Pathways to better diagnostics for **tuberculosis**

A blueprint for the development of TB diagnostics

By the New Diagnostics Working Group of the Stop TB Partnership

Stop TB Partnership

World Health Organization

Pathways to better diagnostics for tuberculosis

A blueprint for the development of TB diagnostics

By the New Diagnostics Working Group of the Stop TB Partnership

Stop TB Partnership

World Health Organization

Contents

⊙ **ANNEXES -** see CD in back cover

The principles of current TB diagnostic tools

1. **Optimizing TB smear microscopy**
2. **Rapid solid and liquid culture**
3. **Antigen detection tests for diagnosis of active TB**
4. **Antibody detection**
5. **T-cell-based interferon-gamma release assays**
6. **Nucleic acid amplification tests**
7. **Molecular drug resistance testing**
8. **Phage-based tests**
9. **Nose technologies**
10. **References and glossary**

Contributors

This volume, published at the initiative of the New Diagnostics Working Group of the Stop TB Partnership, would have been impossible without the invaluable contributions of leading scientists representing a wide spectrum of experience in TB, and specifically in the development of TB diagnostics. They work in academia (universities and research institutions), industry and NGOs. Often, they have experience across all sectors.

The task of coordinating the contributions from some 30 different authors, undertaking the technical editing from a science perspective, guiding the review process and daily keeping the project on course has been assumed by Russell Dacombe of Liverpool Associates in Tropical Health. His broad knowledge of the field proved indispensable in binding together the very specific contributions provided by the authors. The New Diagnostics Working Group is deeply indebted to him for his perseverance and dedication to the task. He was assisted by Tony Murdoch and Martine Guillerm working from WHO/TDR in Geneva.

Contributors are listed in alphabetical order.

Catharina Boehme , FIND (Foundation for Innovative New Diagnostic), Geneva, Switzerland

Patrick Bossuyt, Clinical Epidemiology, Biostatistics and Bioinformatics Department, University of Amsterdam, The Netherlands

Antonio Campos-Neto, Department of Cytokine Biology, Forsythe Institute, Boston, USA

Luis Cuevas, Accessible Quality-assured Diagnostics, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR),World Health Organization, Geneva, Switzerland

Russell Dacombe, Liverpool Associates in Tropical Health, Liverpool, United Kingdom

Mark Doherty, Infectious Disease Immunology, Statens Serum Institut, Copenhagen, Denmark

Martine Guillerm, Accessible Quality-assured Diagnostics, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (TDR),World Health Organization, Geneva, Switzerland and Secretary, Stop TB Partnership's New Diagnostics Working Group

Christy Hanson, Office of Health, Infectious Diseases and Nutrition, U.S. Agency for International Development, Washington, USA

Carole Jefferson, Becton Dickinson Diagnostics, USA

Paul Klatser, KIT Biomedical Research Department, Royal Tropical Institute, Amsterdam, The Netherlands

Arend Kolk, Polymer-Analysis Group, Van't Hoff Institute for Molecular Sciences, University of Amsterdam, The Netherlands

Heather Alexander, FIND (Foundation for Innovative New Diagnostic), Geneva, Switzerland

Suman Laal, Manhattan Veterans Affairs Medical Center, New York University, USA

Gillian Mann, Liverpool School of Tropical Medicine, United Kingdom

Dermot Maher, Research Unit on AIDS, MRC/UVRI Uganda

Ruth McNerney, TARGETS international research consortium on communicable diseases of vulnerability and poverty, London School of Tropical Medicine, United Kingdom

Dick Menzies, Respiratory Division, McGill University Health Centre, Montreal, Canada

David Moore, Imperial College London, Universidad Peruana Cayetano Heredia (UPCH), Lima, Peru

Richard O'Brien, FIND (Foundation for Innovative New Diagnostics), Geneva, Switzerland

Madhukar Pai, McGill University, Montreal, Canada and Co-Chair, Stop TB Partnership's New Diagnostics Working Group

Juan Carlos Palomino, Mycobacteriology Unit, Institute of Tropical Medicine, Antwerp, Belgium

Mark Perkins, FIND (Foundation for Innovative New Diagnostics), Geneva, Switzerland

Françoise Portaels, Mycobacteriology Unit, Institute of Tropical Medicine, Antwerp, Belgium

Andrew Ramsay, UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases,Geneva, Switzerland and Secretary, Stop TB Partnership's New Diagnostics Working Group

Leen Rigouts, Mycobacteriology Unit, Institute of Tropical Medicine, Antwerp, Belgium

Giorgio Roscigno, FIND (Foundation for Innovative New Diagnostics), Geneva, Switzerland and Co-Chair, Stop TB Partnership's New Diagnostics Working Group

Max Salfinger, Bureau of Laboratories, Florida Department of Health, USA

Thomas M Shinnick, Division of Tuberculosis Elimination, National Center for HIV/AIDS, Viral Hepatitis, STD, and TB Prevention, Centers for Disease Control and Prevention, USA

Tido von Shoen Angerer, Access Campaign for Essential Medicines, Médecins sans Frontières, Geneva, Switzerland

Hojoon Sohn, FIND (Foundation of Innovative New Diagnostics), Geneva, Switzerland

Bertie Squire, Clinical Tropical Medicine, Liverpool School of Tropical Medicine

Francis Varaine, TB Working Group, Médecins sans Frontières, Geneva, Switzerland

Glossary

ADA	antigen detection assay
AFB	acid-fast bacilli
AIDS	acquired immunodeficiency syndrome
ATP	adenosine triphosphate
BCG	bacille Calmette-Guerin (current TB vaccine)
BSL	biosafety level
CCD	charged couple device
CDC	Centers for Disease Control and Prevention
CE	Communauté Européenne (European Community)
CFP	culture filtrate protein
CI	confidence interval
CMOS	complementary metal oxide semiconductor
CPC	cetylpyridium chloride
DALY	disability-adjusted life years
DEC	disease-endemic country
DMS	differential mobility spectrometry
DNA	deoxyribonucleic acid
DOTS	directly observed treatment short course
DR	drug resistant
DST	drug susceptibility testing
E	ethambutol
ECD	electrochemical detection
ELISA	enzyme-linked immunosorbent assay
ESAT	early secreted antigenic target
EU	European Union
FDA	(US) Federal Drug Administration
FIND	Foundation for Innovative New Diagnostics
FM	fluorescence microscopy
GC	gas chromatography
GC-MS	gas chromatography-mass chromatography

GCP	good clinical practice
GLI	Global Laboratory Initiative
GRADE	Grading of Recommendations, Assessment, Development and Evaluation
H	isoniazid
HIV	human immunodeficiency virus
HPLC	high-pressure liquid chromatography
HTA	health technology assessment
ICH	international conference on harmonization
IFN	interferon
IgG	immunoglobulin G
IgM	immunoglobulin M
IGRA	interferon gamma release assay
IMD	ion mobility detection
IP	intellectual property
IQC	internal quality control
IRB	institutional review board
ISO	International Organization for Standardization
IU	international unit
IUATLD	International Union Against Tuberculosis and Lung Disease
IVD	in vitro diagnostic medical device
LAM	lipoarabinomannan
LAMP	loop-mediated isothermal amplification
LCR	ligase chain reaction
LED	light-emitting diode
LJ	Lowenstein-Jensen (medium)
LOD	logarithm of the odds
LPA	line-probe assay
LRP	luciferase reporter phage
LTBi	latent tuberculosis infection
MTB	*Mycobacterium tuberculosis*

MDR-TB	multidrug-resistant tuberculosis		**ROC**	receiver operating characteristic
MEMS	micro-electro mechanical system		**RPA**	recombinase polymerase amplification
MODS	microscopic observation drug susceptibility		**RRDR**	rifampicin resistance determining region
MOH	ministry of health		**S**	streptomycin
MOTT	mycobacteria other than tuberculosis		**SAW**	surface acoustic wave
MSH	management sciences for health		**SD**	serodiagnostic test
MTB	*Mycobacterium tuberculosis*		**SDA**	strand displacement amplification
NAA	nucleic acid amplification		**SELDI-TOF-MS**	surface enhanced laser desorption/ionisation time of flight mass spectrometry
NAAT	nucleic acid amplification test		**SOP**	standard operating procedure
NAP	p-nitro-alpha acetylamino-beta-hydroxy-propiophenone		**SPR**	surface plasmon resonance
NPV	negative predictive value		**STAG-TB**	Strategic and Technical Advisory Group for Tuberculosis
NTM	non-tuberculosis mycobacteria		**SVM**	support vector machine
NTP	national tuberculosis programme		**TB**	tuberculosis
OQC	outgoing quality control		**TBSA**	tuberculostearic acid
PAL	(Japan's) Pharmaceutical Affairs Law		**TDR**	UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases
PAS	para-aminosalicyclic acid		**TLA**	thin-layer agar
PBMC	peripheral blood mononuclear cell		**TMA**	transcription mediated amplification
PCR	polymerase chain reaction		**TST**	tuberculin skin test
PNB	phenyl-t-butyl nitrone		**TTD**	time to detection
POC	point of care		**UK**	United Kingdom
PPD	purified protein derivative		**US**	United States of America
PPM-DOTS	public-private mix directly observed treatment strategy		**V&V**	verification and validation
PPV	positive predictive value		**VOC**	volatile organic compound
PTB	pulmonary tuberculosis		**WHO**	World Health Organization
QA	quality assurance		**XDR-TB**	extensively drug resistant tuberculosis
QALY	quality-adjusted life years		**Z**	pyrazinamide
QC	quality control		**ZN**	Ziehl-Neelsen
R	rifampicin			
R&D	research and development			
RD	region of difference			
RIF	rifampicin			
RNA	ribonucleic acid			

Executive summary

This publication is intended to help tuberculosis diagnostics researchers work more effectively with academics, officials and industry professionals. It offers a structure to guide those involved in diagnostic development through the different phases of development – from the discovery of new techniques and tools through to their delivery in previously neglected markets. It is designed to help identify the most promising TB tests, to push them toward alignment with the needs and requirements of the areas where tuberculosis is most prevalent, and to help determine why some are held up in development. The blueprint also is intended to help boost coordination and collaboration across the diagnostic research and development landscape, while leading to greater success in advancing new tests by clarifying that landscape and the various development steps. Finally, considering the paucity of national regulations regarding the quality of diagnostics and the need for a strong evidence base to support health policy decisions, this blueprint seeks to inform and advance independent evaluation and assessment of the new tools' likely impacts based on scientifically sound methodology, in addition to well-defined standards of practice for study design.

