Diagnostic accuracy of interferon gamma-induced protein 10 for tuberculosis: a meta-analysis

Shu-Jin Guo*, Liu-Qun Jia*, Qian-Jing Hu, Hong-Yu Long, Cai-Shuang Pang, Fu-Qiang Wen

Division of Pulmonary Diseases, State Key Laboratory of Biotherapy of China and Department of Respiratory Medicine, West China Hospital of Sichuan University, Chengdu 610041, China. *Equal contributors.

Received October 20, 2013; Accepted November 21, 2013; Epub January 15, 2014; Published January 30, 2014

Abstract: The diagnostic accuracy of tuberculosis (TB) remains a clinical challenge, and a number of studies have used the interferon gamma-induced protein 10 (IP-10) in the diagnosis of TB. The aim of the present meta-analysis was to determine the overall accuracy of IP-10 in the diagnosis of TB. A systematic review of studies published in English from Medline, Embase and Cochrane Library was conducted and the data concerning the accuracy of IP-10 in the diagnosis of TB were pooled. The methodological quality of each study was assessed by QUADAS (quality assessment for studies of diagnostic accuracy). Statistical analysis was performed by employing Meta-Disc 1.4 soft-ware and STATA. The overall test performance was summarized using receiver operating characteristic curves. 14 studies, based on 2075 subjects, met the inclusion criteria. The summary estimates for IP-10 in the diagnosis of TB were: sensitivity 0.73 (95% CI, 0.71-0.76), specificity 0.83 (95% CI, 0.81-0.86), positive likelihood ratio 7.08 (95% CI, 3.94-12.72), negative likelihood ratio 0.26 (95% CI, 0.20-0.35) and diagnostic odds ratio 29.50 (95% CI, 14.43-60.30), and the area under the curve was 0.88. Our findings suggest that IP-10 may improve the accuracy of TB diagnosis, while the results of IP-10 assays should be interpreted in parallel with conventional test results and other clinical findings.

Keywords: Interferon gamma-induced protein 10 (IP-10), tuberculosis, diagnostic, meta-analysis

Introduction

Tuberculosis is a life-threatening disease with the annual death rate more than 2 millions, especially in the areas short of medical and health resources [1, 2]. TB remains a highly morbidity and mortality epidemic, making it crucial for early diagnosis and treatment. The diagnosis of TB is difficult owing to diverse clinical presentations combined with paucibacillary infection, making bacteriological confirmation challenging. Biopsy is invasive and it’s hard for histological diagnosis [3, 4]. For decades there has been little effort to improve techniques for diagnosing tuberculosis. A number of biomarkers have been studied in attempts to improve the accuracy of TB diagnosis but failed to identify a reliable marker with both high sensitivity and specificity. Therefore, it is imperative to identify a novel marker to increase diagnostic accuracy.

IP-10 is a member of the CXC family [5]. IP-10 is a proinflammatory chemokine which are expressed in inflamed tissues by resident and infiltrated cells (primarily monocyte/macrophages) after paracrine stimulation from T-cells by IFNs and other proinflammatory cytokines. Current knowledge showed that IP-10 and its homologues were involved in inflammatory lung injury, and were closely related to TB [6, 7]. An increasing number of studies consider IP-10 to be a marker for the diagnosis of TB [6-9]. However, conflicting results have been reported and the exact role of IP-10 remains unclear. Therefore, we performed the present meta-analysis to establish the overall accuracy of IP-10 for diagnosing TB.

Method

Date source and search strategy

Two investigators independently performed a systematic electronic search of the Pubmed and Embase databases until 1 March 2013 to identify potentially relevant articles. The Cochrane Library database was also searched.
for review and meta-analysis. The following search terms were used: “interferon gamma-induced protein 10” or “IP-10” and “tuberculosis” or “TB”. The restrictions languages were English and Chinese. We reviewed the bibliographies of all selection articles to identify additional relevant studies.

