnostic odds ratio (DOR) was used as the vertical coordinate, and the funnel plot was used as the abscissa to evaluate publication bias. According to the Deeks’ funnel plot results, the linear regression test P = 0.000-0.10, indicating publication bias existed in the included studies (Figure 10). Sensitivity analysis results showed little changes, all of which were within 95% CI of the original summary.
Figure 9 Pooled analysis for the relationship between EpCAM overexpression and clinicopathological features. (A) Gender (man versus women); (B) age of patients (<60 groups versus ≥60 groups); (C) tumour size (≤3 versus >3 cm); (D) T stage (T3-T4 versus T1-T2); (E) TNM stage I+IV versus I+II; (F) poor versus well and moderate differentiation; (G) lymph node metastasis (NO versus YES); (H) distant metastasis (no versus yes); OR, odds ratio; CI, confidence interval.
analysis results, suggesting that the meta-analysis summary results of this study were reliable and stable (Figure 11A,B).

**Discussion**

At present, no study has comprehensively and systematically assessed the relationship between EpCAM and lung cancer using a meta-analysis. Based on our results, we found three reliable conclusions regarding the correlation between EpCAM overexpression and lung cancer. Firstly, EpCAM overexpression had diagnostic significance for lung cancer. Besides, EpCAM overexpression had no significant relationship with the prognosis of lung cancer. Finally, patients with a higher TNM stage of lung cancer had a higher frequency of EpCAM overexpression.

The vast majority of lung cancer deaths result from metastasis and spread of cancer cells. Significantly, control of metastasis can improve the survival and quality of life of lung cancer patients. Therefore, it has been a hot research topic to identify genes that affect lung cancer metastasis. Initially, EpCAM is identified as a type of normal epithelial adhesion molecules of the alveolar epithelium and involved in cell-to-cell contact. Many studies have focused on the relationship between EpCAM and tumor, suggesting that EpCAM overexpression can up-regulate oncogenes, C-Myc, and Cyclin A and E, accelerate cell cycle, and enhance cell proliferation (38). Overexpression of EpCAM can promote the proliferation and invasiveness of tumor cells. In contrast, down-regulation of EpCAM can inhibit the proliferation and invasion of tumor cells (39). EpCAM is a target gene of the wnt/β-catenin signaling pathway, which is considered a critical marker of tumor cells with the characteristics of stem cells (40). Besides, *in vitro* studies have also shown that EpCAM is a downstream target gene of transfer-related protein 1, and overexpression of transfer-related protein-1 can increase the expression of EpCAM, leading to enhanced invasion and metastasis of lung cancer (41). Our findings demonstrated that EpCAM overexpression exhibited positive diagnostic capacity with high sensitivity and excellent specificity, and it could be used as a promising biomarker for lung cancer. Before our study, there are two different opinions on the correlation between EpCAM and prognosis of lung cancer. Some studies (29) have indicated that EpCAM is significantly correlated with a good forecast of lung cancer. In contrast, others (28,30) have reported that EpCAM overexpression has no prognostic value for lung cancer, and our results supported the latter. However, due to the lack of studies on the relationship between EpCAM and lung cancer prognosis, more studies are expected.

There was no significant correlation between EpCAM overexpression and characteristics of lung cancer patients.

![Graph of Deeks’ funnel plot asymmetry test.](image)

**Figure 10** Graph of Deeks’ funnel plot asymmetry test.

![Sensitivity analysis. (A) Cancer vs. normals. (B) Cancer vs. Benign.](image)

**Figure 11** Sensitivity analysis. (A) Cancer vs. normals. (B) Cancer vs. Benign.
in terms of age, sex, tumor size, degree of differentiation, tumor stage, lymph node metastasis, and distant metastasis, while its expression level had a significant correlation with the TNM stage. In the early stage of lung cancer, EpCAM mRNA is involved in the proliferation regulation of lung cancer cells. With the progression of lung cancer, the expression of EpCAM is increased. Besides, EpCAM protein is mainly located on the cell membrane of cancer cells and occasionally dispersed in plasma. Therefore, in our present study, high EpCAM expression was significantly correlated with TNM stage (III+IV) (III+IV versus I+II: OR =0.41, 95% CI, 0.21–0.82, P=0.073, Figure 9E). Different from previous reports, we found that there was no correlation between EpCAM overexpression and tumor size, degree of differentiation, tumor stage, lymph node metastasis, or distant metastasis. These differences might be attributed to our small sample size. Therefore, more studies with a large sample size are required for further verification.

**Limitations**

There were still four limitations in this study: firstly, our results indicate that there is significant evidence of publication bias in our study, which may be related to the inclusion of retrospective case-control studies and may increase selective bias. Secondly, this meta-analysis only included two Chinese databases and four English databases. Therefore, some studies in other databases and languages might be ignored. Then, different studies used different definite criteria for IHC staining. Finally, only three studies reporting OS data were analyzed in this meta-analysis. Due to the shortage and instability of the results, further large-scale and well-designed studies are required to confirm our results, which might provide a higher level of evidence.

**Conclusions**

Collectively, our results suggested that EpCAM had diagnostic value but not prognostic value for lung cancer patients. We also found that TNM staging was correlated with the EpCAM expression in patients with lung cancer. However, further experimental and clinical studies are still required to elucidate the potential biological mechanisms of EpCAM in the diagnosis, prognosis, and treatment of lung cancer.

**Acknowledgments**

Special thanks go to Richard Liang, PhD for his help in language editing.

**Funding:** This work was supported by Natural Science Foundation of Shanghai (Research Grant no.19ZR1452400) and National Natural Science Foundation of China (Research Grant no. 81673743).

**Footnote**

**Reporting Checklist:** The authors have completed the PRISMA reporting checklist. Available at [http://dx.doi.org/10.21037/apm-20-2013](http://dx.doi.org/10.21037/apm-20-2013)

**Conflicts of Interest:** All authors have completed the ICMJE uniform disclosure form (available at [http://dx.doi.org/10.21037/apm-20-2013](http://dx.doi.org/10.21037/apm-20-2013)). The authors have no conflicts of interest to declare.

**Ethical Statement:** The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved.

**Open Access Statement:** This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: [https://creativecommons.org/licenses/by-nc-nd/4.0/](https://creativecommons.org/licenses/by-nc-nd/4.0/).

**References**


