

Evidence-Based Tuberculosis Diagnosis

Madhukar Pai*, Andrew Ramsay, Richard O'Brien

There is great excitement in the tuberculosis (TB) scientific community over the introduction of new tools into TB control activities. The development of new tools is an important component of the Global Plan to Stop TB and the World Health Organization's new global Stop TB Strategy [1,2]. Anticipating the introduction of new tools, the Stop TB Partnership has established a Retooling Task Force to develop a framework for engaging policy makers to foster accelerated adoption and implementation of new tools into TB control programs [3].

While new tools offer great promise in clinical medicine and in public health, limited resources and the movement toward evidence-based guidelines and policies require careful validation of new tools prior to their introduction for routine use. The world spends an estimated US\$1 billion per year on diagnostics for TB [4]. It is important to ensure that such expenditure is backed by strong evidence.

Ideally, clinical and policy decisions must be guided by the totality of evidence on a given topic. This is particularly relevant for TB, where concerns have been raised about the lack of emphasis on evidence of effectiveness in some of the existing TB guidelines and policies [5]. These concerns are being taken seriously [6,7], and the outcome should be evident in upcoming TB guidelines and policies. In fact, the World Health Organization (WHO) recently announced its approach for developing new policies on TB in a document entitled "Moving Research Findings into New WHO Policies" [7]. According to this document, in order to consider a global policy change, WHO must have solid evidence, including clinical trials or field evaluations in high TB prevalence settings. The

steps involved in the policy process include a comprehensive review of the evidence, as well as expert opinion and judgment (Box 1).

High-quality evidence on TB diagnostics is critical for the development of evidence-based policies on TB diagnosis, and, ultimately, for effective control of the global TB epidemic. While primary diagnostic trials are needed to generate data on test accuracy and operational performance, systematic reviews provide the best synthesis of current evidence on any given diagnostic test [8]. Although a large number of trials on TB diagnostics have been published, surprisingly, no systematic reviews were published until recently. In the past few years, at least 30 systematic reviews and meta-analyses have been published on various TB tests [9–38]. These reviews have synthesized the results of more than 1,000 primary studies, providing valuable insights into the diagnostic accuracy of various tests (Table 1, Box 2).

Implications for Clinical and Laboratory Practice

For clinicians, systematic reviews provide several useful insights for diagnosis of latent TB infection, active TB disease, and drug resistance.

For diagnosis of latent TB, clinicians have used the tuberculin skin test (TST) for decades. Recently, interferon-gamma release assays (IGRAs) have emerged as attractive alternatives. While the TST is known to have poor specificity in populations vaccinated with bacille Calmette-Guérin (BCG) [34], meta-analyses have shown that IGRAs have much higher specificity for TB infection than the TST, and IGRA specificity is unaffected by BCG vaccination [21,26,37]. However, another meta-analysis showed that BCG vaccination received in infancy has a minimal effect on the TST, whereas BCG received after infancy produces more frequent, more persistent, and larger TST reactions [35]. Thus, the TST might retain high specificity in some populations, whereas it may perform poorly in others. IGRAs are particularly attractive

in the latter setting. However, meta-analyses on IGRAs have highlighted the lack of evidence on the predictive ability of these assays in identifying those individuals with TB infection who are at highest risk for progressing to active disease. Several cohort studies are ongoing (reviewed elsewhere [39]), and these should provide useful evidence on this unresolved issue.

For active TB, serological tests have been attempted for decades. Two meta-analyses have convincingly shown that existing commercial antibody-based tests have poor accuracy and limited clinical utility [29,30]. Despite this evidence, dozens of commercial serological tests continue to be marketed, mostly in private sectors of countries that lack diagnostic regulatory bodies [4].

Nucleic acid amplification tests (NAATs) were considered to be a major breakthrough in TB diagnosis when they were first introduced. A series of meta-analyses have shown that NAATs have high specificity and positive predictive value, but modest and highly variable sensitivity, especially

Citation: Pai M, Ramsay A, O'Brien R (2008) Evidence-based tuberculosis diagnosis. *PLoS Med* 5(7): e156. doi:10.1371/journal.pmed.0050156

Copyright: © 2008 Pai et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Abbreviations: ADA, adenosine deaminase; BCG, bacille Calmette-Guérin; FIND, Foundation for Innovative New Diagnostics; IFN- γ , interferon-gamma; IGRA, interferon-gamma release assay; LED, light-emitting diode; MDR-TB, multidrug-resistant tuberculosis; NAAT, nucleic acid amplification test; STAG-TB, Strategic and Technical Advisory Group for Tuberculosis; TB, tuberculosis; TDR, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases; TST, tuberculin skin test; WHO, World Health Organization; XDR-TB, extensively drug-resistant tuberculosis

Madhukar Pai is with the Department of Epidemiology, Biostatistics and Occupational Health, McGill University, Montreal, Quebec, Canada. Andrew Ramsay is with the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, World Health Organization, Geneva, Switzerland. Richard O'Brien is with the Foundation for Innovative New Diagnostics, Geneva, Switzerland.

