Commercial serological tests for the diagnosis of active TB: The evidence is reviewed

Karen R Steingart, MD, MPH
New Diagnostics Working Group
Lille, 26 October, 2011
karenst@uw.edu
Disclosure

- I have no financial disclosures to declare
- I have published previous systematic reviews on serological tests
- I serve as Coordinator of the Evidence Synthesis & Policy subgroup of Stop TB Partnership’s New Diagnostics Working Group
Overview

- Background
- The evidence is reviewed
  - Updated systematic review and meta-analysis
  - WHO/TDR evaluation of rapid tests
  - Economic and epidemiological impact of serologic testing for active TB in India
- WHO policy statement on serological tests
- Response to the new policy
Background - definitions

- **Antigen** – any molecule that can bind specifically to an antibody (the name comes from the ability to generate antibody)

- **Antibody** - a protein that binds specifically to a particular substance, its antigen; all antibody molecules belong to a family of proteins called immunoglobulins

- **Serological tests for TB** - tests (such as ELISA, immunochromatographic tests) on a sample of blood serum that detect the humoral immune (antibody) responses to *M. tuberculosis* antigens

- Do not confuse serological tests with IGRAs that measure the T-cell-based interferon-gamma response to *M. tuberculosis* antigens

Janeway, Immunobiology, 6th edition
Background - advantages

- Serological tests could be developed into point-of-care tests
- Serological tests provide rapid results
  - ELISA, within hours
  - immunochromatographic assay, within minutes
- For children, a blood test may be more practical than sputum microscopy
- For patients suspected of extrapulmonary TB, a blood test, if accurate, could replace more invasive tests
Package inserts claim high accuracy

2) Comparison SD Rapid TB vs. a commercial anti-TB ELISA
The SD Rapid TB have tested with positive and negative clinical samples tested by a leading commercial ELISA test. The result shows that the SD Rapid TB is very accurate to other commercial ELISA test.

<table>
<thead>
<tr>
<th>A Commercial PHA</th>
<th>Total Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>Negative</td>
</tr>
<tr>
<td>Positive</td>
<td>112</td>
</tr>
<tr>
<td>Negative</td>
<td>1</td>
</tr>
<tr>
<td>Total Results</td>
<td>113</td>
</tr>
</tbody>
</table>

In a comparison of the SD Rapid TB versus a leading commercial ELISA test, results gave sensitivity of 98.2% (112/114), a specificity of 99.7% (350/351), and a total agreement of 99.35% (462/465).

PERFORMANCE CHARACTERISTICS:

Sensitivity: Sera were collected from patients under anti TB treatment. Results of sputum examination were not available. Among 75 sera collected, samples were positive by the TB onsite Rapid screening Test Thus, the test sensitivity is 93%.

Specificity: In 53 sera derived from Northern America, all the samples were negative.

Sensitivity = 98%
Specificity = 100%

Sensitivity = 93%
Specificity = 100%
Serological tests for TB have not been found to perform well in previous systematic reviews.

Commercial Serological Antibody Detection Tests for the Diagnosis of Pulmonary Tuberculosis: A Systematic Review

Karen R. Steingart, Megan Henry, Suman Laal, Philip C. Hopewell, Andrew Ramsay, Dick Menzies, Jane Cunningham, Karin Weldingh, Madhukar Pai

A systematic review of commercial serological antibody detection tests for the diagnosis of extrapulmonary tuberculosis

Karen R. Steingart, Megan Henry, Suman Laal, Philip C. Hopewell, Andrew Ramsay, Dick Menzies, Jane Cunningham, Karin Weldingh, Madhukar Pai

Performance of Purified Antigens for Serodiagnosis of Pulmonary Tuberculosis: a Meta-Analysis

Deeply troubling…

- Serological tests are being used widely in a majority of high TB burden countries

- “Our survey also confirms the previous observation that companies in western countries (e.g. France, UK, USA, Germany, Australia) are exporting inaccurate and unreliable TB diagnostics to poor countries, while not approving the same tests for domestic use.” Grenier, Eur Respir J, 2011, in press
In 2010, WHO convened a process to develop recommendations about commercial serological tests

- Commissioned an updated systematic review and a decision-analysis model
- Convened an Expert Group to assess the evidence base
- Used the GRADE approach to rate the quality of evidence and determine the strength of recommendations

www.gradeworkinggroup.org
Objective: To obtain summary estimates of the diagnostic accuracy of commercial serological tests for the diagnosis of pulmonary and extrapulmonary TB

Participants: adults and children, all countries

Reference standards
- Pulmonary TB: Culture, solid or liquid
- Extrapulmonary TB: Smear, culture, histopathology
Methods

- Two independent reviewers
- Updated literature search from previous systematic reviews, all languages
- QUADAS to appraise methodological quality
- Prespecified subgroups by test, smear, HIV
- Meta-analysis by hierarchical SROC random effects model
- The GRADE approach to determine quality of the body of evidence
PRISMA Diagram