Increasing the speed, effectiveness and accuracy of diagnostic tests is central to the goal of rolling back the global tuberculosis epidemic that afflicts nearly a third of the world's population. Though recent and ongoing advancements in drug therapies offer great promise for saving lives, the unfortunate fact is that new medicines have limited value in the places where TB is rampant. This is because the principal diagnostic tools used in developing countries for determining whether someone has tuberculosis – microscopic examination of stained sputum and chest X-ray – are simply not accurate enough to identify many TB infections. In addition, many poor and vulnerable people lack access to even these basic diagnostic tools and so do not find out what is wrong with them until it is too late to successfully treat the disease, and until long after they are likely to have transmitted the disease to others.

The problem is that the cost and sophistication of current new diagnostic tools limit their application to specialized reference laboratories, even in industrialized countries. What is required are far simpler, accurate point of care tests that can be used in remote health centres to reach the majority of tuberculosis sufferers. Affordability (to the health system) and accessibility (to individuals with suspected TB) as well as high sensitivity are central to determining the impact of any diagnostic test and its ability to accurately diagnose tuberculosis in as many patients as possible.

OBJECTIVES FOR TB TEST DEVELOPMENT

> **Simplify and improve** detection of TB cases, including smear-negative, extra-pulmonary and childhood TB, through increased sensitivity and specificity and improved accessibility.
> **Create and distribute** simple, accurate, safe and inexpensive tests that can be performed at the point-of-care level of the health care system and that produce same-day results.
> **Enable more effective** monitoring of TB treatment (latent and active).
> **Rapidly identify** drug resistance to both first- and second-line anti-TB medicines.
> **Reliably identify** latent TB infection and determine the risk of progression to active disease, enabling rational use of preventative therapy.

BARRIERS AND CHALLENGES

While there is considerable diagnostic work going on in smaller biotechnology companies and academic research groups, the private-sector diagnostic development community generally views this market as too difficult to enter to ensure a return on investment. This presents a significant barrier to innovation. The need, under the current structure, for academic and commercial partners to own intellectual property in order to leverage funding encourages secrecy and leads to duplication of work. The existence

of so many varying research agendas means the most urgent medical needs don't attract the greatest attention. The lack of validated biomarkers for active disease in adults and children has been identified as a barrier to test development, and greater efforts are needed in this area of research. Too often there is no cohesion between various actors in the development of new diagnostics, with many initiatives working in isolation.

Further, the absence of regulatory entities for diagnostics in many countries allows underperforming or inadequate tests to be marketed, which can lead to inappropriate treatment or delayed diagnosis. The current practice of some test developers to not sufficiently evaluate their own products in unbiased studies is flawed and needs to be addressed by encouraging additional independent assessment and rigorous comparison of competing or alternative technologies.

Market incentive mechanisms, which rely on high prices to fund R&D, do not result in creating advanced diagnostics in the areas of highest need; thus financing approaches should be expanded. Funding levels are still insufficient to adequately support TB diagnostics development, as funding for overall TB research and development stands at barely 20 percent of what is needed to roll back the disease to targeted levels.

So the key challenge in the fight against tuberculosis is to adapt promising new diagnostic tools for use in high-burden settings and to open a pipeline for their development, marketing, distribution and widespread use in the places where they are needed most. Overcoming this challenge, and thereby making faster and more accurate diagnosis the norm even in the most remote settings, will expand effective treatment in the developing world, vastly reducing cases worldwide and saving millions of lives.

Fortunately, a great and growing interest in the multilateral community has led to increased resources for action against TB, including development and distribution of new and more effective diagnostics. Monies are becoming available through national research institutes and donor programmes. Public-private partnerships are emerging for the purpose of improving diagnostic tools

and finding ways to speed them to market, and the international infrastructure for accomplishing this aim is beginning to take shape.

A NEW ERA OF TECHNOLOGICAL INNOVATION

Despite the barriers and challenges, recent developments in diagnostics offer hope. The convergence of multiple disciplines – physics, nanotechnology, nanobiotechnology, molecular biophysics, molecular biology, immunology, genomics and biomedical sciences – is revolutionizing scientific research and shifting diagnostic development from the biological to the molecular level. This holds great promise for the development of new TB diagnostics that can deliver faster and more accurate results, detect and help address drug-resistant strains of the disease and better meet the challenges of implementing effective tests in high-burden, low-resource areas.

Advances in biological test platforms continue to hold enormous promise for improving TB diagnostics on many levels. But the molecular techniques under development, such as the nucleic acid amplification test (NAAT), are yielding excellent results in the detection of minute amounts of genetic material – with potential application for the diagnosis and identification of active mycobacterial infection. Developments in gene mutation and antigen detection assays are also showing great promise for meeting TB diagnostic goals, as are efforts to adapt liquid chromatography methods for use in the field, and to develop successful "E-nose," urinalysis and breath analysis technologies for use in detecting tuberculosis.

Work continues to develop or improve indirect assays for TB diagnosis such as antibody detection tests (which currently do not work well), tuberculin skin tests (which lack specificity needed for highly endemic countries) and IFN-gamma release assays (which cannot now distinguish between latent and active TB). The rapidly growing biological systems research disciplines of

proteomics and metabolomics also offer promising new work.

Accordingly, the overall aim for improving TB diagnostics is to support the development of cost-effective, patient-centred applications on common technology platforms appropriate to different tiers of the health system in developing countries. This should lead to overall improved access to appropriate treatments.

KEY PHASES OF TB DIAGNOSTIC DEVELOPMENT

Needs assessment

A comprehensive needs assessment – systematically measuring the extent and nature of the problems faced by the people on whom the tests will be performed – is a critical but often overlooked aspect of the development process.

Key elements of the needs assessment phase include determining the medical requirements of the local population vis-à-vis the presence of TB, the expectations for the new diagnostic tool within a specific country or regional context, the collection of critical data related to the epidemiological context of the local area, the health structure of the area where the test is to be used, the laboratories that would be involved and the challenges that people face in accessing them. Data collection should be optimized through a collaborative approach that involves patient groups, prescribers, laboratory staff and a broad range of local stakeholders.

This phase produces a *user requirements document* that provides detailed information on the expected performance in real-life conditions, time to results (and preferably time to treatment initiation), technical requirements, users' skills, medical algorithms within which the test is to be used and a clear description of the setting where the test is to be implemented. The underlying aims of the document are to ensure high performance and ease of use amid the limited resources available in the local area. Ultimately, the needs assessment should be clearly described in the product dossier with an eye towards the stakeholders contributing to the development of a new test and to

informing donors and national programmes in their decision-making.

Following the needs assessment, a *concept definition document* will be drawn up. This provides a general vision of the product: what it would do, how it would increase access to diagnosis for the (very) poor, who would be the target users and their needs, how the users would act on the information it provides, in what tier(s) of the health care system the product will be used, what kind(s) of technologies might be employed, how the product could be manufactured and strategies for its marketing. The outcome should include an overall business plan that includes initial estimates of product manufacturing costs, selling prices, unit sales, cost to serve and support customers, profit projections and project timeline.

From here onwards the diagnostic development process goes through a number of phases, each designed to advance the prospects for successful outcome:

Feasibility: The objective of this phase is to determine whether, based on the outcomes of the needs assessment, a specific technical solution can be adapted for the specific diagnostic use required. The goal is to decrease risk before product development begins, to choose among technical solutions or options, and to arrive at a go/no-go decision on whether to proceed with development. Accordingly, definitions and targets must be clear and studies must be robust enough to adequately inform this decision. It is in this phase where specific diagnostic technologies are targeted, and where technical and logistical questions are addressed. It also includes legal assessment of intellectual property concerns and, if necessary, the development of plans to license the IP, purchase the component from a licensed vendor and/or protect any new IP that is developed. Other outcomes of the feasibility phase include a regulatory plan, a clinical trial plan, a preliminary user support plan and a marketing plan.

Development and optimization: The majority of the detailed design and development activity takes place during this phase, which includes refinement of the product through rapid prototyping and user feedback as well as establishment of

Executive summary

manufacturing and user support processes and a detailed launch plan. Estimates of product costs, selling prices, quality/reliability measures, unit sales, cost to serve and support users and project timeline are updated and more critically analyzed in this phase as part of an updated business plan. Manufacturing processes and a final product design that meets the user requirements document will be approved before moving into the next phase. Clinical studies done in this phase are field evaluations using a partially validated prototype product and are part of the optimization process, with each study requiring planning, managing, staffing and data analysis. Labelling will be developed and procedures implemented for development of expiration dating and storage conditions for raw materials, work in process and finished product. Finally, quality assurance planning takes place in this phase, as well as assuring compliance with international product development requirements or standards.

