Special article

Assessing new strategies for TB diagnosis in low- and middle-income countries

Afrânio Kritski\textsuperscript{a,b,∗}, Paula I. Fujiwara\textsuperscript{c}, Maria Armanda Vieira\textsuperscript{b}, Antonio Ruffino Netto\textsuperscript{d}, Martha Maria Oliveira\textsuperscript{a,b}, Gisele Huf\textsuperscript{b,e}, S. Bertel Squire\textsuperscript{f}

\textsuperscript{a} Rede TB, Rio de Janeiro, RJ, Brazil  
\textsuperscript{b} Academic Tuberculosis Program, Medical School, Universidade Federal do Rio de Janeiro, Rio de Janeiro, RJ, Brazil  
\textsuperscript{c} International Union Against Tuberculosis and Lung Disease, Paris, France  
\textsuperscript{d} Medical School, Universidade de São Paulo, Ribeirão Preto, SP, Brazil  
\textsuperscript{e} Liverpool School of Tropical Medicine, Liverpool, United Kingdom  
\textsuperscript{f} National Institute for Quality Control, Fundação Oswaldo Cruz, Rio de Janeiro, RJ, Brazil

\textbf{Article history:}  
Available online 5 March 2013

\textbf{Keywords:}  
Tuberculosis  
Diagnostics  
Impact  
Evidence  
Policy  
Innovation  
Regulatory agency

\textbf{Abstract}

In this manuscript, we report the current situation of tuberculosis globally and in Brazil, the need for new strategies toward tuberculosis control, focusing on new diagnostic technologies. Critical comments are given on the state of the art regarding the evaluation of new health technologies, degree of scientific evidence needed, evaluation of clinical impact, cost-effectiveness of incorporation into the health system and the social impact.

© 2013 Elsevier Editora Ltda. Este é um artigo Open Access sob a licença de CC BY-NC-ND

\textbf{Introduction}

Tuberculosis (TB) is a major public health problem in the world, with incidence differing by regions. It is estimated that in 2010, approximately 8.8–9.2 million new TB cases arose, of which 4 million did not receive treatment. Treatment outcomes have been poor in countries that have not adopted the DOTS strategy proposed by the World Health Organization (WHO) in 1993, and where effective TB control actions were implemented, have been hampered by lack of political commitment and/or disorganization of the health system, associated to poverty, social exclusion, and the burden of human immunodeficiency virus (HIV) or multidrug-resistant (MDR) or extensively drug-resistant (XDR) TB.\textsuperscript{1}

As one of the emerging middle-income countries, Brazil ranks 19th among the 22 countries which account for 80% of all TB cases worldwide and 108th in incidence.\textsuperscript{2,3} According to the Ministry of Health, in 2010, 71,000 cases of TB were reported in the National Reporting Case System (SINAN), for which the cure rate was 66.4% and the defaulting rate was 11% (which ideally should be below 5%). The percentage of cases on Directly Observed Therapy (DOT) was only 38%. It is estimated that among TB patients co-infected with HIV, the cure rate was...
55.7% and death rate was 23% (5.7 times greater than for HIV-seronegative patients). In 2008, among the twelve measures for collaborative TB-HIV control proposed by WHO, only three were adopted by both programs of the Ministry of Health in Brazil.

In 2006, WHO’s Global Plan to Stop TB⁴ expanded the DOTS strategy (renaming it the Stop TB Strategy). This plan prioritized the strengthening of the health system; the pursuit of public–public and public–private partnerships; social mobilization and the reevaluation of the academic role (forgotten since the 1970s for its contribution to TB control efforts); in order to seek new ways to increase the effectiveness of DOTS, improve access to services, and increase the diagnosis and cure rate of TB, including those with MDR-TB and co-infected with HIV.⁵

Despite the fact that acid-fast bacilli (AFB) sputum smear microscopy has a low sensitivity (60%), it remains one of the most frequently used tests for the diagnosis of pulmonary TB in low-income countries. In addition, in HIV-infected patients, children or patients with other immunosuppressive diseases, the sensitivity of smear microscopy is much lower (<30%). In most high burden countries, in practice, the mycobacteria culture, for which diagnostic sensitivity is higher (80-85%) it is performed on solid Lowenstein–Jensen (LJ) medium but is used only in selected clinical cases (cases of treatment failure, patients with persistent smear-negative or extra-pulmonary forms). The major problem in the use of LJ for diagnosing tuberculosis is the long incubation time (4–6 weeks), and since the drug sensitivity testing (DST) is performed from the culture and not from the clinical specimen, several additional weeks are required to obtain the results.⁷

