Case Study: Serological tests for tuberculosis

Karen R Steingart, MD, MPH
McGill University, Montreal, July 2011
karenst@uw.edu
Disclosure

- I have no financial disclosures to declare
- I have published previous systematic reviews on this topic
Overview

- Background
- Commercial serological tests for the diagnosis of active pulmonary and extrapulmonary tuberculosis: An updated systematic review and meta-analysis
- Lessons learned
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 (Ig)

- **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 with IGRAs that measure the T-cell-based interferon-gamma response to *M. tuberculosis antigens*

Janeway, Immunobiology, 6th edition
Antigen Soup

- 38 kDa, Rv0934, Ag 5, Ag 78, phoS, phoS1, LAM, glycolipids, diacyltrehaloses (DAT), triacyltrehaloses (TAT), 2,3-diacyl-trehalose-2'-sulphate (SL-IV), trehalose dimycolate (cord factor), PPD, p90 (Kp90), 34 kDa and 38 kDa, excretory-secretory (ES)-31, excretory-secretory (ES)-41, excretory-secretory (ES)-31 and excretory-secretory (ES)-41, TbF6 polyprotein (Mtb8, 38kDa, Mtb11, Mtb48), TbF6/DPEP (MPT32), 38-26-16 kDa, 38-26-6 kDa, 38-26-16-6 kDa, 26-16-6 kDa, 26-6 kDa, 26-38 kDa, 26-16 kDa, 6 kDa, (ESAT-6), 16 kDa, (Rv1926c), (MPT63), Rv1271c, Rv1804c, Rv2253, 38 kDa and (Rv:0203, 1271c), Rv:0203, 1271c, 38 kDa, and (Rv:0203, 0603, 1271c, 1804c, and 2253), Rv:0203, 0603, 1271c, 1804c, and 2253 (Is this 14 kDa?), Tb68, 19 kDa, Rv3763, TB72 (epitope 38 kDa), Rv3871, Rv3872, Rv1335, CFP-10 (MTSA-10, MPT11)…
Serological (antibody-based detection) TB tests

...around for a long time, successful for infectious disease
...attractive, especially if made into point-of-care tests
- but existing serological TB tests have variable accuracies and a limited clinical role
WHO/TDR Laboratory-based evaluation...2008

Sensitivity range: 1 to 60%
Specificity range: 53 to 99%
Claims of high accuracy

- Package inserts are usually based on internal company data.

<table>
<thead>
<tr>
<th>A Commercial PHA</th>
<th>Total Results</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive</td>
<td>112 2</td>
</tr>
<tr>
<td>Negative</td>
<td>1 350</td>
</tr>
<tr>
<td><strong>Total Results</strong></td>
<td><strong>113 352</strong></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).

Sensitivity = 98%
Specificity = 100%

**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 = 93%
Specificity = 100%
Current situation

- The International Standards for TB Care and Control discourages their use, ITSC 2009
- Serological tests for active TB are in widespread use in a majority of high TB burden countries, Grenier J et al, ERJ 2011
- In India, market for serological tests estimated at least $15 million USD (~25% of RNTCP annual budget 2010), Specter M, The New Yorker 2010
Deeply troubling...

“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 J et al. ERJ 2011, in press
Commercial Serological Tests for the Diagnosis of Active Pulmonary and Extrapulmonary Tuberculosis: An Updated Systematic Review and Meta-Analysis

Karen R. Steingart¹, Laura L. Flores²,³, Nandini Dendukuri⁴, Ian Schiller⁴, Suman Laal⁵,⁶,⁷, Andrew Ramsay⁸, Philip C. Hopewell²,³, Madhukar Pai⁴,*

Protocol and analysis plan

Selection of studies

Types of studies: Diagnostic studies with any study design
Participants: Patients with suspected/confirmed pulmonary or extrapulmonary TB, all countries
Index tests: Commercial serological tests
Comparator tests: None or smear microscopy
Target conditions: Pulmonary and extrapulmonary TB
Reference standards
- Pulmonary TB: Solid or liquid culture
- Extrapulmonary TB: Smear, culture, histopathological examination

