COMMERCIAL SERODIAGNOSTIC TESTS

FOR DIAGNOSIS OF TUBERCULOSIS

EXPERT GROUP MEETING REPORT
22 July 2010

Printed in Switzerland

WHO/HTM/TB/2011.14

Executive summary

Background: An antibody detection-based diagnostic test in a suitable format could potentially replace microscopy and extend tuberculosis diagnosis to lower levels of health services. Dozens of commercial serological tests for tuberculosis are being marketed in many parts of the world, despite previous systematic reviews having reported variable sensitivity and specificity of these tests. Since the publication of these reviews, the evidence base has grown, methods for meta-analyses of diagnostic tests have evolved, and WHO Stop TB Department (STB) has implemented a systematic approach to evidence synthesis for TB diagnostic policy development involving systematic reviews and meta-analyses, assessment of the evidence base by Expert Group review, and implementation of the GRADE process for evidence synthesis.

Methods: An updated systematic review was commissioned to synthesize the evidence on the diagnostic accuracy of commercial serological tests for pulmonary and extrapulmonary tuberculosis. Database searches for relevant studies in all languages were updated through May 2010 and a bivariate meta-analysis that jointly models both sensitivity and specificity was performed. The findings were presented to a WHO Expert Group and the evidence assessed using the GRADE approach. This report reflects the outcomes of the Expert Group meeting and the consensus recommendations.

Results: For pulmonary tuberculosis, 67 unique studies were identified, including 32 studies from low- and middle-income countries. None of these studies evaluated the tests in children. The results demonstrated that (1) for all commercial tests, sensitivity (0% to 100%) and specificity (31% to 100%) from individual studies were highly variable; (2) using bivariate meta-analysis, for Anda-TB IgG (the most commonly evaluated test), the pooled sensitivity was 76% (95% CI 63% to 87%) in studies of smear-positive and 59% (95% CI 10% to 96%) in studies of smear-negative patients, respectively; the pooled specificity in these studies was similar: 92% (95% CI 74% to 98%) and 91% (95% CI 79% to 96%), respectively; (3) for Anda-TB IgG, sensitivity values in smear-positive (54% to 85%) and smear-negative (35% to 73%) patients from individual studies were highly variable; (4) for Anda-TB IgG, specificity values from individual studies were variable (68% to 100%); (5) a TDR evaluation of 19 rapid commercial tests, in comparison with culture plus clinical follow-up, showed similar variability with sensitivity values of 1% to 60% and specificity of 53% to 99%; (6) compared with ELISAs [60% (95% CI 6% to 65%], immuno-chromatographic assays had lower sensitivity [53%, 95% CI 42% to 64%]; and (7) in a single study involving HIV-infected TB patients, the sensitivity of the SDHO test was 16% (95% CI 5% to 34%).

For extrapulmonary tuberculosis, 25 unique studies were identified, including 10 studies from low- and middle-income countries. None of these studies evaluated the tests in children. The results demonstrated that (1) for all commercial tests, sensitivity (0% to 100%) and specificity (59% to 100%) values from individual studies were highly variable; (2) pooled sensitivity was 64% (95% CI 28% to 92%) for lymph node tuberculosis and 46% (95% CI 29% to 63%) for pleural tuberculosis; (3) for Anda-TB IgG, although the pooled sensitivity and specificity were 81% (95% CI 49% to 97%) and 85% (95% CI 77% to 92%) respectively, sensitivity (26% to 100%) and specificity (59% to 100%) values from individual studies were equally variable; and (5) in one study involving HIV-infected TB patients, the sensitivity of the MycoDot test was 33% (95% CI 19% to 39%).

Conclusions: Commercial serological tests provide inconsistent and imprecise estimates of sensitivity and specificity. There is no evidence that existing commercial serological assays improve patient-important outcomes. Overall data quality was graded as very low and the Expert Group strongly recommended that these tests not be used for the diagnosis of pulmonary and extra-pulmonary TB.