In HIV-positive patients and in children, the strategy proposed by WHO to prioritize the assessment of respiratory symptoms (cough for more than 2–3 weeks) to search for pulmonary TB has been inadequate. Recently, Cain et al.⁸ and Marais et al.⁹ proposed an innovative approach regarding the evaluation of clinical scores and radiological diagnosis of pulmonary TB in different epidemiological settings, for adults and children. They highlighted the urgent need for evaluation of new diagnostic approaches that promote greater impact on the region(s) most affected by coinfection TB-HIV and/or MDR-TB.¹⁰ Additionally, due to the absence of laboratories capable of routinely performing culture and DST; there are few reliable data on MDR- or XDR-TB among the 22 high burden TB countries. In 2010, it has been estimated that there were 650,000 cases of MDR-TB, with cure rates less than 60%, higher rates of morbidity/mortality and increased treatment costs. Of these MDR-TB cases, only 8.5% were diagnosed and an even smaller proportion of them had access to appropriate treatment.¹¹ Therefore, the evaluation of new diagnostic technologies is needed urgently, but different detection strategies that also include an analysis of factors associated with access, linked to the patient and/or health system must be considered.¹²

In Brazil, in 2010, cultures for mycobacteria were performed in only 30% of retreatment cases, even though national policy is to have them done for all such cases. In addition, only 22% of TB cases with a positive serology for HIV and 36% of all prisoners had a culture done, even though the same national policy applies. These data demonstrate the difficulty in Brazil of providing adequate coverage for the diagnosis of TB. Due to the low coverage for culture (DST was performed only in 30.7% of the cases that are supposed to have it according to Brazilian policy), the number of MDR-TB cases in the country is probably deeply underestimated. In these patients, the cure rate does not exceed 65%, those who default from treatment is greater than 20% and the proportion of those who die is over 12%. Therefore in Brazil, in order to increase access, equity, the quality of TB diagnosis, reduce defaulting from treatment and provide more social services or health care of their families, it is urgent to prioritize collaborative activities between the TB and AIDS control programs, and also with primary health care (PHC) programs at the three government levels: federal, state and municipal, along with the support of the academic community and civil society. Such activities will enable the development and evaluation of the impact of the adoption of new strategies for TB, TB-HIV, and MDR-TB.

### Evaluation of new technologies and strategies in the control of TB, TB-HIV, MDR-TB

Enabling and promoting research are key components of the Stop TB Strategy, and should be pursued vigorously. Clearly, new and better technologies for the prevention, diagnosis, treatment and care of active and latent TB and its associated conditions and complications are needed. However, this will not be sufficient. Innovative approaches also must continue to be developed to ensure equitable access to these technologies.¹⁴ In addition, those approaches involving operational and cost-effectiveness evaluation should be adapted and adjusted based on the epidemiological context of the local health system.

Most of the innovations cannot be translated into effective action without careful local planning and adaptation. Operational research needs to be conducted in a well-planned manner that is distinct from the routine monitoring carried out to assess the epidemiological situation of the national and local health system. Only then can proper arrangements of different applications for local interventions be identified. However, there are many hurdles to this essential step in the chain of events from basic research to practical application. National TB, AIDS and PHC programs generally have limited capacity to conduct operational research and often do not have a development agenda for it. So there is a need for guidance on what issues to investigate, how to do it and how to strengthen the capacity for operational research. There are several ways and steps to assess the potential value of a diagnostic test/strategy for clinical use, but the choice of the model depends on proper determination of the question that needs to be answered.

### The current focus on accuracy

The first question that arises for a new test/strategy is whether the results obtained in individuals with active TB differ from results found in healthy individuals. To answer this question, the study should be conducted in individuals with known disease and in those who are healthy, checking the results in
each of these groups. This study begins with the evaluation of the sensitivity, specificity and accuracy of a new diagnostic test for a given disease. This type of study does not imply a diagnostic action, but it is an early stage of the process, usually with greater involvement of researchers in the basic and applied basic research arena, linked to research laboratories in universities, research institutes and industry. The next step is to question whether the new test/strategy is able to distinguish individuals with active TB and those TB suspects (with no TB) seeking medical care. In this evaluation phase in more traditional clinical research centers, this is based on comparison of the new test/strategy with a reference test or gold standard, to derive measures of diagnostic accuracy, such as sensitivity and specificity. These studies provide scientific evidence and consist almost entirely of reviews of new diagnostic tests/strategies for TB in the published literature and used as evidence in recent years. These studies have been conducted in the clinical research centers, in universities, and research institutes usually linked to industry support.