Evaluation: The major activity during a product's evaluation phase is to plan and conduct a study – one independent from the commercial goals and regulatory steps of R&D – to assess performance of a commercial product that has completed development and is design-locked and ready for scaled-up manufacture. Results are judged against target product specifications and gold-standard techniques. The high-quality data generated during this phase should be part of the performance data in a commercial package insert and may be used for in vitro diagnostic medical device (IVD) registration with regulatory agencies. Test site staff must be adequately qualified and trained, follow regulations and protocols and provide data integrity. All shipments of product must follow appropriate regulations and have appropriate labelling. Expected outcomes are data on the performance of a test in the conditions of use – which then become critical information for the users of the test within the laboratory services, as well as for providing updates of the user support plan and the marketing launch plan – and the creation of a service training programme and design history file.

Demonstration: The ability of a new diagnostic test to decrease TB morbidity, mortality and disease transmission may not be evident solely on the basis of accuracy data (i.e. sensitivity and specificity) or controlled trials of test performance. Large-scale demonstration projects are required to provide the evidence that new tests which perform well in controlled settings can have an important medical and public health impact when implemented in programmatic settings. These tests, which are typically carried out in close cooperation with national health ministries and relevant tuberculosis laboratories, are designed to assess scaled-up performance, patient outcome and public health impact, cost effectiveness and barriers to implementation. Testing is normally done against a reference standard or existing test or algorithm, and this will include the economic evaluation to identify, measure, value and compare the costs and consequences of the alternatives being considered.

Assessing the impact: Robust and informative impact assessment of a technology's performance under general conditions (as opposed to the controlled environments typical of prior phases in the blueprint) is essential to enabling national health systems make rational decisions on whether to implement a particular diagnostic test. The ultimate impact of any tuberculosis test should be measured by its capacity to generate a beneficial therapeutic outcome in as many patients as possible. Impact assessment therefore goes beyond the test's sensitivity and specificity to include factors such as its affordability for the health system and accessibility (in its fullest sense) to patients, as well as an explicit assessment of its likely societal impact. In the case of TB, the potential impact of new diagnostics on poor people's access to care is critically important.

While the standardized and regulated form of impact assessment that has been carried out for many years in industrialized countries has emphasized the functionality and safety of diagnostic tests, a shift toward health care spending by national governments means that decision-makers are

placing increasing importance on promoting the most efficient and effective use of public resources. This is likely to result in greater emphasis on health care access by different population groups, greater explicitness in the assessment process and greater transparency in decision-making.

With this in mind, the blueprint proposes a five-layer framework for collating evidence in impact assessment. The framework deals in successive order with key issues involving assessment of sensitivity and specificity, collection of quantitative and qualitative data, human resource and infrastructure requirements, economic evaluation of the tests and policy guidelines.

MAXIMIZING ACCESS TO TB DIAGNOSTICS

As current technologies struggle to meet the global target for case detection (notifying at least 70% of new smear-positive cases), the wider goal of the diagnostics community is to rapidly and correctly identify *all* tuberculosis cases. While the discovery, development and refinement of new technologies have the potential to advance towards that goal, getting there will also require an intense focus on optimizing access to the new technologies in geographical, financial and social terms. During the evaluation process, the ease with which diagnostic tests can be delivered and used in poor-country settings will be important factors as governments and health agencies assess their potential to have maximum public health impact. At the same time, in disease-endemic countries the most fundamental factor for ensuring access to a new diagnostic technology is the government's decision to introduce it.

Accordingly, this blueprint seeks to address the dimensions of access that can and should be considered in evaluating new technologies. While direct patient access and its importance to governments is undoubtedly the most important of these dimensions, there are other factors, stakeholders and methods – from local health care settings to diagnostic delivery approaches – that should be considered in

a comprehensive evaluation of how access can be maximized throughout a particular community. The test's cost and acceptability to patients are particularly important access issues as well.

One means to address access issues is by strengthening the capacity of public-sector laboratory networks, an objective in which new diagnostics play an important role as they can strengthen ties between public and private health care networks. Finally, it is vitally important that when a new diagnostic test is introduced, providers across the health care network understand it, accept it, become trained in administering it and promote its use – unless access be suppressed by under-utilization.

THE MEASURE OF SUCCESS

Better diagnostics offer great hope that the long battle against tuberculosis can be won. Success in leveraging today's promising technologies into tomorrow's fast, simple and accurate tests for the disease – and building effective pipelines for delivering these tests to the developing world – can be the linchpin in rolling back TB and reducing the misery it causes in so many poor communities. In turn, freedom from this disease can make a critical difference in generating the economic and social progress that has eluded these communities until now.

But it will take enormous resources, coordination, collaboration and innovative thinking to get the job done. By contributing to a framework that can guide diagnostics development and bring the essential elements together to create tests that will work in developing-country conditions, it is hoped that this blueprint will help advance such tests' technological and market successes.

We are close – so close – to a breakthrough that can change the world. The dream of better health is within our grasp.

Introduction

Delivering diagnostics, from concept to delivery

The difference between what we do and what we are capable of doing would suffice to solve most of the world's problems

(Mohandas Gandhi)

Despite substantial success in implementing standardized care and improving rates of cure in recent years, the global burden of tuberculosis (TB) remains enormous. Nearly 9 million people develop TB disease each year, and an estimated 1.6 million die from the disease. Despite this immense global burden, case detection rates are low, posing serious hurdles for TB control. Lack of rapid and accurate diagnostics to facilitate case detection is widely considered to be a major obstacle to TB control. Even in 2009, TB diagnosis continues to rely heavily on tools such as direct-smear microscopy, solid culture and chest radiography. These tools often perform suboptimally, and require infrastructure that is often unavailable in the periphery of the health system where patients first seek care.

The limitations of the existing portfolio of TB diagnostics were exposed by the HIV epidemic and by the growing threat of multidrug-resistant (MDR) and extensively drug-resistant (XDR) tuberculosis. In addition to inherent limitations of the tools we have, there are health systems failures, diagnostic delays and access to quality care issues that need to be factored in. These factors, combined with inadequate tools, often result in missed or late diagnoses with serious consequences for TB patients and communities.

Thankfully, the situation has dramatically changed. Over the past decade, unprecedented interest and activity have been focused on developing new TB diagnostic tools, largely because of the engagement of agencies such as the Stop TB Partnership's New Diagnostics Working Group (NDWG), the Foundation for Innovative New Diagnostics (FIND), the Global Laboratory Initiative (GLI), the World Health Organization (WHO) and the Special Programme for Research and Training in Tropical Diseases (TDR), several industry partners, non-governmental agencies and national TB programmes. Funding agencies such as the Bill & Melinda Gates Foundation, the Global Fund to Fight AIDS, TB and Malaria (GFATM) and UNITAID have provided much-needed resources and impetus to push the new tools

agenda, in keeping with the Global Plan to Stop TB. As a consequence, we now have a strong pipeline of improved or new tools for TB diagnosis. The technologies range from simple microscopic and growth detection systems to sophisticated molecular and immune-based systems.

The Stop TB Partnership's NDWG was established in 2001 as one of the core working groups within the Partnership. This working group coordinates and facilitates the development, evaluation and implementation of new and improved TB diagnostics in a scientifically acceptable, evidence-based and timely manner by linking all stakeholders involved in the diagnostics development and evaluation pathway. The group's mission is to advocate and implement research and/or operational activities supporting development and implementation of TB diagnostic tools, and to collaborate with other elements of the Partnership so as to create synergy and add value to actions taken.

In 2008, the Stop TB Partnership's Retooling Task Force (RTF) and NDWG produced a detailed brochure, *New laboratory diagnostic tools for tuberculosis control,* about the diagnostics pipeline. This is mainly to provide guidance to national TB programmes (NTPs) and for funding and technical agencies that may wish to support the development, evaluation or implementation of new tools. Figure 1 shows the pipeline, where the tools are stratified as WHO-endorsed or as "tools in late-stage development/evaluation" and "tools in early phase development." The pipeline not only describes the various tests but also provides some information on the commercial kits available, training requirements and estimated costs.

Introduction

Figure 1. Summary of new TB diagnostic technologies

SUMMARY OF TECHNOLOGIES			ESTIMATED COSTS			
Technology	Description	Product	Training[1]	Infra-structure[2]	Equip.[3]	Consumables
WHO-ENDORSED TOOLS (2006-2008)						
Liquid culture	Commercial broth-based culture systems detect TB bacteria (manual and automated systems are available), can be configured for DST	BacT/ALERT 3D; MGIT	Extensive (3 weeks)	▪▪▪	High	High
Molecular line probe assay	Strip test simultaneously detects TB bacteria and genetic mutations that indicate isoniazid and/or rifampicin resistance	GeoType® MTBDR and MTBDRplus; INNO-LiPA Rif.TB	Moderate (3 days)	▪▪ to ▪▪▪	High	High
Strip speciation	Strip speciation test detects a TB-specific antigen from positive liquid or solid cultures to confirm the presence of TB bacteria in culture samples	Capilia TB Rapid Diagnostic Test	Minimal (1 day)	▪▪▪	Low	Medium
TOOLS IN LATE-STAGE DEVELOPMENT/EVALUATION						
Automated detection and MDR screening	Device allows automated sample processing; DNA amplification and detection of M. tuberculosis and screening for rifampicin resistance	Cepheid GeneXpert device and Xpert MTB cartridge	Minimal	▪	High	High
Colorimetric redox indicators	Technique detects isoniazid and rifampicin resistance in culture samples after incubation with redox dyes	Non-commercial method (Resazurin)	Extensive	▪▪▪	Low	Medium
Front-loaded smear microscopy	Based on 2 or 3 specimens but aims to examine specimens on the day that patient presents to the health service (thus identifying 95% of TB cases)	n/a	Minimal	▪	Low	Low
Interferon gamma release assay	Blood test detects specific cellular immune responses indicating TB infection	QuantiFERON®-TB Gold In Tube; T-SPOT.TB®	Moderate	▪	Low	High
LED fluorescence microscopy	Robust fluorescence microscopy (FM) systems based on light-emitting diodes (LEDs) that could allow the advantages of FM at levels of the health system where conventional FM would be impractical	Fraen; LW Scientific; Zeiss	Moderate Moderate Moderate	▪ ▪ ▪	Medium Medium Medium	Low Low Low
Microscopic Observation Drug Susceptibility (MODS)	Manual liquid culture technique uses basic laboratory equipment (incl. an inverted light microscope) and microscopy skills to detect TB bacteria	Non-commercial method	Extensive	▪▪ to ▪▪▪	Medium	Medium
New solid culture methods	Solid culture technique measures nitrate reduction to indicate isoniazid and rifampicin resistance.	Non-commercial method (Nitrate reductase assay)	Moderate	▪▪ to ▪▪▪	Low	Medium
	Solid culture technique simultaneously detects TB bacteria and indicate isoniazid and rifampicin resistance	Non-commercial method (Thin layer agar culture)	Extensive	▪▪ to ▪▪▪	Low	Medium