* To whom correspondence should be addressed. E-mail: madhukar.pai@mcgill.ca

Research in Translation discusses health interventions in the context of translation from basic to clinical research, or from clinical evidence to practice.

Box 1. WHO Policy Process for Tuberculosis

1. Identifying the Need for a Policy Change

The need to formulate new or revised policies may arise from WHO's ongoing monitoring of technical developments or from interested parties submitting requests with supporting documentation for policy or guideline development. WHO receives information about a new technology or approach via many channels, with the most direct lines coming from national TB programs and researchers themselves. To consider a global policy change, WHO must have solid evidence, including clinical trials or field evaluations in high TB prevalence settings.

2. Reviewing the Evidence

WHO may carry out or commission a review of the documentation of the technology's clinical or programmatic performance, including newly published and "grey" research or reviews, "proof of principle" reports, large-scale field trials, and demonstration projects in different resource settings. Standardized evaluation criteria have been and are being developed by the New Diagnostics, New Drugs, and New Vaccines Working Groups of the Stop TB Partnership.

3. Convening an Expert Panel

If the evidence base is compelling, WHO will convene an external panel of experts, excluding all original principal investigators from the studies. The panel will review the evidence and make a recommendation or propose draft policies or guidelines to WHO's Strategic and Technical Advisory Group for Tuberculosis (STAG-TB).

4. Assessing Draft Policies and Guidelines

STAG-TB provides objective, ongoing technical and strategic advice to WHO on TB care and control. STAG-TB's objectives are to provide the Director-General, through the Stop TB Department, with an independent evaluation of the strategic, scientific, and technical aspects of WHO's TB activities; review progress and challenges in WHO's TB-related core functions; review and make recommendations on committees and working groups; and make recommendations on WHO's TB activity priorities. STAG-TB reviews the policy drafts and supporting documentation during its annual meeting. STAG-TB may endorse the policy recommendation with or without revisions, request additional information and re-review the evidence in subsequent years, or reject the recommendation.

5. Formulating and Disseminating Policy

New WHO policies and guidelines will be disseminated through different channels to Member States, including through the World Health Assembly, WHO Web site, listservs, and journal publications. WHO also disseminates its recommendations to other agencies and donors engaged in TB control activities.

Source: World Health Organization [7]

microscopy (with no significant loss in specificity) [31], that sputum processing methods (e.g., bleach or centrifugation) can be effective in increasing the yield of smear microscopy [32], and that liquid cultures are more rapid and sensitive than solid cultures [10].

Implications for Policies and Guidelines

In addition to informing evidence-based TB diagnosis, systematic reviews have been helpful in informing policy decisions. For example, a series of recent reviews has shown that smear microscopy can be optimized using at least three different approaches: chemical and physical processing for concentration of sputum, use of fluorescence microscopy instead of conventional light microscopy, and the examination of two (as compared to three) sputum specimens [20,31,32]. The findings of these reviews were incorporated into the International Standards for TB Care [42], and have informed policy guidance on the diagnosis of smear-negative TB in people living with HIV/AIDS [43].

The review on incremental yield of serial smears showed that the average incremental yield and/or increase in sensitivity of examining a third sputum specimen ranged between 2% and 5% [20]. This suggested that reducing the recommended number of specimens examined from three to two could potentially benefit TB control programs, and potentially increase case detection for several reasons [20]. Partly based on this evidence and expert opinion, WHO recently revised its policies on smear microscopy [44]. It now recommends that the number of specimens to be examined for screening of TB cases be reduced from three to two, in places where a well-functioning external quality assurance system exists, where the workload is very high, and where human resources are limited [44]. The revised WHO definition of a new sputum smear-positive pulmonary TB case is based on the presence of at least one acid fast bacillus in at least one sputum sample in countries with a well-functioning external quality assurance system [45].

These new policies have major implications for resource-poor settings with high TB prevalence where sputum microscopy is the main or

in smear-negative and extrapulmonary TB [9,11,14,18,23,24,28].

Conventional tests such as smears and cultures perform poorly in extrapulmonary TB. A series of reviews have shown that biomarkers such as adenosine deaminase (ADA) and interferon-gamma (IFN- γ) have excellent accuracy for tuberculous pleural effusion [12,13,15,17]. These biomarkers, especially ADA, are easy to measure and inexpensive. Despite this evidence, these tests appear to be underutilized [40].

For the diagnosis of multidrug-resistant TB (MDR-TB), available data suggest that phage-based assays do not perform well when directly applied to clinical specimens [25].