The GRADE system, systematic reviews and their limitations

Making adequate decisions on health, including on the diagnostic method to be used, means not only putting on a scale the available evidence about the risks and benefits of alternative strategies, but also depends on the confidence that their findings can inspire. This recognition led to the emergence of a series of formal systems to categorize the quality of scientific evidence (i.e. from very high to very low), among which was the GRADE (Grades of Recommendation Assessment, Development and Evaluation) system, adopted by WHO in 2007. GRADE was initially proposed in Canada for new recommendations for TB control and became the most widely used system of evaluation in the last decade in developed countries. This system consists of two main parts. The first is the degree to which scientific evidence is derived from studies of models that are more able to prevent the occurrence of systematic errors or bias: clinical trials and systematic reviews of clinical trials are the best evidence; observational studies (cohort and case-control) and then, finally, studies without a comparison group (case studies) and expert opinions are of lesser value in this hierarchy. The second part is dictated by the strength of the recommendation compiled from studies that should be used in changes of clinical guidelines and standards manuals. However, in most instances usually when the GRADE system is adopted, only high-quality studies have been used.

Although recommendations regarding the use of diagnostic tests share the same fundamental logic with recommendations for therapeutic interventions, they have unique characteristics and challenges, brought into sharp focus when applying the GRADE system to the area of testing and diagnostic strategies. In evaluating the results of new diagnostic tests/strategies, it has been described that low accuracy greatly limits the clinical value of a test/strategy. However, a test with high sensitivity and/or specificity alone does not guarantee an improvement in outcomes considered important for doctors and patients. In practice, clinicians want to know how a test/strategy is able to affect clinical judgment to be used in decision making about clinical procedures. What the doctor wants to know is if those tests/strategies do better in terms of clinical outcomes. Does the new diagnostic strategy promote more appropriate therapeutic interventions? Very rarely is this benefit quite clear in the literature. Generally, as in the case of tests for early detection of asymptomatic disease, this evaluation can only be done accurately by tracking individuals who were randomized to undergo the test of interest and another (or no) test.

After two decades of experience, it was observed that the outcomes studied from systematic reviews and meta-analyses in health of most clinical trials, did not respond to the key issues to help deciding whether or not the incorporation of technology to the health system is indicated. Additionally, those studies were conducted in clinical research centers in specific populations that are not representative of the general one. Moreover, these studies have generally not included evaluations of cost effectiveness. At the end of the last decade, a distinction between explanatory and pragmatic trials began to emerge. Explanatory clinical trials seek to answer questions of efficacy, whether and how an intervention works. On the other hand, pragmatic trials are conducted to support decision-making in health care, and therefore are conducted in conditions very close to those provided within the routine health services, in patients who are very similar to those who will need the treatment in the future. In a recent systematic review, which included 168,000 randomized controlled trials conducted in the period 1976–2002, it was found that only 95 (0.05%) met the criteria of pragmatic clinical trials. The authors emphasized the urgent need to prioritize the achievement of pragmatic trials that may also answer questions regarding the applicability of new technology in the health system and not just the issues of efficacy used in explanatory clinical trials, whose main purpose would be to obtain industry product registration followed by the regulatory agencies for their marketing in the private system. Additionally, trials should be designed and reported in such a way that users of the results can make meaningful judgments about applicability to their own context. As the pragmatic-explanatory distinction comprises a continuous spectrum, not an either/or dichotomy of the extremes, we suggest to follow the innovative approach proposed by Thorpe et al. using the pragmatic explanatory continuum indicator summary (PRECIS) with the identification of key domains that distinguish pragmatic from explanatory trials.

The limitation of regulatory approval for private markers and its impact on innovation in the National Guidelines

In the area of diagnostic tests, the situation is not different. New diagnostic tests evaluated with funding from industry research centers through clinical tests of accuracy have been sufficient for approval of registration for marketing by regulatory agencies (the United States: Food and Drug Administration – FDA; in Europe Medicine Agency – EMA and in Brazil, the Agência Nacional de Vigilância Sanitária – ANVISA). These tests are implemented in the private system as they become available, based on experiences with a limited number
of cases; i.e. a subjective expectation of its usefulness. As a result of lobbying by industries, biomedical companies and the media influenced by marketing in the current economic system, the logic has been that the individual seeking care in the health system must be offered all the technological innovations produced with some scientific evidence but without systematic assessment of their impact on the health system.

