Performance of ESAT-6 for serodiagnosis of nonhuman primate tuberculosis: A meta-analysis

Fangui Min, Jing Wang and Yu Zhang*

Guangdong Laboratory Animals Monitoring Institute, Guangdong Provincial Key Laboratory of Laboratory Animals, Guangzhou 510663, Peoples Republic of China.
*Corresponding author’s e-mail: zhangyugzh@hotmail.com

ABSTRACT

ESAT-6 is one of the most studied antigens in vaccine, diagnosis, and pathogenic mechanism of tuberculosis. In the present study, a meta-analysis was performed regarding the use of ESAT-6 based antibody detection test for diagnosing nonhuman primate (NHP) tuberculosis. Studies in English and Chinese were searched and selected strictly. Quality of included studies was assessed using the standardized QUADAS-2 tool. Heterogeneity was explored through meta-regression. Finally, eight studies were included with high degree of homogeneity. Quality of included studies was generally satisfied except the bias of “patient selection” for the majority of serum samples were from experimental infections. Estimates of sensitivity ranged from 69% to 82%, while specificity ranged from 96% to 99%. Area under ROC curves and Q were 0.9503 and 0.8909 respectively, indicating a high diagnostic accuracy. Current evidence suggests that ESAT-6 based serodiagnosis has the potential to become useful diagnostic tools for NHP tuberculosis.

Keywords
ESAT-6, Meta analysis, Nonhuman primate, Serodiagnosis, Tuberculosis

INTRODUCTION

Among infectious diseases of non-human primates (NHP), tuberculosis is the most economically devastating one. Tuberculosis in NHP poses a substantial threat to their neighbors, animal technicians, scientists, and post-mortem technicians (Garcia et al., 2004a; Wolf et al., 2014). Since most of the animals are asymptomatic at early stage of infection, tuberculosis may spread quickly through a colony before the index animal is detected. The tuberculin skin testing (TST) has kept as the mainstay of tuberculosis diagnosis in living NHP since the 1940s. But several limitations of TST significantly reduce its efficiency, such as low specificity, anergic reaction, intermittent positive reaction on repeated TSTs, and even false-negative TST results for the true positive (Walsh et al., 1996; Garcia et al., 2004b). Sputum microscopy, currently the sole diagnostic test in most active tuberculosis of human, presents no values for NHP tuberculosis because of its habit of swallowing sputum. Histological and microbiological investigations are always used to confirm tuberculosis at necropsy. Thus, supplementary diagnostic methods to TST are needed urgently (Maas et al., 2013).

In recent years, many possible auxiliary diagnostic methods for tuberculosis have been developed, including serodiagnosis (Lin et al., 2008; Song et al., 2014), gamma interferon in whole blood (Vervenne et al., 2004), multi-PCR, and RT-PCR (Mukherjee et al., 2012). However, serodiagnosis constitutes an attractive diagnostic method because of its convenience, robustness and easy implementation, and low exposure. But it was subjected to efficient antigen screening. Purified protein derivative (PPD) or old tuberculin (OT) based ELISA were limited for their highly cross-reactivity to mycobacterial species, which might result in both false positive and false negative reported by many studies (Placket et al., 1989; Sayin and Erganis, 2013).
By employing molecular biology techniques, many antigens have been expressed and evaluated (Weldingh et al., 2005; Zhang et al., 2009). Several antigens, such as ESAT-6 and CFP10, have been included in commercial immunochromatographic test kits for NHP tuberculosis diagnosing (Lyashchenko et al., 2007; Parsons et al., 2009, 2010; Nath et al., 2012). Among those antigens, ESAT-6 is one of the most studied antigens with high efficiency for serodiagnosis of NHP tuberculosis. But to date, no systematic review on this topic has been summarized except some reviews mentioned ESAT-6-ELISA (Lerche et al., 2008). In order to obtain a more comprehensive evaluation of ESAT-6 for serodiagnosis of NHP tuberculosis, we planned the current systematic review of the existing evidence.

**METHODOLOGY**

**Search strategy:** Research data on ESAT-6 published in English or Chinese between 2000 and 2014 were screened. The databases included PubMed, Google Scholar, China National Knowledge Infrastructure (CNKI), and Chinese Wanfang Data. The search terms used included the following: “tuberculosis”, “Mycobacterium tuberculosis”, “nonhuman primate”, “monkey”, “macaque”, “ESAT-6”, and “serodiagnosis”. We performed the last search on 18 September 2014. Additional studies were identified by contacting experts in the field and industry representatives.

