

Assessing the impact of new diagnostics on tuberculosis control

THE LAST FEW YEARS have seen an unprecedented effort to develop new diagnostics for tuberculosis (TB), and a number of significant achievements in evidence-based TB diagnosis and translation of evidence into policy.^{1–3} The World Health Organization (WHO) has endorsed more than 10 new or improved tools for the diagnosis of TB since 2007, and four additional tools are currently under review.^{1,2}

The WHO process for policy formulation related to TB is a five-step process, consisting of: 1) identifying the need for a policy change; 2) reviewing the evidence; 3) convening an expert panel; 4) assessing draft policies and guidelines; and 5) formulating and disseminating policy.⁴

Central to this policy-making process is the synthesis of the available evidence on a diagnostic through systematic reviews and the application of the GRADE (grading of recommendations assessment, development and evaluation) approach to guideline development.⁵ The GRADE approach often rates diagnostic studies reporting only test accuracy (e.g., sensitivity, specificity) as low-quality evidence for policy development because the link between diagnostic accuracy and patient-important outcomes is indirect.⁵ A large proportion (>80%) of all TB diagnostic research publications are focused on test accuracy, and there is little published evidence on the impact of TB tests on patient-important outcomes.^{2,6} To fill this gap, expert opinion is sought, through WHO Expert Group Meetings, on the likely effect of diagnostic accuracy on patient-important outcomes for a given diagnostic. As a result, strong recommendations (positive or negative) have been made on the basis of moderate or low-quality evidence,⁷ and the GRADE approach permits this.⁵

There is a need to conduct diagnostics evaluations that assess the impact of new diagnostics on patient-important outcomes, including time to diagnosis, time to treatment, incremental value of new diagnostics, impact of new tests on clinician decision making, appropriateness of the treatment regimen offered on the basis of the diagnostic test result and impact of testing on treatment outcomes. While the methodologies for evidence synthesis and policy recommendation have improved greatly over recent years and the value of these activities is being recognised, most of the original research being synthesised reports only diagnostic accuracy, and policy recommendations continue to be made on the basis of moderate/low-quality evidence.^{2,3}

There are recognised challenges in assessing a number of major patient-important outcomes through a diagnostic trial.² Foremost among these are the ethical considerations that would prevent patient-

management decisions being based upon the result of the trial diagnostic. It is also recognised that most TB diagnostics will not be used alone, but in combination with clinical decisions to test, and with other diagnostics such as smear microscopy and/or culture and drug susceptibility testing. This complexity is likely to influence the patient-important outcomes. It is thus the patient-important outcomes associated with use of a particular diagnostic-intervention package that are the outcomes of interest.

The Stop TB Partnership's New Diagnostics Working Group (NDWG) has recently published a scientific blueprint for TB diagnostics development and evaluation.⁸ This blueprint presents an overview of what evidence is required for the comprehensive assessment of a diagnostic. This document makes it clear that it is not only *patient*-important outcomes that need to be measured, but also *population*-important outcomes (e.g., gender equity in access to the diagnostic) and *health systems*-important outcomes (e.g., demands of new diagnostics on human resources). Most diagnostic accuracy studies, as well as the so-called 'demonstration studies', are conducted in controlled settings with technical and financial inputs far in excess of the resources available in routine programme conditions. Valid measures of some population- and health systems-important outcomes will be difficult to obtain until the tool has been introduced into routine National Tuberculosis Programme (NTP) activities. Beyond populations and health systems, it is necessary to conduct research on the public health and epidemiological impact of introducing new diagnostics to ensure that the intervention is associated with improved case detection and cure rates and reduced TB transmission and incidence. The term 'impact' thus has different interpretations, ranging from the impact of a test on an individual patient's outcome to the epidemiological impact of widespread scale-up of a new diagnostic test in a population.

In a perfect world, we would want all the evidence available before making policy recommendations. However, collecting such a body of evidence will be time-consuming and expensive. Waiting for all the evidence could delay, by several years, the uptake of a new tool with the potential to dramatically improve TB control. The NDWG's scientific blueprint recognises that research is needed both before and after WHO endorsement and introduction of a new tool into NTP activities. Despite the WHO endorsement of several new diagnostic tools in recent years, we are unaware of any peer-reviewed publications describing the impact of new TB diagnostics on patient-, population- or

health systems-important outcomes or on the epidemiological impact. Modelling studies on the likely impact of new diagnostics have been published,^{9,10} but they will need to be supported by real-world empiric data on impact. Such studies would stimulate the scale-up of new diagnostics (if associated with positive impact) or permit the revision or amendment of WHO recommendations (if associated with negative or no impact or shown not to be cost-effective compared to alternative approaches). At present, the five-step policy making process does not include a formal stage in which the WHO reviews its recommendations in the light of experience from routine NTP practice and other impact studies.

A major obstacle to conducting impact assessments is a lack of consensus on what 'impact' really means, and what patient-, population-, health systems-, and epidemiology-important outcomes should be measured to decide on impact. There is also a lack of guidance on the methods (i.e., study designs) to be used to measure them, which methods will be most rapid and cost-effective, and who exactly should be assigned the responsibility of measuring impact (academia, industry, product development partnerships, NTPs, the WHO, technical agencies or other independent bodies).