Line probe assays show great promise for rapid detection of rifampicin resistance in settings with high MDR-TB prevalence [22,38]. Simple tests such as colorimetric redox methods and nitrate reductase assays appear to perform very well, but require culture isolation [19,36]. More evidence is needed on rapid tests for drug resistance, especially since the Global XDR-TB Response Plan calls for wide-scale implementation of rapid methods to screen patients at risk of XDR-TB (extensively drug-resistant TB) and MDR-TB [41].

For laboratory practice, systematic reviews provide strong evidence that fluorescence microscopy is more sensitive than conventional light



doi:10.1371/journal.pmed.0050156.g001

Figure 1. Low-Cost LED-Based Fluorescence Microscopy Being Evaluated at a TDR/WHO Trial Site in Abuja, Nigeria

Photographer: Andrew Ramsay (Courtesy of TDR, Geneva)

only diagnostic test available, and particularly where laboratory services are being overwhelmed with demand for smear microscopy. Omitting the third smear could potentially reduce costs and alleviate the workload of laboratories, particularly in countries with human resource crises. In these settings, laboratories performing smear microscopy often have to deal with anemia, malaria, and other diseases. Thus, the time saved from the inefficient examination of a third smear may be applied toward improving laboratory testing for other diseases [20]. The adoption of the revised case definition and a two-smear approach may create the opportunity to examine both smears during a patient's first presentation to a health facility, and thereby reduce the large numbers of patients known to drop out during the diagnostic process [46]. While these are reasonable assumptions, it is worth emphasizing that there is no hard evidence that the two-smear policy actually improves TB control in the real world. Such data will have to come from programmatic research at the country level and from data collected in routine public health program settings.

There is strong evidence that liquid cultures are more sensitive and rapid than solid media cultures [10]. Based on a review of available evidence and an expert consultation, WHO recently issued policy guidance on the use of liquid TB culture and drug susceptibility testing in low-resource settings [47]. The WHO policy recommends phased implementation of liquid culture systems as a part of a country-specific comprehensive plan for laboratory capacity strengthening that addresses issues such as biosafety, training, maintenance of infrastructure, and reporting of results [47]. These policies are expected to have an important impact in settings with high HIV prevalence [43] and in countries where MDR-TB is an increasing problem [41], helping to inform the needed global scale-up of culture and drug susceptibility testing capacity.

However, implementation of culture testing requires a well-functioning health care system, adequate laboratory infrastructure, and trained personnel. Therefore, emphasis must be placed on capacity building and health system and laboratory strengthening [43,48]. Recognizing this, the Stop

TB Partnership, WHO, and partners have launched a Global Laboratory Initiative to facilitate laboratory policy guidance, technical assistance, quality management, resource mobilization, and advocacy. Again, as in the case of the two-smear strategy, it must be emphasized that there is no strong evidence that the WHO policy on liquid cultures actually improves TB control at the routine programmatic level. Field studies and cost-effectiveness data are needed to better understand the real world implications of this policy.

In June 2008, WHO announced a new policy statement, endorsing the use of line probe assays for rapid screening of patients at risk of MDR-TB (<http://www.who.int/tb/en/>). This policy statement was based in part on evidence summarized in systematic reviews [22,38], expert opinion, and results of field demonstration projects. The recommended use of line probe assays is currently limited to culture isolates and direct testing of smear-positive sputum specimens. Line probe assays are not recommended as a complete replacement for conventional culture and drug susceptibility testing. Culture is still required for smear-negative specimens, and conventional drug susceptibility testing is still necessary to confirm XDR-TB.

Following this new policy, WHO, UNITAID, the Stop TB Partnership, and the Foundation for Innovative New Diagnostics (FIND) announced a new initiative to improve the diagnosis and treatment of MDR-TB in resource-limited settings (http://www.who.int/tb/features_archive/mdrtb_rapid_tests/en/index.html). As part of this initiative, over the next few years, 16 countries will begin using rapid tests to diagnose MDR-TB, including line probe assays. The countries will receive specially priced tests through the Stop TB Partnership's Global Drug Facility, which provides countries with both drugs and diagnostic reagents.

Implications for Research and Development

Systematic reviews have been helpful in identifying key knowledge gaps and defining research agendas. For example, based on the smear microscopy reviews [20,31,32] and expert opinion, the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical



doi:10.1371/journal.pmed.0050156.g002

Figure 2. A Simplified NAAT Being Evaluated at a FIND Trial Site in India
Photographer: Ralf Linke (Courtesy of FIND, Geneva)

Diseases (TDR) recently launched a major research program aimed at the optimization of smear microscopy [49]. Large-scale field studies are ongoing in more than ten countries on issues such as optimum timing and composition of sputum specimen sets; use of lower-cost light-emitting diode (LED) fluorescence microscopy systems (Figure 1); sputum processing methods involving bleach digestion; and potential for reducing time to diagnosis and number of patient visits required by examining two specimens on the day that the patient first presents. The latter can be expected to reduce the considerable patient drop-out rates during diagnosis that are seen in many settings [46].