**Study selection:** All available studies were considered to be included on assessment of ESAT-6 for serodiagnosis of NHP tuberculosis among true positives and negative controls (only studies focused on single ESAT-6 antigen were accepted). Studies as follow were excluded: (a) human studies, (b) case reports, (c) reviews, (d) letters, (e) studies designed to assess the serodiagnostic efficacy of multiply fusion antigen including ESAT-6.

**Data extraction and synthesis:** All included studies were screened and assessed by a reviewer (Min F.) and verified by another reviewer (Wang J). Any disagreement between reviewers was resolved by consensus or discussion with the third reviewer. According to the PRISMA requirements reported previously (Moher et al., 2009), a study selection flow diagram would be generated. Extracted data items comprised publication year, name of first author, sample source, and outcome measures.

**Assessment of study quality:** The quality of included studies was independently assessed by two reviewers using QUADAS-2 (Quality Assessment of Diagnostic Accuracy Studies-2), an updated and validated tool to evaluate the quality of diagnostic accuracy studies (Whiting et al., 2011).

**Exploring heterogeneity:** In test accuracy studies, one of the primary causes of heterogeneity is threshold effect, which arises for different criteria (different cut-offs or thresholds) used in different studies to define a positive/negative result. And threshold effect could be suggested by a typical “shoulder arm” pattern in sensitivity against (1-specificity) in Receiver Operating Characteristics (sROC) space drawn according to accuracy estimates from each study (Zamora et al., 2006). In this study, a sROC space and spearman correlation coefficient between logit of sensitivity and logit of 1-specificity would be used to assess the threshold effect. Other causes (non-threshold effect) that may arise heterogeneity would be evaluated by “diagnostic OR” forest plots and Cochran-Q test.

**Statistical analysis:** For each included study, sensitivity, specificity, positive rate and 95% confidence intervals (CIs) were calculated and summarized in forest plots. Pooled sensitivity and specificity across the included studies were estimated according to total sample size. SROC curves were analyzed to produce a global summary of test accuracy. META-DISC, version 1.4 (Hospital Ramony Cajal, Madrid, Spain), were used to performed the statistical analysis (Zamora et al., 2006). If there was heterogeneity among the included studies, the random effects model (DerSimonian and Laird) would be performed, and the fixed effects model (Mantel–Haenszel) otherwise.

**RESULTS AND DISCUSSION**

**Description of included studies:** Selection process of included studies is shown in **Figure 1**. After primary literature searches, 399 studies (387 in English and 12 in Chinese) and a Unite States patent were identified. A total of 37 duplicate studies were excluded, including 33 English studies, 2 Chinese studies, and 1 patent. Of the 362 studies identified after excluding duplicate ones, only a subset of the 21 studies were left after screening of titles and abstracts. Those 21 studies were further retrieved for full text review, and 8 studies met our inclusion criteria and were considered eligible for meta-analysis. **Table 1** shows the description of included studies.
Figure 1. Selection of included studies from systematic review. Eight studies were included in meta analysis after strict selection.

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

<table>
<thead>
<tr>
<th>Reference</th>
<th>Infection</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>True positive</td>
<td>Total</td>
</tr>
<tr>
<td>Brusasca et al. (2003)</td>
<td>Experimental</td>
<td>17</td>
<td>17</td>
</tr>
<tr>
<td>Kanaujia et al. (2013)</td>
<td>Experimental</td>
<td>22</td>
<td>25</td>
</tr>
<tr>
<td>Ravindran et al. (2014)</td>
<td>Experimental</td>
<td>25</td>
<td>34</td>
</tr>
<tr>
<td>Lyashchenko et al. (2007)</td>
<td>Experimental</td>
<td>10</td>
<td>14</td>
</tr>
<tr>
<td>Khan et al. (2008)</td>
<td>Experimental</td>
<td>10</td>
<td>10</td>
</tr>
<tr>
<td>Min et al. (2013)</td>
<td>Experimental</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td>Min et al. (2011)</td>
<td>Natural</td>
<td>48</td>
<td>71</td>
</tr>
<tr>
<td>Wang (2009)</td>
<td>Natural</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

There are numerous studies focused on the use of ESAT-6 in immunodiagnostic (Salman et al., 2012; Goyal et al., 2014), vaccine (Langermans et al., 2005; Hoang et al., 2013), and pathogenic mechanism of human tuberculosis (Gao et al., 2004; Fortune et al., 2005; Sreejit et al., 2014). But few studies were reported for diagnosis of NHP tuberculosis. In present study, only 8 studies met our search strategy and selection criteria, including 178 positive and 481 negative samples. The limited sample resource from limited studies is a main deficiency of this study, which may cause the bias of the pooled results.