- 4256 citations
- 160 full-text papers

- PTB: 31 papers
  original review (20)
  update (11)

- EPTB: 12 papers
  original review (9)
  update (3)
Characteristics of included studies

- **Pulmonary TB:** 67 studies (5147 participants); 48% studies from low and middle-income countries
  - anda-TB (IgG, IgA, and IgM) was the test most frequently evaluated (16 studies, 24%)

- **Extrapulmonary TB:** 25 studies (1809 participants); 40% from low and middle-income countries
  - anda-TB (IgG, IgA, and IgM) was the test most frequently evaluated (17 studies, 68%)
Pulmonary TB
Sensitivity range: 0 to 100%
Specificity range: 31 to 100%
Extrapulmonary TB

Sensitivity range: 0 to 100%
Specificity range: 59 to 100%
<table>
<thead>
<tr>
<th>Study</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alifano 1994</td>
<td>35</td>
<td>2</td>
<td>7</td>
<td>92</td>
<td>0.83 [0.69, 0.93]</td>
<td>0.98 [0.93, 1.00]</td>
</tr>
<tr>
<td>Alifano 1996 (a)</td>
<td>28</td>
<td>3</td>
<td>5</td>
<td>41</td>
<td>0.85 [0.68, 0.95]</td>
<td>0.93 [0.81, 0.99]</td>
</tr>
<tr>
<td>Kalantri 2005 (a)</td>
<td>84</td>
<td>0</td>
<td>21</td>
<td>40</td>
<td>0.80 [0.71, 0.87]</td>
<td>1.00 [0.91, 1.00]</td>
</tr>
<tr>
<td>Okuda 2004 (a)</td>
<td>26</td>
<td>10</td>
<td>6</td>
<td>101</td>
<td>0.82 [0.65, 0.93]</td>
<td>0.91 [0.84, 0.96]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watanabe 2005</td>
<td>58</td>
<td>4</td>
<td>34</td>
<td>30</td>
<td>0.63 [0.52, 0.73]</td>
<td>0.88 [0.73, 0.97]</td>
</tr>
<tr>
<td>Wu 2005</td>
<td>35</td>
<td>19</td>
<td>30</td>
<td>40</td>
<td>0.54 [0.41, 0.66]</td>
<td>0.68 [0.54, 0.79]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Erer 2001</td>
<td>20</td>
<td>0</td>
<td>23</td>
<td>20</td>
<td>0.47 [0.31, 0.62]</td>
<td>1.00 [0.83, 1.00]</td>
</tr>
<tr>
<td>McConkey 2002</td>
<td>62</td>
<td>13</td>
<td>9</td>
<td>61</td>
<td>0.87 [0.77, 0.94]</td>
<td>0.82 [0.72, 0.90]</td>
</tr>
<tr>
<td>Ongut 2006</td>
<td>21</td>
<td>0</td>
<td>32</td>
<td>54</td>
<td>0.40 [0.26, 0.54]</td>
<td>1.00 [0.93, 1.00]</td>
</tr>
<tr>
<td>Perkins 2003</td>
<td>77</td>
<td>5</td>
<td>43</td>
<td>29</td>
<td>0.64 [0.55, 0.73]</td>
<td>0.85 [0.69, 0.95]</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Study</th>
<th>TP</th>
<th>FP</th>
<th>FN</th>
<th>TN</th>
<th>Sensitivity</th>
<th>Specificity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Alifano 1997 (a)</td>
<td>24</td>
<td>4</td>
<td>8</td>
<td>24</td>
<td>0.75 [0.57, 0.89]</td>
<td>0.86 [0.67, 0.96]</td>
</tr>
<tr>
<td>Alifano 1997 (b)</td>
<td>38</td>
<td>4</td>
<td>18</td>
<td>24</td>
<td>0.68 [0.54, 0.80]</td>
<td>0.86 [0.67, 0.96]</td>
</tr>
<tr>
<td>Conde 2004 (a)</td>
<td>30</td>
<td>7</td>
<td>10</td>
<td>24</td>
<td>0.75 [0.59, 0.87]</td>
<td>0.77 [0.59, 0.90]</td>
</tr>
<tr>
<td>Conde 2004 (b)</td>
<td>33</td>
<td>16</td>
<td>7</td>
<td>15</td>
<td>0.82 [0.67, 0.93]</td>
<td>0.48 [0.30, 0.67]</td>
</tr>
<tr>
<td>Julian 2000</td>
<td>20</td>
<td>19</td>
<td>7</td>
<td>3</td>
<td>0.74 [0.54, 0.89]</td>
<td>0.14 [0.03, 0.35]</td>
</tr>
</tbody>
</table>

- **ICT**
- **TB-EIA**
- **and aTB IgG, Extrapulmonary TB**
Mycobacteria A60 ELISA Kit

Diagnostic kit for the quantitative determination of anti-mycobacteria antibodies in human serum.