Contents

1. BACKGROUND 1

2. EVIDENCE BASE 2

2.1 EVIDENCE SYNTHESIS 2

2.1.1 Systematic review and meta-analyses 2

2.1.2 WHO/TDR laboratory-based evaluation of 19 commercially available rapid diagnostic tests for tuberculosis 3

2.2 CASE STUDY) OF ECONOMIC AND EPIDEMIOLOGICAL IMPACT OF SEROLOGIC TESTING FOR ACTIVE TUBERCULOSIS IN INDIA 3

2.3 GRADE EVALUATION 4

2.4 MEETING PROCEDURAL ISSUES 5

3. RESULTS 7

3.1 COMMERCIAL DIAGNOSTICS IN PULMONARY TB: PERFORMANCE CHARACTERISTICS 7

3.1.1 Studies included 7

3.1.2 Methodological quality 9

3.1.3 Test performance overall 12

3.1.3 Individual test performance 16

3.1.5 Results from meta-analysis 20

3.1.6 Results in HIV-infected TB patients 23

3.2 FINDINGS FROM THE WHO/TDR LABORATORY-BASED EVALUATION 23

3.3 PERFORMANCE CHARACTERISTICS OF COMMERCIAL SERODIAGNOSTICS IN EXTRA-PULMONARY TB 25

3.3.1 Studies included 25

3.3.2 Methodological quality 25

3.3.3 Test performance by site of disease 26

3.3.4 Individual test performance 27

3.3.5 Results from meta-analysis 28

3.3.6 Results in HIV-infected TB patients 29

3.4 ECONOMIC AND EPIDEMIOLOGICAL IMPACT OF TB SERODIAGNOSTICS IN INDIA 29

4. SUMMARY OF FINDINGS AND GRADE EVIDENCE PROFILES 31

4.1 PULMONARY TB 31

4.2 EXTRA-PULMONARY TB 31

4.3 STRENGTHS AND LIMITATIONS OF THE EVIDENCE BASE 32

4.4 GRADE EVIDENCE PROFILES AND FINAL RECOMMENDATIONS 33

5. RESEARCH GAPS 42

6. REFERENCES 42

List of Tables

TABLE 1. QUADAS TOOL 5
TABLE 2. COMMERCIAL SEROLOGICAL TESTS INCLUDED IN THE SYSTEMATIC REVIEW 7
TABLE 3. BIVARIATE META-ANALYSES: POOLED SENSITIVITY AND SPECIFICITY ESTIMATES BY SUBGROUP; POSTERIOR MEANS WITH 95% CREDIBLE INTERVALS IN PARENTHESES 21
TABLE 4. TDR EVALUATION: PERFORMANCE OF 19 RAPID TESTS FOR PULMONARY TUBERCULOSIS 24
TABLE 5. SHOULD COMMERCIAL SEROLOGICAL TESTS BE USED AS A REPLACEMENT TEST FOR CONVENTIONAL TESTS SUCH AS SMEAR MICROSCOPY IN PATIENTS SUSPECTED OF HAVING PULMONARY TUBERCULOSIS? 34
TABLE 6. SHOULD COMMERCIAL SEROLOGICAL TESTS BE USED AS AN ADD-ON TO CONVENTIONAL TESTS SUCH AS SMEAR MICROSCOPY IN PATIENTS SUSPECTED OF HAVING PULMONARY TUBERCULOSIS? 36
TABLE 7. DIAGNOSTIC ACCURACY OF ANDA-TB IGG 38
TABLE 8. DIAGNOSTIC ACCURACY OF ANDA-TB IGG IN STUDIES OF SMEAR-NEGATIVE PATIENTS (IE. AS AN 'ADD ON' TEST TO SMEAR MICROSCOPY) 40