TOOLS IN EARLY PHASE OF DEVELOPMENT			
Tool	Level of health system	Tool	Level of health system
Breathalyser screening test	Community or point-care	Sodium hypochlorite (bleach) microscopy	Peripheral laboratory
First-generation loop-mediated isothermal amplification technology platform (LAMP)	Peripheral laboratory	Sputum filtration	Peripheral laboratory
Lipoarabinomannan (LAM) detection in urine	Peripheral laboratory	TB Patch Test	Health post
Phage-based tests	Reference laboratory	Vital fluorescent staining of sputum smears	Peripheral laboratory

[2]Key	Description
▪	Basic Laboratory [4] ; no specialized biosafety equipment
▪▪	Biosafety level 2. Specialized biosafety equipment required; such as biosafety cabinet
▪▪▪	Biosafety level 3. Biosafety cabinet and other primary safety equipment required. Controlled ventilation that maintains a directional airflow into the laboratory required

[1] Estimate assumes that technicians are already trained in existing TB diagnostic techniques (such as smear microscopy and culturing) and the necessary laboratory safety precautions.
[3] Product prices may vary depending on geographical location and terms of supply. Ranges are indicative only: Low (minimum-2000US$); Medium (2001-7000US$); High (7001 + US$).
[4] Detailed information available on the WHO laboratory Biosafety Manual: http://www.who.int/csr/delibepidemics/WHO_CDS_CSR_LYO_2004_11/en/

From: World Health Organization & Stop TB Partnership. *New laboratory diagnostic tools for tuberculosis control.* World Health Organization, Geneva, 2008

As a next step, the NDWG is pleased to present this document to provide a comprehensive, clear, well-referenced blueprint to guide researchers, clinicians, industry partners, academics and TB controllers in all sectors in aspects of TB diagnostics development, from concept to implementation and delivery. This blueprint will allow an objective, staged and inclusive pipeline to be developed and will help outline the appropriate value chain for the development of new diagnostics to be submitted for scientific review and endorsement to WHO, and to inform global policy development.

As shown in Figure 2, the pathway to TB diagnostics development begins with needs assessment and the concept phase. This is followed by feasibility, development and optimization. Once a prototype product is ready, it is subjected to rigorous evaluation studies to determine test accuracy and performance. If the evaluation study results are promising, the product is subject to demonstration and impact studies. These measure the impact and effectiveness of introducing a new test in routine health care and programmatic settings. Taken together, evidence from various evaluation, demonstration and impact studies should inform the development and/or revision of policies and guidelines on TB diagnostics. In real life, the pathway is not likely to be strictly linear, but a cyclic and repetitive process.

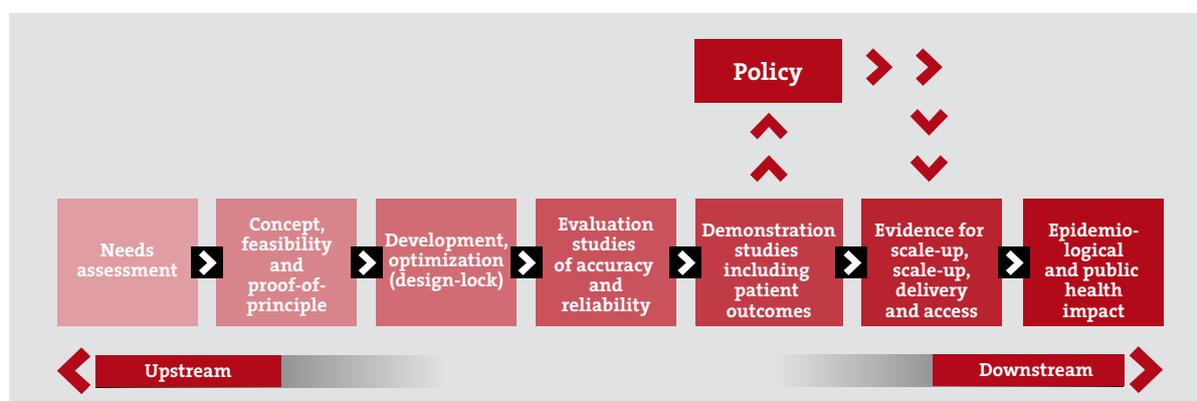## Figure 2. TB diagnostics pathway, from need assessment to delivery

Table 1. Objectives for diagnostics development phases and types of questions that can be addressed at each stage

Description/stage	Objectives	Questions that can be addressed
Pre-clinical: concept and feasibility	> To develop the concept and a certain feasibility	> *What is the vision for the technology? Will it be feasible?*
Pre-clinical: proof-of-principle	> To develop and optimize the prototype	> *Does the test have any discriminatory ability in proof-of-principle studies?*
Laboratory and clinical settings: early evaluation	> To evaluate the test accuracy in a case-control design or using well-characterized specimens. > To evaluate test reliability (reproducibility)	> *Does the test have accuracy in a case-control setting?* > *Is the test reliable (reproducible) in various settings?*
Clinical setting: late evaluation	> To evaluate the test accuracy in a population where the test is clinically indicated (e.g. patients with suspected TB)	> *Does the test have accuracy in a population where the test is clinically indicated?*
Clinical and field settings: demonstration, delivery, impact and access	> To look at the impact on patient important outcomes and globally recognized TB indicators and effectiveness of introducing a new test in routine health care and programmatic settings (generating evidence for scale-up and evidence for post-scale-up impact)	**1.** *How easy and feasible is it to implement the test in programmatic, real-world settings?* **2.** *How does the new test impact the laboratory or health-care delivery system? How does the new test influence the physician's thinking or clinical decision-making?* **3.** *Does the new test have any added (incremental) value over and beyond tests and clinical data or algorithms that are already available and in use?* **4.** *What are the clinical consequences of introducing a diagnostic test? Do patients who get the new test fare better in terms of outcomes than those who do not get the test or get another test?* **5.** *What is the cost-effectiveness of the test in programmatic settings?* **6.** *What are the societal consequences of scaling-up and rolling out a new test in a country or province? What is the post-marketing or post-scale-up impact of a new test?* **7.** *Does the government policy enable access to the new technology? Are appropriate infrastructure and delivery systems essential for implementation available in the country? Are health care providers willing and able to utilize the new test? Does the new test facilitate equitable access to all patients?*

15

Introduction

This blueprint document is structured along the lines of the process shown in Figure 2 and Table 1. It presents the following information in detail:

> **the history** of TB diagnostics;
> **the current status** of TB diagnosis;
> **the rationale** for a diagnostics blueprint;
> **needs** assessment;
> **basic** research and concept development;
> **feasibility,** development and optimization;
> **evaluation** studies;
> **demonstration** and impact studies;
> **maximization** of access and measurement of impact; and
> **barriers** and challenges to test development and implementation.

In addition, principles of various TB diagnostics and their technical descriptions are also included as a technical annex on the CD in the back cover of this publication. Additional resources, including evidence synthesis, policies, guidelines, research agendas and training materials, have been compiled by the NDWG in collaboration with several partners at a new website, *Evidence-based TB diagnosis,* available at *www.tbEvidence.org*

The TB diagnostics pipeline can be seen as a development and evaluation pathway that leads to regulatory approval of a diagnostic product and/or WHO endorsement for use in TB control programmes. This can be a long and expensive process and requires careful planning and costing. Diagnostics may fail at any stage of the process and either be dropped from the pipeline or require additional development work to overcome shortcomings of the original format. However, diagnostics frequently get stuck in the pipeline simply because of inadequate planning or financing. In these cases, the diagnostic often benefits from the involvement of a "champion": a financial donor, industrial partner, technical agency or non-governmental organization that can provide necessary planning and/or financial

support. Potential champions require a clear understanding of the diagnostic pipeline, the stages involved, the criteria for positioning a diagnostic in a specific stage, the reasons why a diagnostic's development may have stalled and the likely costs involved in taking the diagnostic to the end of the pipeline.

It has become clear over the past few years that there was no consensus on what constituted this development and evaluation pathway, nor consensus on what the criteria were for positioning TB diagnostics along this pathway. As this seriously compromised the ability of such champions to engage, it is hoped that this scientific blueprint for TB diagnostics, developed through wide consultation with the many stakeholders involved, will provide the necessary consensus. The TB diagnostics pipeline described in Figure 1 provisionally assigned diagnostics to various stages. The availability of this new blueprint should permit the more accurate staging of products under development, and allow other more accurately staged products into the pipeline.

We are currently witnessing tremendous interest in and support for developing new tools for TB control. After decades of little or no growth the diagnostics pipeline has rapidly grown, with several technologies showing great promise. Indeed, some are already making their way into the field, but there is much work to do. Unless new tools are adequately validated and put through all the phases described in this blueprint, their accuracy, impact and effectiveness cannot be determined and new policies cannot be formulated and implemented at the ground level.