In parallel, FIND recently forged a partnership with Carl Zeiss MicroImaging (<http://www.zeiss.com/micro/>) to develop an inexpensive, robust LED-based microscope that will be extensively evaluated for routine use in high-burden countries [50].

Systematic reviews on existing commercial serological tests and NAATs have shown that these assays have not performed as well as expected [14,18,29,30]. A recent evaluation of 19 rapid commercial serological tests for TB using specimens from the TDR TB Specimen Bank confirmed the poor accuracy of existing serological tests for TB [51]. Such evidence has informed several initiatives to improve serological assays and NAATs. For example,

FIND is supporting the development and evaluation of newer, improved NAATs (Figure 2) [52]. Several groups are working on methods to optimize serological assays, including the use of novel TB-specific antigens, the use of antigen combinations, and the development of point-of-care tests [52].

Systematic reviews on IGRAs have informed the development of guidelines and positions statements in many countries [53,54,55]. They have

also facilitated the development of a comprehensive research agenda with a specific focus on the use of these assays in resource-limited settings [56].

Systematic reviews on TB diagnostics have revealed deficiencies in the quality of TB diagnostic trials. A recent analysis of systematic reviews showed that trials of TB diagnostics lack methodological rigor, and studies are often poorly reported [57]. Lack of methodological rigor in trials is a cause for concern, as it may prove to be a major hurdle for effective application of diagnostics in TB care and control. Biased results from poorly designed trials can lead to premature adoption of diagnostics that may have little or no benefit. The situation is exacerbated by the fact that most developing countries have poor regulatory mechanisms for licensing and post-marketing surveillance of diagnostics. For example, dozens of commercial serological tests are marketed in developing countries, despite lack of evidence on their utility [29,30,51].

It is clear that efforts are needed to improve both methodological quality and reporting of TB diagnostic trials [57,58]. TDR has developed guidelines for researchers on assessing the performance and operational characteristics of diagnostics for infectious diseases [59], and the STARD (Standards for Reporting

Box 2. Five Key Papers in the Field

Dinnes et al., 2007 [10]. The most comprehensive systematic review of several rapid diagnostic tests for the detection of TB, sponsored by the UK Health Technology Assessment Programme.

Mase et al., 2007 [20]. This review on incremental yield of serial smears showed that the average incremental yield and/or the increase in sensitivity of examining a third sputum specimen ranged between 2% and 5%. This evidence partly informed the new WHO policy on smear microscopy.

Menzies et al., 2007 [21]. This meta-analysis showed that IGRAs for TB infection have excellent specificity (higher than the conventional TST), and are unaffected by prior BCG vaccination. This review also highlighted the key unresolved questions regarding the use of these assays in clinical practice. An update to this meta-analysis was published recently (Pai et al., 2008 [37]).

Steingart et al., 2007 [30]. This meta-analysis showed that serological tests for TB produce highly inconsistent estimates of sensitivity and specificity, and none of the currently available commercial assays perform well enough to replace microscopy. Several initiatives are now ongoing to develop improved point-of-care immune-based rapid tests for TB.

Steingart et al., 2006 [31]. This systematic review reported strong evidence that fluorescence microscopy is more sensitive than conventional microscopy. Several initiatives are now ongoing to develop simple, low-cost fluorescence microscopy systems to optimize smear microscopy.