List of Figures

FIGURE 1. SUMMARY OF METHODOLOGICAL QUALITY SUMMARY: STUDIES ON PULMONARY TB 11
FIGURE 2. METHODOLOGICAL QUALITY BY AUTHOR, PULMONARY TB, ALL COMMERCIAL TESTS, STUDIES OF SMEAR-NEGATIVE PATIENTS 12
FIGURE 3. FOREST PLOT OF SENSITIVITY, ALL INCLUDED STUDIES, ALL COMMERCIAL TESTS 13
FIGURE 4. FOREST PLOT OF SPECIFICITY, ALL INCLUDED STUDIES, ALL COMMERCIAL TESTS 14
FIGURE 5. FOREST PLOT OF SENSITIVITY, ALL COMMERCIAL TESTS, SMEAR-NEGATIVE PATIENTS 15
FIGURE 6. FOREST PLOT OF SPECIFICITY, ALL COMMERCIAL TESTS, SMEAR-NEGATIVE PATIENTS 15
FIGURE 7. METHODOLOGICAL QUALITY: ANDA-TB, IGG, SMEAR-POSITIVE PULMONARY TB 16
FIGURE 8. METHODOLOGICAL QUALITY: ANDA-TB, IGG, SMEAR-NEGATIVE/CULTURE-POSITIVE PULMONARY TB 17
FIGURE 9. FOREST PLOTS: ANDA-TB, IGG, SMEAR-POSITIVE PULMONARY TB 17
FIGURE 10. FOREST PLOTS: ANDA-TB, IGG, SMEAR-NEGATIVE/CULTURE-POSITIVE PULMONARY TB 18
FIGURE 11. METHODOLOGICAL QUALITY: PATHOZYME MYCO, IGG, IGM AND IGA. 18
FIGURE 12. FOREST PLOTS: PATHOZYME MYCO-G, M, AND A. G = IGG 19
FIGURE 13. METHODOLOGICAL QUALITY GRAPH FOR DETERMINER TB GLYCOLIPID ASSAY 20
FIGURE 14. FOREST PLOTS: DETERMINER TB GLYCOLIPID ASSAY 20
FIGURE 15. SUMMARY ROC PLOTS OF SENSITIVITY AND SPECIFICITY FOR ANDA-TB IGG 22
FIGURE 16. SUMMARY ROC PLOTS OF SENSITIVITY AND SPECIFICITY BY ASSAY TECHNIQUE 23
FIGURE 17. METHODOLOGICAL QUALITY: EXTRAPULMONARY TB 25
FIGURE 18. METHODOLOGICAL QUALITY: LYMPH NODE TB 26
FIGURE 19. FOREST PLOTS: ALL COMMERCIAL TESTS, LYMPH NODE TB 26
FIGURE 20. METHODOLOGICAL QUALITY GRAPH, PLEURAL TB 27
FIGURE 21. FOREST PLOTS: ALL COMMERCIAL TESTS, PLEURAL TB 27
FIGURE 22. METHODOLOGICAL QUALITY: ANDA-TB IGG, EXTRAPULMONARY TB, ALL SITES OF DISEASE 28
FIGURE 23. FOREST PLOTS: ANDA-TB IGG, EXTRAPULMONARY TB, ALL SITES OF DISEASE 28
FIGURE 24. COST FOR 1.5 MILLION TB SUSPECTS UNDERGOING SEROLOGIC TESTING IN INDIA 29
FIGURE 25. GAINS FROM INVESTMENT IN DIFFERENT DIAGNOSTIC TESTS IN INDIA 30

Annexes

Annex 1: Meeting background and objectives
Annex 2: Agenda
Annex 3: Meeting participants
Annex 4: Declarations of Interest
Annex 5: Included studies
Annex 6: Excluded studies

COMMERCIAL SERODIAGNOSTIC TESTS FOR DIAGNOSIS OF TUBERCULOSIS

1. BACKGROUND

Tuberculosis (TB) serological tests almost exclusively rely on antibody recognition of antigens of *M. tuberculosis* by the humoral immune response, as opposed to antigen recognition by the cellular immune response (e.g. interferon-gamma release assays). An accurate serological test could provide rapid diagnosis of TB and in a suitable format (e.g. point-of-care) would be particularly useful both as a replacement for laboratory-based tests and for extending TB diagnosis to lower levels of health services, especially those without on-site laboratories. Although no serological TB test is recommended by international guidelines for clinical use nor approved by the US Food and Drug Administration, dozens of distinct commercial serological tests (also referred to as commercial serodiagnostics in this document) are marketed in many parts of the world, especially in developing countries with weak regulatory systems.