Improvement and dissemination of new tools should have a profound effect on TB control throughout the world. We hope this blueprint will succeed in bringing together people and groups to make a difference, including diagnostics research scientists, professionals in industry, academics,

national and public health officials and business leaders. Ultimately, we hope it will help create or identify the most promising TB tests, rapidly push them through the pathway outlined in this blueprint and generate evidence for policies to make the best tests available and easily accessible to those who need them most.

On behalf of the NDWG, we thank all the contributors who played a key role in putting together this blueprint, and would like to specifically acknowledge the efforts of the editorial and production teams. We encourage all members of the NDWG and the wider TB community to send feedback, and to contribute to and support future endeavours and activities of our working group and the Stop TB Partnership.

Chapter 1

The search for *tubercle bacilli*
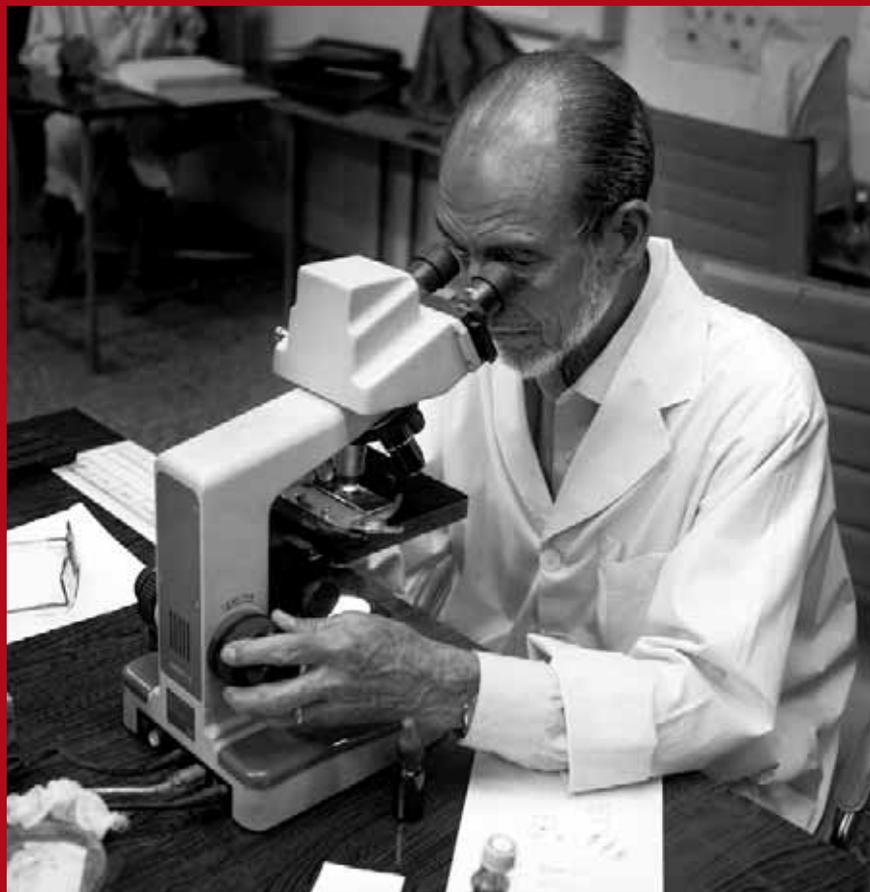
EARLY DETECTIVE WORK

Tuberculosis is nearly as old as human history. Traces of it have been found from Neolithic burial sites some 7000 years old and in ancient Egyptian mummies. Hippocrates identified the disease, calling it *phthisis* or "consumption," and noted that it was fatal. However it was not until the 17th century that more exact pathological and anatomical descriptions of the disease began to appear.

In 1720 an English physician, Benjamin Marten, suggested that consumption – so called because it seemed to consume people from within – could be caused by "wonderfully minute living creatures." It would finally be left to German physician Robert Koch to actually "see" them by use of a staining technique. The early bacteriologist's work on isolating the *tubercle bacilli* was first presented in Berlin on 24 March 1882. At the time of Koch's work, solid agar medium was not widely in use. In fact, Walther and Angelina Hesse helped contribute to the use of agar medium in his laboratory [1] and, using very laborious methods, Koch was able to isolate *Mycobacterium tuberculosis* by using Tindall's blood serum medium [2; 3].

GUINEA PIGS AND X-RAYS

Koch and others infected guinea pigs with samples containing TB (such as sputum, for instance) and seeded their media with tissues obtained during autopsy of the guinea pigs in order to obtain pure *tubercle bacilli* culture, quite a tedious procedure[4]. Following Koch's discovery, the guinea pig was used as the primary culture means for TB. As reported by Baldwin et al. at the Trudeau Sanatorium Research and Clinical Laboratory, for routine cultivation and preserving stock cultures for study or experimental inoculation, the various forms of egg media became universally used as well as coagulated beef or sheep serum, glycerol agar, Dorset's egg medium and Petroff's potato broth [5].

However, for the first half of the 20th century X-ray remained the number one tool worldwide for the diagnosis of TB,

particularly pulmonary tuberculosis. More than 50 years ago, Middlebrook and Cohn stressed the need for bacteriologic analyses: "Today, more than ever, there is an increased need for bacteriologic diagnosis and monitoring of treatment in tuberculosis" [6]. In 1948, when the first report of a WHO Expert Committee on Tuberculosis was published, laboratory diagnosis of tubercle bacilli was one of six problems that needed action to establish uniform procedures, after tuberculin production and testing, BCG preparation, classification of tuberculosis, X-ray interpretation and testing (including mass radiography) and evaluation of new chemotherapeutic agents such as streptomycin [7].

Chadwick and Pope's *The Modern Attack on Tuberculosis* [8] mentions that the expense of X-rays limits their general use. Sputum that is negative on examination of smears should be further studied if TB is suspected. Tubercle bacilli that are missed in smears may be found in concentration of sputum or in culture; sputum or gastric contents injected into a guinea pig may furnish positive evidence of TB that cannot be obtained in any other way.

FROM GUINEA PIGS TO EGG YOLKS

In 1902 Dorset suggested the first solid egg medium (just coagulated egg yolk and white). In 1915 this was modified by Petroff, who added gentian-violet to the medium and introduced the decontamination-homogenization of specimens by sodium hydroxide. The latter was quite a revolution since it replaced the need to use guinea pigs as the primary means to isolate TB.

In 1941, the American Public Health Association published its first edition of *Diagnostic procedures and reagents: Technics for the laboratory diagnosis and control of the communicable diseases* [9]. The egg yolk medium is described as easy to prepare but expensive, requiring a large number of eggs. The egg yolk alone appears to yield a more prolific growth than do media made with whole eggs, and for the isolation of *Mycobacterium bovis*, egg yolk

media are better if made without glycerin. Egg-based Lowenstein-Jensen with glycerin is described as good for the isolation of *Mycobacterium tuberculosis* [9]. John Abbott of the New York State Department of Health demonstrated that factors such as prompt collection, care in minimizing contamination, suitable storage and prompt delivery weigh greatly on the successful bacteriologic examination [10].

A PROLIFERATION OF CULTURE MEDIA

In the late 1950s, Dufault [11] reported that routine chest X-ray of all hospital admissions had already proved to be a productive and comparatively inexpensive method of case-finding which, once generally adopted, would uncover thousands of unsuspected early-stage cases. During this time, numerous different culture media were in use. In 1947, Dubos and Middlebrook created a synthetic medium of an oleic acid-albumin mixture that was not only as effective as any of the more complex organic media, but produced faster growth and allowed earlier detection. Whereas the Dubos-Middlebrook medium offered advantages in research laboratories, the modified Lowenstein medium was generally used in clinical work. A still more sensitive method was reported to be that of guinea pig inoculation.

In 1960, the US Communicable Disease Center reported that drug susceptibility testing could be performed either by the direct or indirect methods [12]. The direct method had the advantage of saving time and of inoculating the test medium with a bacterial population in which the proportion of resistant and susceptible organisms is that of those existing in the exudates sampled. Drugs tested included streptomycin, para-aminosalicyclic acid (PAS), isoniazid, viomycin, pyrazinamide and cycloserine. To improve accuracy, modified Herrolds egg yolk-agar medium in tubes was used since it did not require inspissation. Culture media (at least two tubes of two different media were recommended for each specimen) included modified Lowenstein's medium, Petragnani medium, Trudeau Committee medium, Peizer's medium, Dubos-Middlebrook media, Tarshis blood agar, Middlebrook's 7-H3 medium and Middlebrook's 7-H9 and 7H-10 media.

A SYNTHETIC DEVELOPMENT: THE LIQUID MEDIUM

In 1965, Middlebrook [13] described synthetic liquid medium for research use. In the first Manual of *Clinical Microbiology*, published in 1970 by the American Society for Microbiology, Runyoun et al. specified the use of liquid medium (Dubos, Proskauer-Beck, or 7H9) for sterile sites without the need for prior treatment [14]. Tubed or bottled Lowenstein-Jensen and a clear agar plate such as 7H10 agar were recommended for sputum to allow for microscopic examination and to permit earlier detection of colonies. Five to ten percent CO_2 was considered essential when using 7H10 agar and was markedly stimulatory to growth of *Mycobacterium tuberculosis* on egg medium. Guinea pig inoculation was recommended for contaminated specimens.

TOWARDS FASTER DIAGNOSIS

Mitchison et al. in 1971 described the use of liquid culture medium containing antimicrobials to greatly improve results [15]. In 1977, a semi-automated method incorporating a liquid culture medium containing radioactive palmitic acid and antimicrobials to kill non-acid fast microorganisms was described by Middlebrook et al. [16]. The widespread adoption of this 7H12 medium was due to the US Centers for Disease Control (CDC) recommendation in 1993 prompted by the resurgence of tuberculosis [17]. The radiometric 7H12 provided higher yields and shorter turnaround times.