Kit components:
- Sample Diluent
- Anti-Human IgG Conjugated Aff
- TMB Solution
- Substrate Solution
- Stop Solution
- ELISA plates

IgG 192

Anda Biologicals
### Methodological quality summary with QUADAS, anda-TB IgG, smear-negative patients

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Okuda 2004 (b)</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

Summary HSROC plots for anda-TB IgG: (A) smear-positive and (B) smear-negative pulmonary TB patients


Smear Positive
Sensitivity = 76% (63,87)
Specificity = 92% (74,98)

Smear Negative
Sensitivity = 59% (10,96)
Specificity = 91% (79,96)
HROC plots by assay technique
(A) ELISA and (B) Rapid tests

ELISA
Sensitivity = 60% (6, 65)
Specificity = 98% (96, 99)

Rapid tests
Sensitivity = 53% (42, 64)
Specificity = 98% (76, 99)

### Head-to-head comparison SDHO and smear microscopy, HIV-infected persons

<table>
<thead>
<tr>
<th>Test</th>
<th>Sensitivity % (95% CI)</th>
<th>Specificity (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td>SDHO (Saint-Sauveur des Monts, Canada)</td>
<td>16 (5, 34)</td>
<td>90 (74, 98)</td>
</tr>
<tr>
<td>Smear microscopy</td>
<td>68 (49, 83)</td>
<td>100 (89, 100)</td>
</tr>
</tbody>
</table>

- 55 HIV-infected individuals suspected of having pulmonary TB, inpatient and outpatient
- 31 culture-confirmed TB cases
- Median age 31
- Central African Republic

Quality of evidence can be decreased by 5 factors

1. Study limitations (QUADAS criteria)
2. Inconsistency (unexplained heterogeneity)
3. Indirectness
4. Imprecision (width of confidence intervals)
5. Publication bias
Table 3. GRADE Evidence Profile: should commercial serological tests be used as a replacement test for conventional tests such as smear microscopy in patients of any age suspected of having pulmonary tuberculosis?

<table>
<thead>
<tr>
<th>Outcome</th>
<th>Number of Studies (Participants)</th>
<th>Study Design</th>
<th>Limitations</th>
<th>Indirectness</th>
<th>Inconsistency</th>
<th>Imprecision</th>
<th>Publication Bias</th>
<th>Final Quality</th>
<th>Effect per 1,000a</th>
<th>Importanceb</th>
</tr>
</thead>
<tbody>
<tr>
<td>True Positives</td>
<td>67 (5,147)</td>
<td>Cross-sectional and case-control</td>
<td>Very seriousc (-2)</td>
<td>No serious indirectnessd</td>
<td>Very seriousd (-2)</td>
<td>Seriousf</td>
<td>Likelyg</td>
<td>Very low ⊕〇〇〇</td>
<td>Prevalence 10%: 64; prevalence 30%: 192</td>
<td>Critical</td>
</tr>
<tr>
<td>True Negatives</td>
<td>67 (5,147)</td>
<td>Cross-sectional and case-control</td>
<td>Very seriousc (-2)</td>
<td>No serious indirectnessd</td>
<td>Very seriousd (-2)</td>
<td>Seriousf</td>
<td>Likelyg</td>
<td>Very low ⊕〇〇〇</td>
<td>Prevalence 10%: 819; prevalence 30%: 637</td>
<td>Critical</td>
</tr>
<tr>
<td>False Positives</td>
<td>67 (5,147)</td>
<td>Cross-sectional and case-control</td>
<td>Very seriousc (-2)</td>
<td>No serious indirectnessd</td>
<td>Very seriousd (-2)</td>
<td>Seriousf</td>
<td>Likelyg</td>
<td>Very low ⊕〇〇〇</td>
<td>Prevalence 10%: 81; prevalence 30%: 63</td>
<td>Critical</td>
</tr>
<tr>
<td>False Negatives</td>
<td>67 (5,147)</td>
<td>Cross-sectional and case-control</td>
<td>Very seriousc (-2)</td>
<td>No serious indirectnessd</td>
<td>Very seriousd (-2)</td>
<td>Seriousf</td>
<td>Likelyg</td>
<td>Very low ⊕〇〇〇</td>
<td>Prevalence 10%: 36; prevalence 30%: 108</td>
<td>Critical</td>
</tr>
</tbody>
</table>

Based on sample size = 8,318, sensitivity median = 64%, specificity median = 91%.