The Impact Assessment Framework (IAF) described by Mann and colleagues in this issue of the *Journal*¹¹ is a welcome step in dealing with this gap in our knowledge and practice. Suggesting an initial definition of 'impact', the authors present, for the first time, a systematic, multi-layered approach to collecting relevant data on the overall impact of introducing new diagnostic technologies for TB. A number of methods and study designs exist that could be used to collect data in the different layers proposed in the IAF. However, some kind of consensus and guidance is necessary on what outcomes to measure and how to use the different methodologies to comprehensively collect the evidence required. Efforts are underway to convene an expert group to provide concrete guidance on assessing the impact of new diagnostics on TB control, especially for WHO-approved tests.

Although the methods and approaches for systematic reviews of diagnostic accuracy have improved greatly in the last few years,¹² those used to synthesise the data from studies on equity in health services, for example, are much less developed. In anticipation of new types of data becoming available for policy guidance, evidence synthesis experts need to develop the appropriate tools for their review, and the GRADE approach will need to evolve based on accumulated experience.

It is critical that donors and institutions supporting TB control be aware of the need to continue research within NTP activities beyond WHO endorsement of a new diagnostic tool. Significant funding will be required for such implementation research. There also needs to be a formal mechanism by which the WHO reviews post-endorsement evidence on impact, and through which endorsements or recommendations on TB diagnostics can be revised, expanded or retracted.

ANDREW RAMSAY, PhD*

KAREN R. STEINGART, MD, MPH†

MADHUKAR PAI, MD, PhD‡

*UNICEF/UNDP/World Bank/WHO

Special Programme for

Research and Training in Tropical Diseases

World Health Organization

Geneva, Switzerland

†Curry International Tuberculosis Center

University of California, San Francisco

San Francisco, California, USA

‡Department of Epidemiology,

Biostatistics & Occupational Health

McGill University

Montreal, Quebec, Canada

e-mail: madhukar.pai@mcgill.ca

Acknowledgements

All the authors are involved with the Stop TB Partnership's New Diagnostics Working Group (NDWG). AR serves as the secretariat, KRS serves as co-chair of the evidence synthesis subgroup, and MP serves as co-chair of NDWG. MP is a consultant for the Bill & Melinda Gates Foundation. The views and opinions expressed by the authors are their own and do not necessarily reflect those of the Stop TB Partnership or the institutions for whom they work.

References

- Wallis R S, Pai M, Menzies D, et al. Biomarkers and diagnostics for tuberculosis: progress, needs, and translation into practice. *Lancet* 2010; 375: 1920–1937.
- Pai M, Minion J, Steingart K, Ramsay A. New and improved tuberculosis diagnostics: evidence, policy, practice, and impact. *Curr Opin Pulm Med* 2010; 16: 271–284.
- Pai M, Ramsay A, O'Brien R. Evidence-based tuberculosis diagnosis. *PLoS Med* 2008; 5: e156.
- World Health Organization. Moving research findings into new WHO policies. Geneva, Switzerland: WHO, 2008. <http://www.who.int/tb/dots/laboratory/policy/en/index4.html> Accessed September 2010.
- Schunemann H J, Oxman A D, Brozek J, et al. Grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ* 2008; 336: 1106–1110.
- Brunet L, Minion J, Lienhardt C, Pai M. Mapping the landscape of tuberculosis diagnostic research. *Am J Respir Crit Care Med* 2010; 181: A2255.
- WHO Expert Group meeting reports on sputum smear microscopy and non-commercial culture and bacteriophage-based assays. Geneva, Switzerland: WHO, 2009. http://www.who.int/tb/dots/laboratory/policy_statements/en/index.html Accessed October 2010.
- Stop TB Partnership's New Diagnostics Working Group/World Health Organization. Pathways to better diagnostics for tuberculosis: a blueprint for the development of TB diagnostics. Geneva, Switzerland: World Health Organization, 2009.
- Abu-Raddad L J, Sabatelli L, Achterberg J T, et al. Epidemiological benefits of more-effective tuberculosis vaccines, drugs, and diagnostics. *Proc Natl Acad Sci USA* 2009; 106: 13980–13985.
- Dowdy D W, Chaisson R E, Maartens G, Corbett E L, Dorman S E. Impact of enhanced tuberculosis diagnosis in South Africa: a mathematical model of expanded culture and drug susceptibility testing. *Proc Natl Acad Sci USA* 2008; 105: 11293–11298.
- Mann G, Squire S B, Bissell K, et al. Beyond accuracy: creating a comprehensive evidence base for TB diagnostic tools. *Int J Tuberc Lung Dis* 2010; 14: 1518–1524.
- Leeflang M M, Deeks J J, Gatsonis C, Bossuyt P M. Systematic reviews of diagnostic test accuracy. *Ann Intern Med* 2008; 149: 889–897.