Today, a number of non-radiometric growth detection systems utilizing liquid culture media and continuous monitoring are already in use. However, all currently known culture systems need at least eight days to detect growth in smear-positive specimens. Together with line probe assays (which combine mycobacterial species identification and drug susceptibility tests to the two most important drugs, isoniazid and rifampin) these technologies considerably decreased times necessary for diagnosis of pulmonary tuberculosis and for detection of multidrug-resistant (MDR) or extensively drug-resistant (XDR) strains of the *M. tuberculosis* complex.

In a recent *Morbidity and Mortality Weekly Report* [18], *Updated guidelines for the nucleic acid amplification (NAA) tests in the diagnosis of tuberculosis,* the US CDC recommends that NAA testing be performed on at least one respiratory specimen from each patient with signs and symptoms of pulmonary TB for whom a diagnosis of TB is being considered but has not yet been established, and for whom the test result would alter case management or TB control activities. However, until now the positive result of most of the commercially available NAA tests offers a very high predictive value (100%), whereas the negative predictive value is significantly lower (85%) [19]. As a consequence, it is now not possible to completely replace the use of culture media, and even less so in low-income, high-TB-burden countries which harbour ~85 % of TB cases [20], even though these countries often have the most difficulties in establishing reliable culture facilities.

The culture method founded in the early days after Koch's discovery of *tubercle bacilli* soon established itself and was widely accepted. The introduction of nucleic acid amplification, including point of care real-time PCR assays and other non-molecular POC platforms, is expected to result in more timely detection of *Mycobacterium tuberculosis* in TB suspects, thus hastening the elimination of tuberculosis.

Key points

✱ In close to 130 years since Koch's discovery of *tubercle bacilli*, the search for effective diagnostic tools has moved through a succession of varied and ultimately more effective methods.

✱ However, in the majority of the world smear microscopy, a technique developed over 100 years ago, is still the mainstay of tuberculosis diagnosis.

✱ Globally, culture is still the gold standard for tuberculosis diagnosis, but it has been improved by the use of automated liquid culture technology.

Chapter **2**

TB diagnosis today: *the search for improved diagnostics continues*

THE TELLTALE SIGNS

Good patient care, with prompt and accurate diagnosis and effective treatment, is the basis of the public health approach to tuberculosis control [21]. Accordingly, all health care providers who diagnose and treat patients with TB not only provide individual care but also contribute to controlling the disease [22]. The clinician's first step in diagnosing TB is to recognize the disease's clinical features, prompted by an index of suspicion ("Could TB be the cause of this patient's illness?") while mindful of the main groups at risk [23]. Patients with TB usually present with clinical features related to systemic disturbance (e.g. fever, night sweats, weight loss and tachycardia) and to the local site or sites of disease (e.g. cough and haemoptysis in pulmonary TB, enlarged lymph nodes in TB lymphadenopathy, headache and neck stiffness in meningeal TB [24]. Because of the barriers they face in access to care, patients in developing countries tend to present at a more advanced stage of illness and therefore with more pronounced clinical features.

THE DIFFICULTY OF GETTING A STRAIGHT ANSWER

Since the clinical features of TB are non-specific, answering the question "Could TB be the cause of this patient's illness?" involves diagnostic tests of various degrees of sensitivity and specificity. Diagnostic evaluation in all patients who may have TB involves seeking microbiological, or sometimes histopathological, confirmation of TB [23]. In some cases, such as sputum smear-negative pulmonary TB, the disease is presumed as the most likely diagnosis without "gold standard" confirmation [25]. The usefulness of a test in confirming or excluding the possibility of TB depends on the predictive value of a positive or negative test result, determined by the prevalence of TB in the population to which the patient belongs. *Toman's tuberculosis: case detection, treatment and monitoring* provides a comprehensive account of the evidence base underpinning the performance of diagnostic tests commonly used for TB case detection, i.e. sputum microscopy and chest X-ray [26].

Inadequate evaluation of patients who may have TB results in missed opportunities for earlier detection of the disease and leads to increased disease severity for patients and, in the case of pulmonary TB, a greater likelihood of transmission of *M. tuberculosis*. Failure to perform a proper diagnostic evaluation before initiating anti-TB treatment potentially exposes the patient to the risks of unnecessary or erroneous treatment with no benefit, and sometimes with worsening consequences.

THE COMPLICATING HIV FACTOR

The impact of HIV on TB is twofold – TB is both more common and more difficult to diagnose among people with HIV infection [24]. The diagnosis of TB in HIV-infected patients may be difficult not only because of the increased frequency of sputum-smear negativity (up to 40% in culture-proven pulmonary cases) but also because of atypical radiographic findings, a lack of classic granuloma formation in the late stages and a negative tuberculin skin test. Delays in treatment may prove fatal.

NO TEST IS PERFECT, THE SEARCH MUST GO ON

The reliability of the use of diagnostic tests depends on quality control. Specifically, the reliability of sputum-smear microscopy and culture for *M. tuberculosis* depends on an effective national system of laboratory quality assurance. Investment in the development of new tests more suited to the epidemiological, societal and economic conditions in developing countries are a priority. Yet efforts to make the most of the currently available tests should continue and even be expanded [29]. Each step towards the ideal of a low-cost, quick, user-friendly, highly specific and highly sensitive diagnostic test is likely to contribute to more effective and efficient diagnosis and case finding, with better health outcomes for individuals with TB and better TB control.

Five technologies are commonly used in TB diagnosis in disease-endemic countries – radiography, microscopy, culture, histopathology and tuberculin skin testing.

Radiography

Nearly all patients with chronic cough have a chest radiograph during their diagnostic evaluation. Chest radiography is sensitive in the diagnosis of pulmonary TB but of low specificity, especially in people with HIV infection. Radiography is important in the diagnosis of several forms of extrapulmonary TB, e.g. pleural, vertebral and joint.

Bacteriology

Bacteriological confirmation of TB diagnosis is the gold standard and should be sought by obtaining specimens from the suspected sites of involvement, at least for microscopy and preferably also for culture. Pulmonary TB is the commonest form of TB, and diagnostic evaluation should include sputum microscopy. Extra-pulmonary TB (without associated lung involvement) accounts for 15–20% of TB in populations with low HIV prevalence and an even higher proportion in populations with a high HIV prevalence. Since there are fewer *M. tuberculosis* organisms present in extra-pulmonary sites, the yield of microscopy is lower than in pulmonary TB, and culture and histopathological examination of tissue specimens is important. For example, microscopic examination of pleural fluid in pleural TB detects acid-fast bacilli in only about 5-10% of cases, and the diagnostic yield is similarly low in meningeal TB. Because specimens may be difficult to obtain from some extra-pulmonary sites, bacteriological confirmation of extra-pulmonary TB is often more difficult than for pulmonary TB. Compared to adults, children more often have extra-pulmonary TB, and when they do have pulmonary TB sputum positivity is less common. Bacteriological confirmation is therefore more difficult in children than adults [27].

Microscopy

In populations with high TB incidence, identification of acid-fast bacilli (AFB) in samples from a suspected site of disease provides immediate diagnosis with high specificity and is the practical gold standard. In addition to being highly specific for *M. tuberculosis*, AFB identification by sputum microscopy is particularly important among patients with pulmonary TB for four reasons: a) it is the most rapid method for determining if a person has TB; b) it identifies the most likely transmitters of infection; c) among HIV-negative persons it identifies those who are at greatest risk of dying from the disease (although in persons with HIV infection, mortality rates are greater in patients with clinically diagnosed TB who have negative sputum smears than among HIV-positive patients who have positive sputum smears) [28]; and d) it is the most expeditious method of monitoring response to treatment. Ideally, where quality-assured laboratories capable of TB culture are available, sputum cultures as well as smears should be performed for monitoring. The main drawbacks of traditional light microscopy are that it is, laborious and of low sensitivity. Improvements on basic microscopy for TB diagnosis include better techniques for microscopy (e.g. using a fluorescence microscope to detect fluorochrome-stained TB bacilli) and for sample preparation, e.g. sputum concentration.

Culture

Bacteriological confirmation of TB is possible from a wide range of samples (depending on the suspected site of disease) including sputum, lymph node aspirate, cerebrospinal fluid, bone marrow and blood. *M. tuberculosis* complex grows on LJ in around 3-8 weeks, and so often does not provide useful diagnostic information at the time clinicians need to make the diagnosis. Proving a diagnosis of TB by culture is useful in cases of diagnostic doubt, especially in populations with low TB incidence, and in the management of suspected drug-resistant TB. Automated culture systems enable a quicker result, although a culture result provided even within a couple of weeks often doesn't help the clinician at the time of pressing diagnostic need.

Histopathology

Histopathological examination of tissue specimens (e.g. needle biopsy of lymph nodes) is important in confirming the diagnosis of many forms of extra-pulmonary TB. The sensitivity mainly depends on ease of sampling, e.g. diagnostic yield is higher in lymph nodes than pleural TB. Specificity is high, although reduced in people with HIV infection, which is associated with decreased granuloma formation. Facilities and resources for histopathology are often unavailable in many developing-country settings.

Tuberculin skin testing

Reactivity to tuberculin indicates *M. tuberculosis* infection and can help diagnose TB disease among individuals belonging to a group with low prevalence of infection, e.g. people in developed countries and young children in developing countries. In groups with high prevalence of infection, the predictive value of a positive test is too low to be of much value in confirming diagnoses of TB. The sensitivity of the tuberculin skin test is high, although particular conditions (e.g. HIV infection and malnutrition) can suppress tuberculin reactivity, resulting in decreased sensitivity of the test among the very people who are at increased risk of TB. The value of a positive tuberculin skin test in the diagnosis of TB is often decreased in developing countries, where BCG immunization and exposure to environmental mycobacteria result in decreased test specificity.