By 2007, in the vast majority of countries, with only the universal availability of sputum smear microscopy, about 20–30% of patients treated in low-income countries were treated for TB without bacteriological confirmation. In 2007, in order to respond more effectively to the co-infection with TB and HIV and MDR-TB globally, WHO recommended new TB diagnostic technologies, such as the use of liquid culture for detection of *Mycobacterium tuberculosis* and DST, based on a review of available scientific evidence and expert consultation. In 2008, WHO recommended the use of molecular tests for rapid screening of patients suspected of drug-resistant TB. This recommendation was based on systematic reviews, expert opinion and preliminary results of effectiveness obtained in demonstration projects (phase III/IV) clinical research centers. Such tests should only be used in respiratory specimens smear-positive or culture-positive for mycobacteria. Pai et al. conducted a systematic review of studies evaluating new diagnostic tests for TB, and demonstrated the lack of methodological rigor in most studies. The authors emphasized that biased results of poorly designed studies could lead to the adoption of early diagnostic tests that may have little or no benefit. In recent years, guidelines were issued in the standardization of model studies in the area of infectious diseases, and to evaluate the accuracy of the new diagnostic tests, evaluations of different algorithms (not just individual tests, but also their relative contributions to the system health care); their incremental value, impact on clinical practice over the choices of decision making, studies of cost-effectiveness under routine conditions, and the impact of new tools for the patient and society should all be included. In 2008, Wei et al. analyzed the data published in the literature, concluding that the new recommendations included in the national recommendations or guidelines for TB in developed countries, used the best scientific evidence grade (GRADE), and were rapidly incorporated into books for review and/or Clinical Guidelines in middle-income countries with few changes, without adjusting them to the local needs of the health of each country and that most of the clinical guidelines held in these countries did not have certification by the Appraisal of Guidelines for Research & Evaluation (AGREE).

A recent survey conducted in sixteen high TB burden countries on the adoption of seven new tools for the diagnosis of TB approved by WHO since 2007, confirmed the deployment of new diagnostic tests in TB control policies in half of the countries. Interestingly, none of the seven countries carried out an impact assessment (IA) before incorporating these new technologies, as proposed by WHO and the STOP TB Partnership.

In the evaluation of new diagnostic tests for drug-sensitive and drug-resistant TB, tests using liquid culture (i.e. Bactec®960) or molecular tests (EMTD-GenProbe, Amplicor®Roche, Biometrix, MTBDR plus-Life Science) have been recommended by the United States FDA and the corresponding body in the European Union (and marketed there). These molecular tests have been commercialized in the private sector in countries with an intermediate level of economic development (as in Brazil, since 2009). Although there is no report in the literature of pragmatic clinical trials and cost-effectiveness in the use of molecular tests in the diagnostic approach of drug-resistant TB and TB in developing countries, the test Xpert MTB/RIF was recommended by WHO in December 2010.

Xpert MTB/RIF is a fully automated molecular testing device with an integrated processing model designed to purify, concentrate, amplify and identify the rpoB target sequences for the diagnosis of rifampin resistance, providing results in 120 min from sputum samples without requiring the presence of an expert in molecular biology. The results obtained by demonstration studies (phase III) confirm the specificity for the diagnosis of TB and drug resistance to rifampicin. The 72% sensitivity of the test with one sample from sputum smear-negative patients is similar to that observed with other molecular tests such as the Roche Amplicor and the EMTD GenProbe. Despite the high specificity (>95%) of Xpert MTB/RIF for detection of rifampicin resistance, the test should be used only for clinical decision where the prevalence of resistant TB is more than 15%. However, this test can be decentralized to health facilities at the secondary level, as it does not require a molecular biology laboratory for its implementation.

Impact assessment framework as evidence for scale up

Globally, in recent years, consensus has developed among policymakers that middle-income countries like Brazil should lead processes in the field of Health Technology Assessment and impact analysis for the incorporation of new technologies, focusing on the ability of new technology to improve or maintain health. Diagnostic tests should not escape this principle.

Until recently, prolonged economic growth and democratic stability seen in industrialized nations allowed increasing investment and improvement in the fields of health and education. Additionally, improving public management allowed the anticipation of more care, better performance, and attention to health institutions. In 2004, in Brazil, a Law on Technological Innovation (Law 10973 of 02/12/2004), regulated by Decree No. 5563 of 11.10.2005 was approved. There is visible progress in investment in research in Brazil, with growing space for initiatives in the field of innovation, an area in which the country is still behind the progress. The public–private partnerships, the close interaction of companies with the university arena seems to indicate a welcome progress.