Table 1. Findings from Systematic Reviews on TB Diagnostic Tests

Diagnostic Test [References]	Number of Reviews	Disease/Site	Major Findings/Results of Systematic Reviews
Diagnosis of active TB			
Sputum smear microscopy [20,31,32]	3	Pulmonary TB	<ul style="list-style-type: none"> Fluorescence microscopy is on average 10% more sensitive than conventional microscopy. Specificity of both fluorescence and conventional microscopy is similar. Centrifugation and overnight sedimentation, preceded with any of several chemical methods (including bleach), is more sensitive than direct microscopy; specificity is unaffected by sputum processing methods. When serial sputum specimens are examined, the mean incremental yield and/or increase in sensitivity from examination of 3rd sputum specimen ranges between 2% and 5%.
NAATs [9,10,11,14,18,23,24,28]	8	Pulmonary and extrapulmonary TB	<ul style="list-style-type: none"> NAATs have high specificity and positive predictive value. NAATs, however, have relatively lower (and highly variable) sensitivity and negative predictive value for all forms of TB, especially in smear-negative and extrapulmonary disease. In-house ("home-brew") NAATs produce highly inconsistent results as compared to commercial, standardized NAATs.
Commercial serological antibody detection tests [10,29,30]	3	Pulmonary and extrapulmonary TB	<ul style="list-style-type: none"> Serological tests for both pulmonary and extrapulmonary TB produce highly inconsistent estimates of sensitivity and specificity; none of the assays perform well enough to replace microscopy.
ADA [12,13,17,27,33]	5	TB pleuritis, pericarditis, peritonitis	<ul style="list-style-type: none"> Measurement of ADA levels in pleural, pericardial, and ascitic fluid has high sensitivity and specificity for extrapulmonary TB.
IFN- γ [13,15]	2	TB pleuritis	<ul style="list-style-type: none"> Pleural fluid IFN-γ determination is a sensitive and specific test for the diagnosis of TB pleuritis.
Phage amplification assays [16]	1	Pulmonary TB	<ul style="list-style-type: none"> Phage-based assays have high specificity but lower and variable sensitivity. Their performance characteristics are similar to sputum microscopy.
Automated liquid cultures [10]	1	Pulmonary TB	<ul style="list-style-type: none"> Automated liquid cultures are more sensitive than solid cultures. Time to detection is more rapid than solid cultures.
Diagnosis of latent TB infection			
TST [34,35]	2	Latent TB infection	<ul style="list-style-type: none"> Individuals who receive BCG vaccination are more likely to have a positive TST; the effect of BCG on TST results is less after 15 years; positive TST with indurations of >15 mm are more likely to be the result of TB infection than of BCG vaccination. The effect on TST of BCG received in infancy is minimal, especially 10 years after vaccination. BCG received after infancy produces more frequent, more persistent, and larger TST reactions. Non-tuberculous mycobacterial (NTM) infection is not a clinically important cause of false-positive TST, except in populations with a high prevalence of NTM sensitization and a very low prevalence of TB infection.
T cell-based IGRAs [21,26,37]	3	Latent TB infection	<ul style="list-style-type: none"> IGRAs have excellent specificity (higher than the TST), and are unaffected by prior BCG vaccination.
Diagnosis of drug-resistant TB			
Phage amplification assays [25]	1	Rapid detection of rifampicin resistance	<ul style="list-style-type: none"> When used on culture isolates, phage assays have high sensitivity, but variable and lower specificity. In contrast, evidence is lacking on the accuracy of these assays when they are directly applied to sputum specimens.
Line probe assays: INNO-LiPA Rif. TB (LiPA) [22] and GenoType MTBDR assays [38]	2	Rapid detection of rifampicin resistance	<ul style="list-style-type: none"> LiPA is a highly sensitive and specific test for the detection of rifampicin resistance in culture isolates, with relatively lower sensitivity when used directly on clinical specimens. The GenoType MTBDR assays have excellent sensitivity and specificity for rifampicin resistance even when directly used on clinical specimens.
Colorimetric redox-indicator methods [19] and nitrate reductase assays [36]	2	Rapid detection of rifampicin and isoniazid resistance	<ul style="list-style-type: none"> Colorimetric methods and nitrate reductase assays are highly sensitive and specific for the rapid detection of rifampicin and isoniazid resistance in culture isolates.

doi:10.1371/journal.pmed.0050156.t001

Diagnostic Accuracy) initiative was launched to improve the quality of reporting of diagnostic studies [60].

Conclusions

With the publication of several systematic reviews, there is now a strong evidence base to support global policy on TB diagnostics. A key challenge is to maintain the momentum gained in the past few years, and expand the scope and role of evidence synthesis to outcomes that go beyond conventional diagnostic accuracy. These outcomes include: accuracy of diagnostic algorithms (rather than single tests) and their relative contributions to the health care system; incremental or added value of new tests; impact of new tests on clinical decision-making and therapeutic choices; cost-effectiveness in routine programmatic settings; impact on patient-centered outcomes; and societal impact of new tools. Indeed, the GRADE (Grading of Recommendations Assessment, Development and Evaluation) approach to grading the quality of evidence and strength of recommendations for diagnostic tests recognizes that diagnostic accuracy results are surrogates for patient-centered outcomes, and emphasizes that diagnostic tests are of value only if they result in improved outcomes for patients [61].

In addition to expanding the scope of evidence synthesis, it is also important to ensure that systematic reviews stay current by including new literature. Periodic updates are needed to ensure that systematic reviews provide the most current evidence available for clinical and policy decisions. For example, the literature on IGRAs has exploded in the past few years, and this necessitated an updated meta-analysis on this topic [37].

Recognizing the growing importance of evidence-based TB diagnosis and policy making, the Stop TB Partnership's New Diagnostics Working Group has recently created a new subgroup on Evidence Synthesis for TB Diagnostics [62]. This subgroup will support the development of new systematic reviews, facilitate the development and dissemination of evidence summaries on new diagnostics, and actively promote their use in guideline and policy development processes, along the lines of the GRADE approach. ■

Acknowledgments

The authors acknowledge the excellent contributions made by authors of the systematic reviews cited in this work. Their efforts have made evidence-based TB diagnosis a reality. The authors are grateful to Professor S. P. Kalantri for helpful comments on a draft of this manuscript. **Funding:** MP is a recipient of a New Investigator Career Award from the Canadian Institutes of Health Research. This agency had no involvement in the preparation of this manuscript. **Competing Interests:** The authors have no financial conflicts of interest. All the authors are members of the Stop TB Partnership's Working Group on New Diagnostics. AR is the secretary of the Working Group. MP and ROB are co-chairs of the Working Group's subgroup on Evidence Synthesis for TB Diagnostics. ROB works for the Foundation for Innovative New Diagnostics, a nonprofit agency that collaborates with several industry partners for the development of new diagnostics for neglected infectious diseases. No industry partner was involved in the preparation of this manuscript.