TB LABORATORY DIAGNOSTIC TOOL DEVELOPMENT: THE OBJECTIVES

The overall aim for improving TB diagnostics is to support the development of cost-effective, patient-centred applications on common technology platforms appropriate to different tiers of developing country health systems. This should lead to overall improved access and to appropriate treatments.

Objectives for TB test development

> **To simplify and improve** detection of TB cases, including smear-negative and extra-pulmonary and childhood TB, through increased sensitivity and specificity and improved accessibility.
> **To develop** simple, accurate, safe, rapid and inexpensive tests that can be performed at the point-of-care level of the health care system and that produce quick results on the same day.
> **To enable more** effective monitoring of TB treatment.
> **To rapidly identify** drug resistance to both first- and second-line anti-TB medicines.
> **To reliably identify** latent TB infection and determine the risk of progression to active disease, enabling the rational use of preventative therapy.

Key points

✱ A complex variety of symptoms can indicate the presence of TB. Depending on the stage of infection and the target population, the diagnostic approach will vary.

✱ HIV and TB have together created an additional complication in which the capacity of health care infrastructures is a deciding element.

✱ The search for a low-cost, quick to result, user-friendly, specific and highly sensitive test remains a priority.

Chapter 3

The current
TB epidemic

In 2006, WHO updated the Stop TB strategy to be in line with new global targets for 2015. The most recent complete WHO report on TB epidemiology relates to the global picture from 2007 and clearly links to these global targets. The major indicators from this report are highlighted below. "Global tuberculosis control – epidemiology, strategy, financing" 2009 WHO/HTM/TB/2009.411

INCIDENCE

In 2007 there were an estimated 9.27 million incident cases of TB, with the majority of these occurring in Asia and Africa. This is an increase from the previous year, when 9.24 million cases were recorded. Though the absolute number of TB cases is rising, the number of TB cases per capita has been falling slowly (less than 1% per year) since a peak of 142 cases per 100 000 population in 2004. This fall has been reflected in all WHO regions except Europe, where numbers have remained stable.

PREVALENCE

The prevalence of TB cases dropped from 13.9 million in 2006 (210 per 100 000 population) to 13.7 million in 2007 (206 per 100 000 population).

MORTALITY RATES

Mortality rates decreased globally in 2007, with 1.3 million HIV-negative TB patients dying in 2007 and 456 000 deaths occurring amongst individuals who were infected with both TB and HIV. The reported number of TB deaths in HIV-positive individuals has doubled in 2007 compared to previous years. However, this is related to the availability of new data from the regions rather than an actual dramatic rise in deaths.

MDR- AND XDR-TB

A half-million cases of MDR-TB were estimated in 2007, similar to estimates made in 2006. In addition, extensively drug-resistant cases were reported in 55 countries. However, reliable data in this area is scarce, particularly in Africa.

HIV-TB CO-INFECTION

It is estimated that 15% (1.37 million) of the total number of TB cases are HIV-positive, with the vast majority of these cases (79%) occurring in Africa. Data collected from 64 countries shows that the risk of developing TB among HIV-positive patients is 20 times higher than for HIV-negative people in settings where HIV prevalence is high.

In the vast majority of countries, diagnostic and treatment services are considered to be integrated at the primary health care level. Nevertheless, integration of HIV-TB activities still needs to be scaled up so that all TB patients are tested for HIV and HIV-positive patients are given information and, where possible, tested for TB.

Slow progress in improving health indicators is occurring globally, although Africa still falls behind the rest of the world due to weaker health systems and the impact of the HIV epidemic. Three major indicators (incidence, prevalence and death rate per 100 000 population) are all falling since a peak in 2004. However, if the

present trends continue, the target to halve the prevalence and death rates compared with 1990 will not be met globally, mainly due to the situation in Africa. Improvements in TB diagnosis are at the forefront of what is needed to meet the global targets for 2015.

MEETING THE TARGETS

During 2006-7, 59 countries met the target of successfully treating 85% of smear-positive cases; the majority of these countries were in the eastern Mediterranean, Western Pacific or South-East Asia. Globally the figure for successful treatment is also 85%, an increase of less than 1% compared with 2005, but Africa and Europe are still below this target. The target of detecting 70% or greater of smear-positive TB cases was met by 74 countries, the majority of them in the Americas and South-East Asia. Worldwide, the overall figure for smear-positive case detection is 63%, an increase of 1% from the previous year. Thirty-six countries met both of these targets, with the majority of these in South-East Asia and the Western Pacific.

Key points

* Whether the targets set by the Stop TB Partnership (in the context of the Millennium Development Goals) for the year 2015 remain achievable is uncertain.

* The absolute number of TB cases is rising, but there is a small reduction in the number of cases per capita since 2004. However, prevalence is dropping marginally.

* Mortality rates for those infected with TB have grown, chiefly as a consequence of co-infection with HIV.

* The overall worldwide reduction in the TB indicators of incidence, prevalence and death rate per 100 000 population is slowed by the poor health infrastructure and by the low socio-economic status of much of the population of sub-Saharan Africa.

Chapter 4

The rationale for the diagnostics pipeline

Often thought of as a linear process from idea inception to diagnostic test, "the diagnostic pipeline" of research and development is actually a cyclic, repetitive process that begins with the recognition of a problem and continues through an expansive thinking phase to experimentation, assessment and adoption. This process may be repeated as the technology is improved or modified for new uses. The process also cycles and moves from laboratory to applied research and, ultimately, to (clinical) application. It sometimes falls back to address unanticipated problems and then advances again as those problems are resolved.

A cornerstone of therapeutic drug development is the phasing of research steps. This phasing is thought to provide an efficient and rigorous approach to drug development. A phased approach to diagnostics development might similarly lead to a more efficient and rigorous research process. This blueprint for TB diagnostics development offers a structure which guides those involved in development through the different phases from discovery to delivery. It will allow analysis of promising tests (from concept) in terms of needs as well as where they are in the development process and why some are stuck. The blueprint will help to coordinate diagnostics R&D and ultimately to stimulate diagnostics development by clarifying the landscape and the development steps. Considering the lack and/or inconsistency of national regulations on diagnostics quality in most countries, it is hoped that this blueprint adds needed rigour in diagnostics evaluation. An independent evaluation according to scientifically sound methodology accompanied by assessment of likely impacts and factors required for national-level implementation should feed the evidence base on which health policy decisions are made.

This blueprint for the development of diagnostic tools for the detection of TB is structured into the following phases:

> **Needs assessment** – measuring the extent and nature of the problems faced by the people on whom the tests will be performed. The *user requirements document* is a key output of this phase.
> **Feasibility** – determining whether, based on the outcomes of the needs assessment and using a concept definition document, a specific technical solution can be adapted for the specific diagnostic use required; this decreases risk before product development begins.
> **Development and optimization** – implementing all technical design and development activities in accordance with the *user requirements document*, including regulatory/registration, launch plans, clinical trials and quality control.
> **Evaluation** – assessing the performance of a commercial product that has completed development and is design-locked and ready for scaled-up manufacture.
> **Demonstration** – carrying out large-scale demonstration projects to provide evidence that new tests which perform well in controlled settings can have important medical and public health impacts when implemented in programmatic settings.
> **Access** – viewing the diagnostic from a patient point of view as defined in geographical, financial and social terms.
> **Impact** – measuring the ability of a diagnostic test to accurately diagnose TB in as many patients as possible, to improve the care and outcome for patients and generally improve public health.

Although this involves multiple steps, the process is fairly rational in selecting promising markers and technology early, while recognizing that small studies do not answer the ultimate questions that need to be addressed. One advantage of the blueprint is that by addressing focused questions at each phase it yields well-defined standards of practice for study design.

Chapter 5

Assessing the needs

THE SEARCH FOR ANSWERS BEGINS HERE

Improved diagnostic tools and approaches are central to reversing TB's spread and making real progress toward eradicating the disease. Despite advancements over the years, none of the available tests provides all the answers related to the suspicion of TB infection or disease. In addition, most of these tests are not suitable or effective in many of the developing-country settings where they are most needed. While diagnostic equipment and techniques continue to advance, there does not seem to be a "magic bullet" on the horizon that will address the most serious limitations all at once.

Because a simple, universal approach for diagnosing TB does not exist, the diagnostics community has moved to apply a tiered approach that takes into account various populations and their living and working conditions, as well as the levels of care in settings where the new tests are to be used.

To foster the further development of the best-adapted tests, the research and development process should consider a critical although often-bypassed phase – needs assessment – which by definition measures in a systematic way the extent and nature of problems faced by the population where the new test will be implemented.

TARGET POPULATION NEEDS IN THE EPIDEMIOLOGICAL CONTEXT

Given the diversity of TB epidemiology, the needs will vary between regions of the world. For example, in some countries HIV co-infection is the most important medical condition among TB patients, while in others drug-resistant TB is the greatest concern – and in some instances, both of these problems are of great concern. In other countries (mostly low-incidence settings), the greatest need is to diagnose cases of latent TB.

There are four main situations where a TB test will be used:

1. **For case detection:** to detect active forms of TB (smear +/-, pulmonary/extra-pulmonary) in HIV-positive and -negative patients, in children and adults.
2. **For diagnosis of drug resistance patterns:**
   To detect drug-resistant TB (smear +/-, pulmonary/extra-pulmonary) in HIV-positive and -negative patients, in children and adults; and
   To identify resistance patterns for first-line and second-line drugs in drug-resistant TB.
3. **For treatment follow-up.**
4. **For diagnosis of latent TB.** Although a test is needed, its diagnostics relate to the biology of the host rather than the biology of the TB bacilli alone; therefore the diagnostic process differs from that for active TB.