Selection of studies

Two reviewers independently screened titles and abstracts of all studies for relevance. Disagreements were resolved by a third opinion. The strength of the individual studies was weighed for relevance, based on following items: 1. The clinical domains should include patients with suspected tuberculosis; 2. The reference diagnostic standards were clearly described and all specimens were diagnosed by using the reference standards; 3. Completeness of data (numbers of true-positive, false-positive, true-negative, false negative) were reported, to allow reconstruction of the diagnostic 2 by 2 table; 4. The studies were written in English and Chinese.

Methods appraisal and data extraction

The final set of articles was assessed independently by two reviewers. The retrieved data included author, publication year, the number of included specimens (true-positive, false-positive, true-negative, false negative), sensitivity and specificity. The methodological quality of included studies was evaluated using the Quality Assessment for Studies of Diagnostic Accuracy (QUADAS) [10] Tool. This is an evidence-based approach to quality assessment intended for use in systematic reviews of diagnostic accuracy studies. A quality index is generated, with a maximum value of 14.

Statistical analysis

We used standard methods recommended for meta-analyses of diagnostic test evaluations. Analyses were performed using Meta-disc 1.4 [11] and Stata Version 12 software [12]. Sensitivity; specificity; PLR (positive likelihood ratio); NLR (negative likelihood ratio); and DOR (diagnostic odds ratio) were computed for each study. The analysis was based on a summary ROC (SROC) curve. The sensitivity and specificity for the single test threshold identified for each study were used to plot an SROC curve. Q test was used to determine whether there was heterogeneity and $I^2$ to estimate the degree of heterogeneity. According to the result of heterogeneity analysis, the appropriate statistical analysis model for meta-analysis was chosen.

From the studies included, we extracted the numbers of patients with a true-positive, false-positive, true-negative, false negative test result either directly or through recalculation based on reported measures of accuracy in combination with the incidence and specimen size of the study. Sensitivity, specificity and diagnostic odds ratios (DOR) together with 95% CI (confidence interval) were calculated for each study based on the reconstructive 2 by 2 table. We plotted all results from the included studies on a receiver operating characteristic (ROC) plot of sensitivity against specificity, with the specificity axis reversed. In addition, area under the summary ROC curve (AUC)-ROC values were determined.

Since publication bias is of concern for meta-analyses of diagnostic studies, the publication bias of included studies was assessed by using Deeks test, which was analyzed by using Stata Version 12 software.

Result

After we evaluated these citations and the bibliographies of the potential studies, 14 unique studies [13-26] were eventually included in our meta-analysis. The main reasons of excluding studies were as follows: the study was a duplicate between the Pubmed and Embase database, the study was not diagnostic, or the study cannot reconstruct the diagnostic 2 by 2 table. The study characteristics, along with QUADAS scores, were shown in Table 1.

Study characteristics and quality assessment

Overall, the selected 14 case-control studies included 2075 cases, in which 1171 cases were tuberculosis, 904 cases were non-tuberculosis. The gold standard diagnosis method, TB culture or smear positive, was used to diagnose TB. In several of the included studies, the diagnosis standard of TB was combined the gold standard with clinical data. The patients included pulmonary TB and extrapulmonary TB, and the specimen included blood and pleural effusion. The quality of the 14 studies was generally high, satisfying the majority of the criteria.
Interferon gamma-induced protein 10 for tuberculosis

Diagnostic accuracy in TB

The forest plot of sensitivity and specificity for IP-10 in diagnosing TB was shown in Figures 1 and 2. The heterogeneity analysis showed I^2 of 88.7% for sensitivity and 92.9% for specificity, represented a high heterogeneity, thus the random effects model approach was selected in this study. The overall pooled sensitivity was 0.73 (95% CI, 0.71-0.76), and pooled specificity was 0.83 (95% CI, 0.81-0.86). We also noted that PLR was 7.08 (95% CI, 3.94 to 12.72), NLR was 0.26 (95% CI, 0.20 to 0.35), and the DOR was 29.50 (95% CI, 14.43-60.30). The SROC (Figure 3) curve presents a global summary of test performance, and shows the tradeoff between sensitivity and specificity. The sensitivity, specificity and 95% confidence region (precision of estimation of pooled sensitivity and specificity) of 14 studies were showed in a summary ROC curve (pooled sensitivity against 1-(pooled specificity)). As a global measure of