aWhat do these results mean given 10% or 30% prevalence among individuals being screened for TB?
bOutcomes were ranked by their relative importance as critical, important, or of limited importance. Ranking helped to focus attention on those outcomes that were considered most important.
cThe majority of studies lacked a representative patient population and were not blinded.
dAlthough diagnostic accuracy is considered a surrogate for patient-important outcomes, we did not downgrade.
GRADE defines quality as confidence in the estimates of effect
Objective: To compare performance and reproducibility of rapid MTB-specific antibody detection tests using archived serum samples from the WHO/TDR TB Specimen Bank

Reference standard: culture and clinical follow-up
Methods

- Rapid test - result < 15 minutes
- Simple - 1 or 2 steps, minimal training and no equipment
- Easy to interpret - card or strip format with visual readout
- Archived specimens from Uganda, The Gambia, Canada, Tanzania, Brazil, and Spain
- ROC plots
WHO/TDR Laboratory-based evaluation

All samples, $n = 355$

HIV negative samples, $n = 198$

HIV positive samples, $n = 157$

Sensitivity = 1 to 60%
Specificity = 53 to 99%
Objective: to estimate costs and effectiveness of sputum microscopy (US$3.62 for two smears), microscopy plus automated liquid culture (MGIT, US$20/test), and serological testing (anda-tb ELISA, US$20/test)
Hypothetical study population

- 1.5 million TB suspects
- 1/7 with TB
- 53% TB patients are highly infectious
- 5% HIV prevalence
- 10% with access to ART
- Accuracy estimates from the updated systematic review

Simplified version of study decision tree

Compared with no testing
- **Sputum smear**: additional 44,000 TB cases, 36,000 false positives (FPs)
- **Serology as replacement test**: additional 58,000 TB cases, 157,000 FPs
- **Smear estimated to avert 102,000 more DALY*s, 32,000 more secondary cases** than serology, at ~ 1/4 the incremental cost

\*DALY, disability-adjusted life year

<table>
<thead>
<tr>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>Balance between desirable and undesirable effects</td>
</tr>
<tr>
<td>Quality of evidence</td>
</tr>
<tr>
<td>Values and preferences</td>
</tr>
<tr>
<td>Costs (resource allocation)</td>
</tr>
</tbody>
</table>

Guyatt GH et al. BMJ 2008
Commercial serological tests provide inconsistent and imprecise findings resulting in highly variable values for sensitivity and specificity…high proportions of false-positive and false-negative results adversely impact patient safety. Overall data quality was graded as very low and it is strongly recommended that these tests not be used for the diagnosis of pulmonary and extra-pulmonary TB.

Targeted further research to identify new/alternative point-of-care tests for TB diagnosis and/or serological tests with improved accuracy is strongly encouraged.

Reactions to the WHO policy against the use of TB serological tests

“Responses from governments of high-burden countries have been overwhelmingly positive,” Karin Weyer, WHO Stop TB Department. Morris K, The Lancet Infect Dis 2011

The (Indian) Union Health Ministry has asked all state tuberculosis (TB) officers to endorse the recommendations of the World Health Organization (WHO), urging countries to ban “unapproved” blood tests to diagnose the disease http://www.indianexpress.com/news/tb-battle-states-told-to-follow-who-guideli/820721/

“These tests were discouraged from use almost 20 years ago globally but 10 out of 18 types of strips are still in use in the private sector in Kenya today,” Ms Lucy Chesire, one of the two Kenyan TB experts involved in a WHO study of the problem http://www.nation.co.ke/News/TK+tests+done+at+private+clinics+not+accurate/-/1056/1207098/-/item/0/-/nxponp/-/index.html
References


Acknowledgements

Jane Cunningham (WHO Special Programme for Research and Training in Tropical Diseases), Chris Gilpin (WHO Stop TB Department), Krystal Kobasic (University of California, San Francisco), Megan Henry (Sacramento, California), Christian Lienhardt (Stop TB Partnership Research Movement), Anna Meddaugh (Portland, Oregon), Dick Menzies (McGill University), Alan Mishchenko (University of California, Berkeley), Carl-Michael Nathanson (WHO Special Programme for Research and Training in Tropical Diseases), Marek Perkowski (Portland State University), John Phillips (Fishbon Memorial Library, University of California, San Francisco), Irina Rudoy (University of California, San Francisco), Holger Schünemann (McMaster University), Karin Weldingh (Novo Nordisk), Karin Weyer (WHO Stop TB Department), Gloria Won (Fishbon Memorial Library, University of California, San Francisco), and George Yen (University of California, San Francisco). In addition, we are grateful to Mien Pathey and Jia Si Duan for administrative support and the members of the WHO Expert Group Meeting on Serodiagnostics, July 2010.

Funders: USAID through grant administered by UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases TDR); WHO Stop TB Department; Bill & Melinda Gates Foundation; New Diagnostics Working Group of the Stop TB Partnership
Are we going to see THE END OF TB in our lifetimes?

Merci!