References

1. Stop TB Partnership, World Health Organization (2006) The global plan to stop TB 2006–2015. Available: <http://www.stoptb.org/globalplan/>. Accessed 16 June 2008.
2. Raviglione MC, Uplekar MW (2006) WHO's new Stop TB Strategy. *Lancet* 367: 952–955.
3. World Health Organization (2007) New technologies for tuberculosis control: A framework for their adoption, introduction and implementation. Available: http://whqlibdoc.who.int/publications/2007/9789241595520_eng.pdf. Accessed 16 June 2008.
4. World Health Organization, Special Programme for Research and Training in Tropical Diseases (2006) Diagnostics for tuberculosis. Global demand and market potential. Available: <http://www.who.int/tdr/publications/publications/tbdi.htm>. Accessed 16 June 2008.
5. Oxman AD, Lavis JN, Fretheim A (2007) Use of evidence in WHO recommendations. *Lancet* 369: 1883–1889.
6. Hill S, Pang T (2007) Leading by example: A culture change at WHO. *Lancet* 369: 1842–1844.
7. World Health Organization (2008) Moving research findings into new WHO policies. Available: <http://www.who.int/tb/dots/laboratory/policy/en/index4.html>. Accessed 16 June 2008.
8. Pai M, McCulloch M, Enanoria W, Colford JM Jr. (2004) Systematic reviews of diagnostic test evaluations: What's behind the scenes? *ACP J Club* 141: A11–A13.
9. Daley P, Thomas S, Pai M (2007) Nucleic acid amplification tests for the diagnosis of tuberculous lymphadenitis: A systematic review. *Int J Tuberc Lung Dis* 11: 1166–1176.
10. Dinnes J, Deeks J, Kunst H, Gibson A, Cummins E, et al. (2007) A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. *Health Technol Assess* 11: 1–196.
11. Flores LL, Pai M, Colford JM Jr., Riley LW (2005) In-house nucleic acid amplification tests for the detection of *Mycobacterium tuberculosis* in sputum specimens: meta-analysis and meta-regression. *BMC Microbiol* 5: 55.
12. Goto M, Noguchi Y, Koyama H, Hira K, Shimbo T, et al. (2003) Diagnostic value of

- adenosine deaminase in tuberculous pleural effusion: A meta-analysis. *Ann Clin Biochem* 40: 374–381.
13. Greco S, Girardi E, Masciangelo R, Capocetta GB, Saltini C (2003) Adenosine deaminase and interferon gamma measurements for the diagnosis of tuberculous pleurisy: A meta-analysis. *Int J Tuberc Lung Dis* 7: 777–786.
14. Greco S, Girardi E, Navarra S, Saltini C (2006) The current evidence on diagnostic accuracy of commercial based nucleic acid amplification tests for the diagnosis of pulmonary tuberculosis. *Thorax* 61: 783–790.
15. Jiang J, Shi HZ, Liang QL (2007) Diagnostic value of interferon-gamma in tuberculous pleurisy: A meta-analysis. *Chest* 131: 1133–1141.
16. Kalantri S, Pai M, Pascopella L, Riley L, Reingold A (2005) Bacteriophage-based tests for the detection of *Mycobacterium tuberculosis* in clinical specimens: A systematic review and meta-analysis. *BMC Infect Dis* 5: 59.
17. Liang QL, Shi HZ, Wang K, Qin SM, Qin XJ (2008) Diagnostic accuracy of adenosine deaminase in tuberculous pleurisy: A meta-analysis. *Respir Med* 102: 744–754.
18. Ling DI, Flores LL, Riley LW, Pai M (2008) Commercial nucleic-acid amplification tests for diagnosis of pulmonary tuberculosis in respiratory specimens: Meta-analysis and meta-regression. *PLoS ONE* 3: e1536. doi:10.1371/journal.pone.0001536
19. Martin A, Portaels F, Palomino JC (2007) Colorimetric redox-indicator methods for the rapid detection of multidrug resistance in *Mycobacterium tuberculosis*: A systematic review and meta-analysis. *J Antimicrob Chemother* 59: 175–183.
20. Mase SR, Ramsay A, Ng V, Henry M, Hopewell PC, et al. (2007) Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: A systematic review. *Int J Tuberc Lung Dis* 11: 485–495.
21. Menzies D, Pai M, Comstock G (2007) Meta-analysis: New tests for the diagnosis of latent tuberculosis infection: Areas of uncertainty and recommendations for research. *Ann Intern Med* 146: 340–354.
22. Morgan M, Kalantri S, Flores L, Pai M (2005) A commercial line probe assay for the rapid detection of rifampicin resistance in *Mycobacterium tuberculosis*: A systematic review and meta-analysis. *BMC Infect Dis* 5: 62.
23. Pai M, Flores LL, Hubbard A, Riley LW, Colford JM Jr. (2004) Nucleic acid amplification tests in the diagnosis of tuberculous pleuritis: A systematic review and meta-analysis. *BMC Infect Dis* 4: 6.
24. Pai M, Flores LL, Pai N, Hubbard A, Riley LW, et al. (2003) Diagnostic accuracy of nucleic acid amplification tests for tuberculous meningitis: A systematic review and meta-analysis. *Lancet Infect Dis* 3: 633–643.
25. Pai M, Kalantri S, Pascopella L, Riley LW, Reingold AL (2005) Bacteriophage-based assays for the rapid detection of rifampicin resistance in *Mycobacterium tuberculosis*: A meta-analysis. *J Infect* 51: 175–187.
26. Pai M, Riley LW, Colford JM Jr. (2004) Interferon-gamma assays in the immunodiagnosis of tuberculosis: A systematic review. *Lancet Infect Dis* 4: 761–776.
27. Riquelme A, Calvo M, Salech F, Valderrama S, Patillo A, et al. (2006) Value of adenosine deaminase (ADA) in ascitic fluid for the diagnosis of tuberculous peritonitis: A meta-analysis. *J Clin Gastroenterol* 40: 705–710.
28. Sarmiento OL, Weigle KA, Alexander J, Weber DJ, Miller WC (2003) Assessment by meta-analysis of PCR for diagnosis of smear-negative pulmonary tuberculosis. *J Clin Microbiol* 41: 3233–3240.
29. Steingart KR, Henry M, Laal S, Hopewell PC, Ramsay A, et al. (2007) A systematic review of commercial serological antibody detection