Table 1. Summary of the studies included in the meta-analysis

<table>
<thead>
<tr>
<th>Author</th>
<th>Year</th>
<th>Country</th>
<th>TB/NTB</th>
<th>Cut-off (pg/ml)</th>
<th>Gold Standard</th>
<th>Specimen</th>
<th>test method</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>QUA-DAS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Okamoto</td>
<td>2005</td>
<td>Japan</td>
<td>11/32</td>
<td>7620</td>
<td>Bacteriology/Histology</td>
<td>PE</td>
<td>ELISA</td>
<td>8</td>
<td>1</td>
<td>3</td>
<td>31</td>
<td>10</td>
</tr>
<tr>
<td>Ruhwald</td>
<td>2008</td>
<td>Denmark</td>
<td>74/124</td>
<td>673</td>
<td>Bacteriology</td>
<td>blood</td>
<td>xMAP</td>
<td>59</td>
<td>3</td>
<td>15</td>
<td>121</td>
<td>10</td>
</tr>
<tr>
<td>Ruhwald</td>
<td>2008</td>
<td>Nigeria</td>
<td>59/23</td>
<td>635</td>
<td>Bacteriology</td>
<td>blood</td>
<td>xMAP</td>
<td>41</td>
<td>3</td>
<td>18</td>
<td>20</td>
<td>10</td>
</tr>
<tr>
<td>Supriya</td>
<td>2008</td>
<td>India</td>
<td>38/24</td>
<td>84173</td>
<td>Bacteriology/Imaging/clinical</td>
<td>PE</td>
<td>ELISA</td>
<td>29</td>
<td>1</td>
<td>9</td>
<td>23</td>
<td>11</td>
</tr>
<tr>
<td>Dheda</td>
<td>2009</td>
<td>South Africa</td>
<td>55/19</td>
<td>28170</td>
<td>Bacteriology/Histology/clinical</td>
<td>PE</td>
<td>ELISA</td>
<td>44</td>
<td>3</td>
<td>11</td>
<td>16</td>
<td>12</td>
</tr>
<tr>
<td>Goletti</td>
<td>2010</td>
<td>India</td>
<td>28/38</td>
<td>350</td>
<td>Bacteriology</td>
<td>plasma</td>
<td>ELISA</td>
<td>21</td>
<td>16</td>
<td>7</td>
<td>22</td>
<td>10</td>
</tr>
<tr>
<td>Kabeer</td>
<td>2010</td>
<td>India</td>
<td>173/100</td>
<td>300</td>
<td>Bacteriology</td>
<td>plasma</td>
<td>EIA</td>
<td>160</td>
<td>52</td>
<td>13</td>
<td>48</td>
<td>14</td>
</tr>
<tr>
<td>Kellar</td>
<td>2011</td>
<td>American</td>
<td>12/12</td>
<td>460</td>
<td>Bacteriology</td>
<td>plasma</td>
<td>MMA</td>
<td>12</td>
<td>0</td>
<td>0</td>
<td>12</td>
<td>11</td>
</tr>
<tr>
<td>Sutherland</td>
<td>2012</td>
<td>Gambia</td>
<td>30/11</td>
<td>36695</td>
<td>Bacteriology/Histology/clinical</td>
<td>PE</td>
<td>ELISA</td>
<td>26</td>
<td>2</td>
<td>4</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Kabeer</td>
<td>2012</td>
<td>India</td>
<td>200/186</td>
<td>300</td>
<td>Bacteriology and clinic</td>
<td>blood</td>
<td>EIA</td>
<td>136</td>
<td>53</td>
<td>64</td>
<td>133</td>
<td>10</td>
</tr>
<tr>
<td>Vanini</td>
<td>2012</td>
<td>Italy</td>
<td>37/40</td>
<td>1096</td>
<td>Bacteriology</td>
<td>blood</td>
<td>EIA</td>
<td>28</td>
<td>1</td>
<td>9</td>
<td>39</td>
<td>13</td>
</tr>
<tr>
<td>Wang H</td>
<td>2012</td>
<td>China</td>
<td>78/44</td>
<td>44</td>
<td>Bacteriology/Histology</td>
<td>PE</td>
<td>ELISA</td>
<td>65</td>
<td>6</td>
<td>13</td>
<td>38</td>
<td>10</td>
</tr>
<tr>
<td>Aabye</td>
<td>2013</td>
<td>Denmark</td>
<td>72/97</td>
<td>1500</td>
<td>Bacteriology</td>
<td>plasma</td>
<td>ELISA</td>
<td>61</td>
<td>2</td>
<td>11</td>
<td>95</td>
<td>12</td>
</tr>
<tr>
<td>Mohammed</td>
<td>2013</td>
<td>UK</td>
<td>304/154</td>
<td>3022</td>
<td>Bacteriology</td>
<td>plasma</td>
<td>ELISA</td>
<td>167</td>
<td>9</td>
<td>137</td>
<td>145</td>
<td>10</td>
</tr>
</tbody>
</table>