Assessment of study quality: All included studies used a case-control study design. A summary of the quality of the studies, judged using the QUADAS-2 criteria for diagnostic studies, was displayed graphically in Figure 2. Detailed describing the execution of the index test was provided sufficiently for all included studies. The assessment by the QUADAS-2 tool of study quality was totally satisfied except the part of “patient selection”, which showed high risk of bias for sample resources. In the present study, six studies were based on experimental tuberculosis infection models that had high risk of bias, and only two studies performed on natural tuberculosis
Figure 2. Quality of studies using QUADAS-2 criteria. The study quality was totally satisfied except the part of “patient selection” showing high risk of bias for sample resources.

Figure 3. Forest plots of “diagnostic OR”.

Infection monkeys showing low (Min et al., 2011) and unclear risk of bias (Wang, 2009).

**Heterogeneity of accuracy estimates:** Heterogeneity of accuracy estimates caused by threshold effect was assessed by spearman correlation coefficient. The spearman correlation coefficient was - 0.357 with a p-value of 0.385, indicating no existence of heterogeneity coming from threshold effect. Heterogeneity caused by non-threshold effect was evaluated by “diagnostic OR” forest plots (Figure 3). Result showed that there was no heterogeneity from non-threshold effect (Heterogeneity chi-squared=8.51, p=0.290). Spearman correlation coefficient and “diagnostic OR” forest plots results proved that all included studies were statistically homogeneous, indicating high quality of included studies.

**Accuracy estimates:** Statistical pooling was appropriate in our systematic review, for all of the included studies and results were reasonably homogeneous. Sensitivity and specificity estimates were shown in Figure 4 respectively. The pooled average sensitivity was 76%, ranging from 69% to 82%. And the specificity estimates ranged from 96% to 99% with a pooled average specificity of 98%. Ideal diagnostic method or gold standard (test) is expected to have sensitivity and specificity as high as possible. A sensitivity of 83% and a specificity of 87% are always regarded as basic criteria for gold standard (test) of diagnostic accuracy. The pooled specificity of ESAT-6 in serodiagnosis was high and reliable, while the pooled sensitivity was moderate to high. Overall, the ESAT-6 based serodiagnosis has the potential to become useful diagnostic tools for NHP tuberculosis. What is more, our review analysis reflected an actual event that serodiagnosis for NHP tuberculosis based on one single antigen was insufficient and antigen cocktail might improve the serodiagnostic efficiency. For example, the PrimaTB STAT-PAK assay, as a commercial serodiagnostic test kit for NHP tuberculosis, has employed the antigen cocktail of ESAT-6, CFP10, MPB83, and TBF10 (Lyashchenko et al., 2007).
Figure 4. Forrest plots of sensitivities and specificities from the included studies.

Figure 5. SROC curve for all studies included in systematic review.
Though all of the included studies were proved to be statistically homogeneous, high heterogeneities were observed in both sensitivity estimates (heterogeneity chi-squared=21.32, p=0.003) and specificity estimates (heterogeneity chi-squared=17.53, p=0.014), which were largely due to the sample sources.

**SROC curves:** Sensitivity against 1-specificity (false-positive rate) in sROC space and sROC curves were constructed from the bivariate fixed effects regression models used to calculate the pooled estimates (Figure 5). No presentation of a typical "shoulder arm" pattern was seen in sROC space, which suggested no presence of threshold effect too. In sROC curves, the middle curve was the fitted sROC curve, and the other two curves were 95% CIs. For the fitted sROC curve, the values of area under curves (AUC) and Q were 0.9503 and 0.8909 respectively, indicating a high accuracy of ESAT-6 based serodiagnosis of tuberculosis.

**CONCLUSION**

ESAT-6 based antibody detection of tuberculosis bacteria is one potentially revolutionary diagnostic for NHP tuberculosis, which could be a supplement to TST. Cocktails of ESAT-6 and other antigens may raise the diagnostic efficiency. Further studies are warranted to determine the value of specific antibody detection base on more antigens for NHP tuberculosis. Considering the rapidness and low expense of antibody detection tests, researches to improve their performances are urgently needed.

**ACKNOWLEDGMENT**

This work was supported by grants 2013B020307005 of Guangdong Provincial Science & Technology Project.

**REFERENCES**