In recent decades, the health biotechnology sector in Brazil has made considerable progress toward innovation through institutes controlled by the government and the private sector. However, despite changes in the national scene in the public institutions of higher education, researchers have worked in a very fragmented manner, especially in health, without a unity of expression and an effective and coordinated strategy to address common challenges. With respect to the subject of publications, in middle-income countries, the predominance of basic research studies in the areas of vaccines, immunology,
Fig. 1 – Pathway for the development of tuberculosis diagnostics, from needs assessment to delivery.

In general, in middle-income countries, the participation of civil society and financial support toward investments in health care, integrating education, research and extension, especially in universities, is very welcome. Recently in Brazil, lack of technological innovation and its interface with postgraduate programs were considered as one of the major challenges in the 2012–2020 Plan prepared by Ministry of Education. On the other hand, within the Ministry of Health in 2008, the Commission for the Incorporation of Technology (CITEC) of the Ministry of Health, linked to the Department of Science, Technology and Strategic Inputs/Ministry of Health was created, according to decree No. 2587 of October 30, 2008.

The incorporation of new, or removal of antiquated equipment/technology, drugs, biological supplies (diagnostic tests) in the Unified Health System (SUS) will occur only when the assessment of such products addresses the following issues: (a) impact of technology on health, and (b) the technological relevance established through studies of evaluation of health technologies, such as technical-scientific, systematic reviews, meta-analysis, economic studies and pragmatic trials.

To assist in the discussion on this issue, the International Union Against Tuberculosis and Lung Disease and its regional partners, including the Brazilian Network of Tuberculosis Research (REDE-TB) and the Medical School of the Federal University of Rio de Janeiro recently published a new proposal for a platform to evaluate the impact of new diagnostic technologies for tuberculosis. Using this new platform of technology assessment described in Fig. 1, operational research can be prioritized through randomized pragmatic clinical trials, cost analysis and scale-up, and also include issues related to equity, access, qualitative studies with users, health professionals, managers, and representatives of local and international industries to identify barriers or facilitators for the incorporation of new technology in different health systems. Using this platform, it is intended that the discussion on the merger or dissemination of a new test/strategy for TB diagnosis can at least answer the simple question: will it be better for patients and/or the current health system in the country?

Individual health values can no longer be the only criterion on which decisions are based in industry. It is also necessary to take into account the social cost, the individual acceptance or time spent by the end-user and/or health care professional. As mentioned earlier, new diagnostic tools are often immediately incorporated into the routine of services as soon as they are approved for marketing by regulatory agencies, based on their performance through the analysis of sensitivity, specificity or “receiver operating curves”.

More recently, Cobelens et al. proposed a pathway for evaluating new tuberculosis diagnostics separating the technical and programmatic policy recommendations. This pathway allows all stakeholders to distinguish between the statement about whether a particular test has sufficient potential to be used in tuberculosis control, and the statement about how, where, and under which conditions this test should be implemented. It acknowledges that these statements require different types of evidence that can only be collected in a phased manner. Clarity for policy makers, tuberculosis program managers, and donors about this 2-step process will allow them to choose between scaling up after technical policy recommendation(s) only (early adoption by regulatory agencies as described above) versus waiting until the WHO has issued a detailed programmatic policy recommendation. Countries that adopt a new technology can play an important role in collecting the evidence needed in the stage before scale-up.

Closing remarks

The pragmatic approach presented in this manuscript indicates that it is not appropriate to conduct an investigation on the incorporation of new technologies only in a “purely experimental” manner in clinical research centers. Research and practice clinical processes become intertwined and the main outcomes to be considered are the patient’s health and actions creating a more effective health system in which the new technology will be incorporated. In this new scenario, it is essential that the academic biomedical areas reformulate their undergraduate curricula to include courses that address the development of new technologies, including the assessment of clinical impact, and economic and social incorporation of these new technologies into the current health system that will influence the future practice of
their graduate students. In parallel, only through collaborative activities between academics, health service providers (public or private), producers of raw materials, laboratories and representatives of civil society will it be possible to conduct such studies under routine conditions in demonstration areas to enable an analysis appropriate to the relevance of the incorporation of new technologies in the country.

**Conflict of interest**

The authors have no conflict of interest to declare.

**REFERENCES**

36. Van Rie A, Page-Shipp L, Scott L, Sanne I, Stevens W. Xpert (R) MTB/RIF for point-of-care diagnosis of TB in high-HIV