Excluded: Studies published before 1990 and studies with < 10 TB cases
Methods - 2

**Electronic searches:** NLM, Embase, WOS, Biosis
- original search (1 January 1990 to 30 May 2006) restricted to English
- updated search (1 May 2006 to 29 June 2010), all languages

**Other searches:** Reference lists of eligible papers and related reviews were reviewed; researchers in the field were contacted

**Data extraction**
- Details of study: first author; publication year; case country of residence; World Bank country income status; clinical setting; study design; manner of participant selection; number of participants or samples
- Characteristics of participants: age group; HIV status; smear status
- Details of index test: test name, antibody class, assay type (e.g., ELISA or immunochromatographic test)
- Details of reference standards: type of reference, type of culture
- Details of outcomes: TP, FP, FN, TN
Methods - 3

**Quality assessment:** QUADAS 11 core items, industry sponsorship

**Statistical analysis and data synthesis**
- Descriptive analyses with SPSS
- Sensitivity/specificity were calculated (95% CI) and forest plots displayed with RevMan
- Heterogeneity was determined by visual inspection
- Bivariate random effects meta-analysis (WinBUGS)
- HSROC plots with R

**Heterogeneity**
- Prespecified subgroups according to serological test, antibody class, smear, HIV
- Bivariate random effects meta-analysis (at least 4 studies)

**The GRADE approach**
RESULTS
Pulmonary TB
Study characteristics - 1

- 4256 citations; 160 full-text papers
- 31 papers (20 from the original review, 11 from update)
- 67 studies (5147 participants)
- 48% studies from low/middle-income countries
- No RCTs; 55% cross-sectional; 45% case-control
18 serological tests, anda-TB (IgG, IgA, and IgM) most frequently evaluated (24%)

One study involved HIV-infected individuals

No studies involved children

Median number of TB patients in each study 41 (interquartile range 33 to 54)
Flow of studies:

1. Potentially relevant citations identified from electronic databases: 4256
2. Excluded screen1: 4,101
   - Reasons: Duplicate publication: 1509, Relevance: 2592
3. Papers added from reference review and contact with experts: 5
4. Full papers retrieved for more detailed evaluation: 160
5. Excluded screen2: 149
   - Reasons: Active TB unspecified 1, Antigen detection 16, Could not obtain 1, Duplicate 3, Editorial, commentary 2, Extrapolatory TB 2, Fewer than 10 TB cases 2, Insufficient data 9, Latent TB infection 2, Noncommercial 98, Other specimen 2, Pulmonary TB 10, Reference standard lacking 9, Relevance 1, Review 1
6. Papers added from 2007 systematic review: 20
7. Papers (studies) included in the systematic review of commercial serological tests for extrapulmonary tuberculosis: 12 (25)
8. Papers (studies) included in the systematic review of commercial serological tests for pulmonary tuberculosis: 31 (67)
<table>
<thead>
<tr>
<th>Author (year, study)</th>
<th>Smear Status</th>
<th>Country</th>
<th>Comparison Group</th>
<th>Immuno-globulin Detected</th>
<th>No. With/Without TB</th>
<th>Sensitivity % (95% CI)</th>
<th>Specificity % (95% CI)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Active TB</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anderson (2008, a)</td>
<td>Not reported</td>
<td>USA</td>
<td>Healthy</td>
<td>IgG</td>
<td>11/87</td>
<td>100 (72-100)</td>
<td>99 (94-100)</td>
</tr>
<tr>
<td><strong>Anda TB</strong> (Anda Biologics, Strasbourg)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alifano (1994)</td>
<td>Smear positive</td>
<td>Italy</td>
<td>Healthy</td>
<td>IgG</td>
<td>42/44</td>
<td>83 (69-92)</td>
<td>98 (93-100)</td>
</tr>
<tr>
<td>Alifano (1998, a)</td>
<td>Smear positive</td>
<td>Italy</td>
<td>ORD</td>
<td>IgG</td>
<td>33/44</td>
<td>83 (68-89)</td>
<td>93 (81-99)</td>
</tr>
<tr>
<td>Alifano (1998, b)</td>
<td>Smear positive</td>
<td>Italy</td>
<td>ORD</td>
<td>IgA</td>
<td>33/44</td>
<td>82 (65-89)</td>
<td>98 (75-96)</td>
</tr>
<tr>
<td>Alifano (1998, c)</td>
<td>Smear negative</td>
<td>Italy</td>
<td>ORD</td>
<td>IgG</td>
<td>36/44</td>
<td>64 (46-79)</td>
<td>93 (81-99)</td>
</tr>
<tr>
<td>Alifano (1998, d)</td>
<td>Smear negative</td>
<td>Italy</td>
<td>ORD</td>
<td>IgA</td>
<td>36/44</td>
<td>64 (46-79)</td>
<td>98 (75-96)</td>
</tr>
<tr>
<td>Anderson (2008, c)</td>
<td>Not reported</td>
<td>USA</td>
<td>Healthy</td>
<td>IgG</td>
<td>11/75</td>
<td>100 (72-100)</td>
<td>72 (60-82)</td>
</tr>
<tr>
<td><strong>Budlery (2007)</strong></td>
<td></td>
<td>Saudi Arabia</td>
<td>ORD</td>
<td>IgG</td>
<td>48/22</td>
<td>100 (92-100)</td>
<td>32 (14-55)</td>
</tr>
<tr>
<td><strong>Kajari (2005, a)</strong></td>
<td>Smear positive</td>
<td>India</td>
<td>ORD</td>
<td>IgG</td>
<td>105/40</td>
<td>80 (71-87)</td>
<td>100 (91-100)</td>
</tr>
<tr>
<td><strong>Kajari (2005, b)</strong></td>
<td>Smear positive</td>
<td>India</td>
<td>ORD</td>
<td>IgM</td>
<td>105/40</td>
<td>80 (71-87)</td>
<td>100 (91-100)</td>
</tr>
<tr>
<td><strong>Liu (1995, b)</strong></td>
<td>Smear negative</td>
<td>Taiwan</td>
<td>ORD</td>
<td>IgG</td>
<td>70/203</td>
<td>71 (59-82)</td>
<td>99 (85-92)</td>
</tr>
<tr>
<td>Okuda (2004, a)</td>
<td>Smear positive</td>
<td>Japan</td>
<td>ORD</td>
<td>IgG, IgM, IgA</td>
<td>84/111</td>
<td>82 (65-93)</td>
<td>91 (84-98)</td>
</tr>
<tr>
<td>Okuda (2004, b)</td>
<td>Smear negative</td>
<td>Japan</td>
<td>ORD</td>
<td>IgG, IgM, IgA</td>
<td>26/111</td>
<td>73 (52-88)</td>
<td>91 (84-98)</td>
</tr>
<tr>
<td><strong>Tranmuller (2005)</strong></td>
<td>Smear positive</td>
<td>Saudi Arabia</td>
<td>ORD</td>
<td>IgG</td>
<td>36/79</td>
<td>84 (69-94)</td>
<td>73 (62-82)</td>
</tr>
<tr>
<td>Wu (2004, a)</td>
<td>Smear positive</td>
<td>Taiwan</td>
<td>ORD</td>
<td>IgG</td>
<td>92/34</td>
<td>63 (52-73)</td>
<td>88 (73-97)</td>
</tr>
<tr>
<td>Wu (2004, b)</td>
<td>Smear negative</td>
<td>Taiwan</td>
<td>ORD</td>
<td>IgG</td>
<td>86/34</td>
<td>65 (53-76)</td>
<td>91 (75-97)</td>
</tr>
<tr>
<td>Wu (2005)</td>
<td>Smear positive</td>
<td>Taiwan</td>
<td>ORD and Healthy</td>
<td>IgG</td>
<td>65/59</td>
<td>64 (41-66)</td>
<td>68 (54-79)</td>
</tr>
<tr>
<td><strong>Assure TB</strong> (MedTek, Pleasanton City)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alifano (2006, a)</td>
<td>Smear positive</td>
<td>Singapore</td>
<td>ORD</td>
<td>IgG</td>
<td>54/139</td>
<td>72 (58-84)</td>
<td>77 (69-84)</td>
</tr>
<tr>
<td>Alifano (2006, b)</td>
<td>Smear positive</td>
<td>Singapore</td>
<td>ORD</td>
<td>IgG</td>
<td>25/139</td>
<td>48 (27-69)</td>
<td>77 (69-84)</td>
</tr>
<tr>
<td><strong>Detect TB</strong> (Adaxia, Rome)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Antociogato (1999, a)</td>
<td>Smear positive</td>
<td>Italy</td>
<td>Healthy</td>
<td>IgG</td>
<td>54/150</td>
<td>81 (71-92)</td>
<td>97 (92-99)</td>
</tr>
<tr>
<td>Antociogato (1999, b)</td>
<td>Smear positive</td>
<td>Italy</td>
<td>Healthy</td>
<td>IgG</td>
<td>46/150</td>
<td>65 (50-79)</td>
<td>97 (82-99)</td>
</tr>
</tbody>
</table>
Figure. Methodological quality graph, all studies, pulmonary TB. Review authors' judgments about each methodological quality item.