- tests for the diagnosis of extra-pulmonary tuberculosis. *Thorax* 62: 911-918.
30. Steingart KR, Henry M, Laal S, Hopewell PC, Ramsay A, et al. (2007) Commercial serological antibody detection tests for the diagnosis of pulmonary tuberculosis: A systematic review. *PLoS Med* 4: e202. doi:10.1371/journal.pmed.0040202
 31. Steingart KR, Henry M, Ng V, Hopewell PC, Ramsay A, et al. (2006) Fluorescence versus conventional sputum smear microscopy for tuberculosis: A systematic review. *Lancet Infect Dis* 6: 570-581.
 32. Steingart KR, Ng V, Henry M, Hopewell PC, Ramsay A, et al. (2006) Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: A systematic review. *Lancet Infect Dis* 6: 664-674.
 33. Tuon FF, Litvoc MN, Lopes MI (2006) Adenosine deaminase and tuberculous pericarditis—A systematic review with meta-analysis. *Acta Trop* 99: 67-74.
 34. Wang L, Turner MO, Elwood RK, Schulzer M, FitzGerald JM (2002) A meta-analysis of the effect of Bacille Calmette Guerin vaccination on tuberculin skin test measurements. *Thorax* 57: 804-809.
 35. Farhat M, Greenaway C, Pai M, Menzies D (2006) False-positive tuberculin skin tests: What is the absolute effect of BCG and non-tuberculous mycobacteria? *Int J Tuberc Lung Dis* 10: 1192-1204.
 36. Martin A, Panaiotov S, Portails F, Hoffner S, Palomino JC, et al. (2008) The nitrate reductase assay for the rapid detection of isoniazid and rifampicin resistance in *Mycobacterium tuberculosis*: A systematic review and meta-analysis. *J Antimicrob Chemother* 62: 56-64.
 37. Pai M, Zwerling A, Menzies D (2008) Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection—An update. *Ann Intern Med* 149: 177-184.
 38. Ling DI, Zwerling A, Pai M (2008) GenoType MTBDR assays for the diagnosis of multidrug-resistant tuberculosis: A meta-analysis. *Eur Respir J*. E-pub 9 July 2008. doi:10.1183/09031936.00061808
 39. Andersen P, Doherty TM, Pai M, Weldingh K (2007) The prognosis of latent tuberculosis: Can disease be predicted? *Trends Mol Med* 13: 175-182.
 40. Trajman A, Pai M, Dheda K, van zyl Smit R, Zwerling A, et al. (2008) Novel tests for diagnosing tuberculous pleural effusion: What works and what does not? *Eur Respir J* 31: 1098-1106.
 41. World Health Organization (2007) The global MDR-TB and XDR-TB response plan 2007–2008. Available: http://whqlibdoc.who.int/hq/2007/WHO_HTM_TB_2007.387_eng.pdf. Accessed 16 June 2008.
 42. Hopewell PC, Pai M, Maher D, Uplekar M, Ravighione MC (2006) International standards for tuberculosis care. *Lancet Infect Dis* 6: 710-725.
 43. Getahun H, Harrington M, O'Brien R, Nunn P (2007) Diagnosis of smear-negative pulmonary tuberculosis in people with HIV infection or AIDS in resource-constrained settings: Informing urgent policy changes. *Lancet* 369: 2042-2049.
 44. World Health Organization (2007) Reduction of number of smears for the diagnosis of pulmonary TB. Available: <http://www.who.int/tb/dots/laboratory/policy/en/index2.html>. Accessed 16 June 2008.
 45. World Health Organization (2007) Definition of a new sputum smear-positive TB case. Available: <http://www.who.int/tb/dots/laboratory/policy/en/index1.html>. Accessed 16 June 2008.
 46. Squire SB, Belaye AK, Kashoti A, Salaniponi FM, Mundy CJ, et al. (2005) 'Lost' smear-positive pulmonary tuberculosis cases: Where are they and why did we lose them? *Int J Tuberc Lung Dis* 9: 25-31.