Figure 1. Forest plots of sensitivity for IP-10 in the diagnosis of TB.
Interferon gamma-induced protein 10 for tuberculosis

Figure 2. Forest plots of specificity for IP-10 in the diagnosis of TB.

Figure 3. Summary receiver operating characteristic (SROC) curve of IP-10 in the diagnosis of TB.
Interferon gamma-induced protein 10 for tuberculosis

Table 2. Sensitivity and specificity for IP-10 in subgroup analysis

<table>
<thead>
<tr>
<th></th>
<th>blood</th>
<th>p</th>
<th>I²</th>
<th>pleural effusion</th>
<th>p</th>
<th>I²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sensitivity</td>
<td>0.72 (0.69-0.74)</td>
<td>0.0000</td>
<td>92.4%</td>
<td>0.81 (0.75-0.86)</td>
<td>0.7436</td>
<td>0</td>
</tr>
<tr>
<td>specificity</td>
<td>0.82 (0.79-0.85)</td>
<td>0.0000</td>
<td>95.4%</td>
<td>0.92 (0.84-0.95)</td>
<td>0.2624</td>
<td>23.8%</td>
</tr>
<tr>
<td>PLR (95% CI)</td>
<td>6.71 (3.30-13.67)</td>
<td>0.0000</td>
<td>93.6%</td>
<td>7.01 (4.23-11.61)</td>
<td>0.5172</td>
<td>0</td>
</tr>
<tr>
<td>NLR (95% CI)</td>
<td>0.28 (0.20-0.40)</td>
<td>0.0000</td>
<td>86.3%</td>
<td>0.22 (0.16-0.30)</td>
<td>0.8962</td>
<td>0</td>
</tr>
<tr>
<td>DOR (95% CI)</td>
<td>27.34 (10.76-69.52)</td>
<td>0.0000</td>
<td>86.8%</td>
<td>34.77 (17.75-68.10)</td>
<td>0.8236</td>
<td>0</td>
</tr>
<tr>
<td>AUC</td>
<td>0.88</td>
<td>0.91</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

PLR, positive likelihood ratio; NLR, negative likelihood ratio; DOR, diagnostic odds ratio; AUC, the area under the summary receiver operating characteristic curve.

Figure 4. Linear regression test of funnel plot asymmetry.

test efficacy we used the Q-value, the intersection point of the SROC curve with a diagonal line from the left upper corner to the right lower corner of the ROC space, which corresponds to the highest common value of sensitivity and specificity for the test. This point does not indicate the only or even the best combination of sensitivity and specificity for a particular clinical setting but represents an overall measure of the discriminatory power of a test. Our data showed that the SROC curve is positioned near the desirable upper left corner of the SROC curve, and that the maximum joint sensitivity and specificity (Q value) was 0.82, while the area under the curve (AUC) was 0.88, indicating a moderate level of overall accuracy.