Sensitivity range: 0 to 100%
Specificity range: 31 to 100%
# 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, hospitalized and outpatient
- 31 culture-confirmed TB cases
- Median age 31
- Central African Republic

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)

Methodological quality summary, studies of 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>Wu 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>
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)

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 serious (−2)</td>
<td>No serious indirectness</td>
<td>Very serious (−2)</td>
<td>Serious</td>
<td>Likely</td>
<td>Very low 0000</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 serious (−2)</td>
<td>No serious indirectness</td>
<td>Very serious (−2)</td>
<td>Serious</td>
<td>Likely</td>
<td>Very low 0000</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 serious (−2)</td>
<td>No serious indirectness</td>
<td>Very serious (−2)</td>
<td>Serious</td>
<td>Likely</td>
<td>Very low 0000</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 serious (−2)</td>
<td>No serious indirectness</td>
<td>Very serious (−2)</td>
<td>Serious</td>
<td>Likely</td>
<td>Very low 0000</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.

Footnotes

Steingart KR et al. PLoS Medicine, in press
Limitations

- Small number of studies for a particular serological test
- anda-TB IgG only test with enough studies for meta-analysis
- In some cases, we assumed multiple results carried out on same sample were independent; our meta-analysis model may have underestimated heterogeneity and overestimated precision of the pooled sensitivity and specificity estimates
- Language bias (the original literature search was limited to studies published in English)
- Publication bias - no formal assessment
Cost-effectiveness study - TB serological tests

- The findings from the systematic review were used as input for a cost-effectiveness study of serological testing for active TB in India.

- “In comparison with sputum microscopy, serological testing resulted in fewer disability-adjusted life years averted and more false-positive diagnoses and secondary infections, while increasing costs to the Indian TB control sector approximately four-fold.”

Conclusions - 1

- Despite expansion of the literature since 2006, commercial serological tests continue to produce inconsistent and imprecise estimates of sensitivity and specificity.
- Quality of evidence remains very low.
Conclusions - 2

Re: Future research

- These conclusions should be reconsidered if, in the future, methodologically adequate research evaluating serological tests becomes available.
WHO issues a strong recommendation against the use of serological tests

- WHO Expert Group recommended that serological tests should not be used in adults and children suspected of active pulmonary or extrapulmonary TB, irrespective of their HIV status; stressed the importance of continued research on serological tests (7/22/10)
- WHO policy on TB serology expected in July 2011

‘Negative’ policy guidance is a first for WHO
Lessons learned

- We focused the updated systematic review on anda-TB IgG because this was the only test with enough studies for meta-analysis...we did not capture the totality of evidence on serologic tests in the PLoS Medicine paper
- A protocol and an analysis plan were critical
- The GRADE approach was helpful
- What about publication bias?
- How should we define the future research agenda?
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). We thank David Dowdy (University of California, San Francisco) and the reviewers for their helpful comments on the manuscript. 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.

UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases; WHO Stop TB Department; Bill & Melinda Gates Foundation; New Diagnostics Working Group of the Stop TB Partnership
References