 47. World Health Organization (2007) The use of liquid medium for culture and DST. Available: <http://www.who.int/tb/dots/laboratory/policy/en/index3.html>. Accessed 16 June 2008.
 48. Ridderhof JC, van Deun A, Kam KM, Narayanan PR, Aziz MA (2007) Roles of laboratories and laboratory systems in effective tuberculosis programmes. *Bull World Health Organ* 85: 354-359.
 49. World Health Organization, Special Programme for Research and Training in Tropical Diseases (2006) Request for applications: Diagnostic trial sites: Improving the diagnosis of tuberculosis through optimization of sputum smear microscopy. Available: http://www.who.int/tdr/grants/grants/trial_sites.htm. Accessed 16 June 2008.
 50. Foundation for Innovative New Diagnostics (2008) Market launch of improved fluorescence microscope scheduled for later this year. Available: http://www.finddiagnostics.org/news/press/zeiss_mar08.pdf. Accessed 16 June 2008.
 51. Cunningham J (2005) Rapid serological-based TB test evaluation: Prelim analysis. Stop TB Working Group on New Diagnostics. Available: http://www.stoptb.org/wg/new_diagnostics/assets/documents/jane_cunningham.pdf. Accessed 16 June 2008.
 52. Perkins MD, Cunningham J (2007) Facing the crisis: Improving the diagnosis of tuberculosis in the HIV era. *J Infect Dis* 196 (Suppl 1): S15-S27.
 53. Canadian Tuberculosis Committee (2007) Interferon gamma release assays for latent tuberculosis infection. An Advisory Committee Statement (ACS). *Can Commun Dis Rep* 33: 1-18.
 54. HPA Tuberculosis Programme Board (2007) Health Protection Agency Position Statement on the use of Interferon Gamma Release Assay (IGRA) tests for tuberculosis (TB): Draft for consultation. Available: http://www.hpa.org.uk/web/HPAwebFile/HPAweb_C/1204186168242. Accessed 16 June 2008.
 55. National Tuberculosis Advisory Committee Australia (2007) Position statement on interferon-gamma release immunoassays in the detection of latent tuberculosis infection, October 2007. *Commun Dis Intell* 31: 404-405.
 56. Pai M, Dheda K, Cunningham J, Scano F, O'Brien R (2007) T-cell assays for the diagnosis of latent tuberculosis infection: Moving the research agenda forward. *Lancet Infect Dis* 7: 428-438.
 57. Pai M, O'Brien R (2006) Tuberculosis diagnostics trials: Do they lack methodological rigor? *Expert Rev Mol Diagn* 6: 509-514.
 58. Small PM, Perkins MD (2000) More rigour needed in trials of new diagnostic agents for tuberculosis. *Lancet* 356: 1048-1049.
 59. Banoo S, Bell D, Bossuyt P, Herring A, Mabey D, et al. (2006) Evaluation of diagnostic tests for infectious diseases: General principles. *Nat Rev Microbiol* 4: S21-S31.
 60. Bossuyt PM, Reitsma JB, Bruns DE, Gatsonis CA, Glasziou PP, et al. (2003) Towards complete and accurate reporting of studies of diagnostic accuracy: The STARD Initiative. *Ann Intern Med* 138: 40-44.
 61. Schünemann HJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, et al. (2008) Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 336: 1106-1110.
 62. Stop TB Partnership Working Group on New TB Diagnostics (2007) Revised strategic plan of the Stop TB Partnership's Working Group on New Diagnostics. Available: http://www.stoptb.org/wg/new_diagnostics/assets/documents/Draft%20NDWG%20Strategic%20Plan%20for%20Cape%20Town%20Meeting.pdf. Accessed 16 June 2008.