We also explored the diagnostic accuracy in different specimen, blood and pleural effusion. In the 14 included studies, the specimen of 9 were blood/plasma, the overall pooled sensitivity was 0.72 (95% CI, 0.69-0.74), pooled specificity was 0.82 (95% CI, 0.79-0.85), and PLR was 6.71 (95% CI, 3.30 to 13.67), NLR was
Interferon gamma-induced protein 10 for tuberculosis

0.28 (95% CI, 0.20 to 0.40). The specimen of 5 studies were pleural effusion, the overall pooled sensitivity was 0.81 (95% CI, 0.75-0.86), pooled specificity was 0.90 (95% CI, 0.84-0.95), and PLR was 7.01 (95% CI, 4.23 to 11.61), NLR was 0.22 (95% CI, 0.16 to 0.30). The results were showed in Table 2.

Publication bias

Deeks’ funnel plot asymmetry test was used to evaluate the final set of studies for potential publication bias. The slope coefficient was associated with a p value of 0.17, suggesting symmetry in the data and no publication bias (Figure 4).

Discussion

The results of these 14 studies analyzed by patients showed that IP-10 plays a role in diagnosing TB. Using the random-effects approach, we found a summary estimate of 73% for sensitivity and 83% for specificity, and the maximum joint sensitivity and specificity (Q value) was 0.82 while the AUC was 0.88, indicating a moderate level of overall accuracy. The DOR is a single indicator of test accuracy that combines the data from sensitivity and specificity into a single number. The DOR of a test is the ratio of the odds of positive test results in the patient with disease relative to the odds of positive test results in the patient without disease. The value of a DOR ranges from 0 to infinity, with higher values indicating higher accuracy. In this meta-analysis we found that the mean DOR was 29.50, indicating a moderate level of overall accuracy. Likelihood ratios are considered to be more clinically meaningful, and we also presented both PLR and NLR as our measures of diagnostic accuracy. Likelihood ratios of >10 or <0.1 generate large and often conclusive shifts from pretest to posttest probability (indicating high accuracy). The pooled PLR 7.08 suggests that patients with TB have an approximately 7-fold higher chance of being IP-10 positive compared with patients without TB. On the other hand, the pooled NLR 0.26 suggests that if the IP-10 test was negative, the probability that this patient has TB was 26 percent, which is not low enough to rule out TB. These data suggest that a negative IP-10 result should not be used alone to diagnosis TB.

In this meta-analysis, we found that the diagnostic accuracy was high in pleural effusion than in blood. However, the diagnostic accuracy was moderate in both kinds of specimen when used IP-10 alone. It’s reported that when IP-10 was combined with other test, the diagnostic accuracy increased [13, 16]. One of the study showed that when IP-10 combined with INF-y/ tuberculin skin test, the sensitivity increased significantly (91.0% vs. 96.5%/98.3%) [16]. Therefore, even if IP-10 cannot diagnosis TB alone, but also can be used as a powerful reference marker.

An exploration of the reasons for heterogeneity rather than the computation of a single summary measure is an important goal of meta-analysis. In the present study, QUADAS scores was used to assess the effect of study quality, we observe that the studies had high quality (QUADAS score of ≥10). The possible reasons of heterogeneity included the different cutoff values of the studies and the bias of the selected cases. In one of the subgroup analysis (analyzed by the specimen of pleural effusion) showed no heterogeneity, suggesting the different specimen of the included studies may also cause heterogeneity.

Some limitations should be considered when interpreting the results. Firstly, the sample sizes of several included studies are rather small and they do not have adequate ability to assess the diagnostic accuracy. Secondly, the included studies use different cutoff values, which may contribute to the heterogeneity. Thirdly, this meta-analysis limited to published studies that may miss some of the gray literature.

Conclusion

IP-10 plays a role in the diagnosis of TB, and the diagnostic accuracy was moderate. IP-10 can increase diagnostic accuracy when combined with other tests. The results of IP-10 should be interpreted in parallel with clinical findings and the results of other conventional tests.

Acknowledgements

This study was supported by grants 31171103 and 81230001 from the National Natural Science Foundation of China; and grant 06-834 from the China Medical Board of New York. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.
References


Interferon gamma-induced protein 10 for tuberculosis