Several systematic reviews and one laboratory-based evaluation on this topic have been published. Two reviews evaluating commercial tests for pulmonary TB (68 studies) and extrapulmonary TB (21 studies) found sensitivity and specificity of these tests to be highly variable.[1-3] A meta-analysis of non-commercial tests for pulmonary TB (254 datasets including 51 distinct single antigens and 30 distinct multiple-antigen combinations) identified potential candidate antigens for inclusion in an antibody detection based TB test in HIV-uninfected and -infected individuals; however, no antigen or antigen combination achieved sufficient sensitivity to replace smear microscopy.[2] A previous systematic review by Dinnes and colleagues of rapid TB diagnostic tests (literature search through 2003, seven datasets) reported pooled sensitivity and specificity values of 34% and 91%, respectively in studies meeting at least two design-related criteria.[4]

In 2005, the UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR) performed an evaluation of 19 commercially available rapid diagnostic TB tests ('rapid' defined as having a test result available in less than 15 minutes).[5] The evaluation reported that, in comparison with culture plus clinical follow-up, commercial tests provided sensitivity and specificity values of 1% to 60% and 53% to 99%, respectively.

Since the publication of previous reviews, the evidence base has grown and approaches to meta-analyses of diagnostic tests have evolved. WHO-STB and TDR therefore commissioned an updated systematic review to synthesize new evidence since 2006 on the diagnostic accuracy of commercial tests for pulmonary and extrapulmonary TB. In addition, the findings from the previous TDR evaluation are summarised below.

The systematic review and this document are limited to commercial serological tests only. In-house tests are likely to be less standardised, have less quality assurance during manufacture, and are prone to be more operator dependent. As a result, the quality issues of limitations, precision, consistency, directness and probably publication bias are expected to be more severe.

2. EVIDENCE BASE

2.1 Evidence synthesis

The systematic, structured, evidence-based process for TB diagnostic policy generation developed by WHO-STB was followed: The first step constituted a systematic review and meta-analysis of available data (published and unpublished) using standard methods appropriate for diagnostic accuracy studies. The second step involved the convening of an Expert Group to a) evaluate the strength of the evidence base; b) recommend operational and logistical considerations for mainstreaming such the methods/approaches into national TB control programmes; and c) identify gaps to be addressed in future research. Based on the Expert Group findings, the third and final step involves WHO policy guidance on the use of these tools/approaches, presented to the WHO Strategic and Technical Advisory Group for TB (STAG-TB) for consideration, and eventual dissemination to WHO Member States for implementation.

The Expert Group (Annex 1) consisted of researchers, clinicians, epidemiologists, end-users (programme and laboratory representatives), community representatives and evidence synthesis experts. The Expert Group meeting followed a structured agenda (Annex 1) and was co-chaired by WHO and a clinical epidemiologist with expertise and extensive experience in evidence synthesis and guideline development.

To comply with current standards for evidence assessment in formulation of policy recommendations, the GRADE system (www.gradeworkinggroup.org), adopted by WHO for all policy and guidelines development, was used. The GRADE approach, assessing both the quality of evidence and strength of recommendations, aims to provide a comprehensive and transparent approach for developing policy guidance. Started about 10 years ago to assess treatment interventions, the GRADE approach has recently been refined for diagnostics; however, while the latter process shares the fundamental logic of recommendations for other interventions (notably treatment), it also presents unique challenges, most often due to study limitations related to a lack of data on patient-important outcomes and impact (see below).

Recognising that test results may be surrogates for patient-important outcomes, the Expert Group evaluated diagnostic accuracy while also drawing inferences on the likely impact of these approaches on patient outcomes, as reflected by false-negatives (i.e. cases missed) or false-positives. In addition, the Expert Group was presented with an epidemiological and economic model on the cost-effectiveness and cost-benefit of commercial serodiagnostics using a case study from India, where an estimated 1.5 million TB commercial (ELISA) tests are performed every year. These tests are used mostly by the private sector (the primary source for TB care) in India, predominantly using imported TB ELISA kits at an expenditure conservatively estimated at 15 million US dollars per