WHAT IS EXPECTED OF A NEW DIAGNOSTIC TOOL?

For a given country and context, the main questions in identifying needs are:

> What are the delays resulting from a diagnosis based on the existing tools, and what are the consequences for public health and for patients in terms of death or default?

> Who are the patients misdiagnosed by existing tests and strategies, and what are the consequences for public health and for these patients in terms of death, default before diagnosis or over-treatment?

In practice, the only simple and rapid test that is available at peripheral level in developing countries is smear microscopy. Nevertheless:

> It requires a laboratory with skilled personnel;
> It is unable to detect less than 5000 to 10 000 AFB per ml;
> It requires sputum (so is not suitable for children);

Chapter 5

> It is not appropriate for extra-pulmonary TB forms;
> It does not allow diagnosis of resistant TB; and
> It does not distinguish between live and dead bacilli.

An essential characteristic of any new test, in addition to improving performance over existing tests, is that its use should result in a significant improvement in management of TB patients. A number of associated parameters should be considered of equal importance. For instance, in many situations simply improving sensitivity will not necessarily result in better diagnoses if the improvement also means longer time to results, less user-friendliness and/or increased complexity.

In addition, medical needs should be considered together with the constraints encountered at the various levels of infrastructure where the patients are presenting. A specific strategy for diagnosis and follow-up should be offered to subgroups of patients according to the specific needs they face in terms of time to result, type of sample and location of sample collection. In addition, the development of a test should pay special attention to the local laboratory infrastructures. Force-fitting technology into an unprepared setting has proven highly unsuccessful in the past. This reinforces the need for a flexible approach as evidence shows that one size does not fit all.

Médecins Sans Frontières (MSF) has generated a set of minimum specifications for a TB point-of-care test which can be freely accessed *(http://www.msfaccess.org/ TB_POC_Parismeeting/)*.

GATHERING EXISTING INFORMATION SOURCES

As it is likely that developers have to compromise between these parameters for a test to be developed, priorities must be set. The collection of the following data will help in adapting the technical constraints to the medical context.

Information relating to the epidemiological context of the region or the specific population to be served could include:

> Prevalence of latent and active TB, case notification rates, sex and age distribution;
> Breakdown data on TB cases: smear-positive/smear-negative/pulmonary/extrapulmonary/new cases/re-treatment cases;
> Treatment outcomes: cure rate and treatment success, failure rate, death rate, defaulter rate;
> HIV/AIDS incidence and prevalence;
> Data on resistance: MDR- and XDR-TB prevalence among new and retreated cases;
> Information related to the health care structure where the test is likely to be used;
> Availability of health facilities providing TB services;
> Availability of human resources: clinicians, prescribers, nurses, other trained personnel;
> User fees system applied to consultations and diagnostics;
> Average costs incurred by poor and less poor patients in accessing diagnostic and treatment services;
> Clinical algorithms and national protocols used locally;
> Endorsement of WHO policies;
> Average distance from patients' homes to health centres at which diagnostic and treatment services are available;
> Implementation of infection control measures; and
> Other barriers faced by patients in accessing care facilities.

Specific information about lab services is also useful, although new generations of tests might be used outside the classical laboratory premises:

> Organization of the national laboratory services in endemic areas;
> Availability of human resources and technical skills;
> Acceptable workload for properly trained staff in the context of all diagnostic services provided at their facility level, not only those for TB;
> Speed with which laboratories can provide results to clinicians requesting the test;
> Availability of biosafety equipment at each level of care;
> Supervision and quality assurance processes;
> Organization of the logistic supply chains;
> Transport capacities to next-level laboratories;
> Training capacities; and
> Package of lab tests routinely offered at each level of care and available technical platforms.

Critical information is available from international institutions (WHO, Stop TB Partnership, IUATLD) that can facilitate investigations. Health ministries and national TB programmes are also essential sources of information. Data from the most peripheral settings will often require field assessment.

INVOLVING TARGET POPULATIONS AND SERVICE PROVIDERS

To achieve a comprehensive assessment of needs it is essential to optimize data collection by meeting with patient groups, prescribers, laboratory staff and local stakeholders. A consultation of the entire range of stakeholders involved in the different steps of TB control would include:

> Caretakers in public and private sectors;
> Programme managers at all levels of health care;
> MOH central and local representatives;
> Patients' group representatives;
> National TB lab service staff and managers;
> National and international reference laboratories; and
> National and international academic institutions.

PUTTING NEEDS IN PRIORITY ORDER

Once medical and programmatic constraints are placed in perspective, the identification of priority subgroups of patients is based on the most critical gaps to be tackled:

Medical needs: Among the target populations mentioned above, the absolute priority remains the diagnosis of the patients at higher risk of death by misdiagnosis or delay, as in cases where there may be HIV/TB co-infection, where the detection of active TB for smear-negative patients (either adults or children) and extra-pulmonary TB cases is critical. In a context of high resistance, the fast detection of MDR-TB cases is essential for timely patient management.

Patient needs: Considering how important it is that tests be available at facilities located close to where patients live, many people in developing countries are out of reach of reference labs and district hospitals. Systems for referring patients to higher-level health care also are often weak. In these cases, the technology needs to adapt to the local infrastructure so the diagnostic strategy has maximum opportunity for success.

Public health needs: In situations where transmission can be induced by poor environmental health, one needs to be able to rapidly detect smear-positive cases. This applies to those living in crowded urban slum conditions as well as those in rural households with many persons living under the same roof. Categories such as prisoners, migrant populations and indigenous peoples should also be taken into account.

NEEDS ASSESSMENT: A CONDITION FOR PRODUCT EVALUATION

Tuberculosis is marked by diversity among target populations, disease presentations and health care infrastructures, as well as by diverse diagnostic and treatment capacities in settings where TB is a major health problem. This diversity calls for a systematic and comprehensive needs assessment to be performed early in the R&D process. It is only against this initial step that the relevance of a new product will be properly evaluated. Such data, if collected comprehensively, could provide a baseline for an assessment of impact.

The outcome of the needs assessment

The user requirements document is the expected outcome of a well-conducted needs assessment. This should be a list of specifications for the required tests according to different settings and indications. It would include the expected performance in real-life conditions, time to results, technical requirements, users' skills, medical algorithm within which the test is to be used and a clear description of the setting where the test is to be implemented. These specifications illustrate the tensions between the users' needs and the products' performance.

The most immediate outcomes are linked with the need to ensure high performance, ease of use for limited human resources, robustness and capacity to absorb a high number of samples using scarce laboratory infrastructures and equipment.

The main criteria to be considered are listed below.

Criteria for detection of active TB:
> time to result;
> ease of use;
> improved sensitivity and specificity (for ruling out latent TB and mycobacteria other than tuberculosis, or MOTT);
> type of sample; and
> safety (sample collection, laboratory use and disposal).

For diagnosis of drug resistance patterns:
> ease of use and time to result for rifampicin (R) and isoniazid (H) resistance;
> improved time to result of DST for second-line drugs;
> standardization of procedures in DST for E, Z, S;
> use at district level;
> safety (sample collection, laboratory use and disposal).

For treatment follow-up:
> differentiation between live and dead bacteria;
> detection of R and H resistance;
> detection of treatment failure in extra-pulmonary and smear-negative cases;
> test of cure; and
> safety (sample collection, laboratory use and disposal).

For diagnosis of latent TB:

> suitable performance in sub-groups (e.g. HIV patients, household contacts, children);
> ability of the test to predict future risk of active disease, and to identify individuals who are most likely to benefit from preventive therapy;
> suitable performance in sub-groups (HIV patients); and
> safety (sample collection, laboratory use and disposal).

The needs assessment should be clearly described in the product dossier (methodology and findings) with an eye toward the stakeholders contributing to the development of a new test. These critical findings will help donors and national programmes to decide about the relevance of a given project or final product.

Key points

* The diversity in TB epidemiology leads inevitably to diversity in needs, and therefore rules out a one-size-fits-all solution.

* Needs must be prioritized through a collaborative approach involving a wide range of stakeholders, including patient groups.

* The diversity among target populations, disease-presentations, health care infrastructures, diagnostics and treatment capacity of the different settings where tuberculosis is a major health problem calls for a comprehensive needs assessment to be performed early in the R&D process. It is only against this initial step that the relevance of a new product will be properly evaluated.

* The main outcome of the needs assessment is development of a list of specifications known as the user requirements document.

CUSTOMER REQUIREMENTS DOCUMENT

Example of a Customer Requirement document developed by FIND:

FIND
foundation
for innovative new diagnostics

Customer Requirement document
TUBERCULOSIS DETECTION, HEALTH POST LEVEL

Intended use:

Detection of active infection with *M. tuberculosis*. This is presently not confirmed in peripheral health clinics, and patients must be referred for confirmation. The primary goal is to provide a test that gets more patients on correct therapy sooner. This could happen through a test that extends the confirmatory capacity of microscopy, even with imperfect sensitivity, to the health post, or through a highly sensitive test that would allow active case finding and referral, and elimination of TB work-up in non-TB patients with suggestive symptoms.

Feature	Desired	Minimum	Comments
A - Workflow			
1. Sample			
1.1. Sample type	Sputum, skin, breath, urine	Sputum, skin, breath, blood, serum, urine	*Should work with at least one of mentioned sample types*
1.2. sample preparation	None	simple 1 to 10 two step procedure < 20 min per sample hands on time	
2. Time-to-result	< 10 min	< 2 hours	*Incl. sample prep and readout*
3. Instrumentation	None	None, or small, maintenance-free device	
4. Additional equipment required	None	< 2 additional devices, no maintenance required, portable, small total footprint (total space for TB testing <2m2)	
5. Biosafety	No need for biosafety cabinet Disposal of clinical material without further treatment	No need for biosafety cabinet	
B - Performance			
1. Sensitivity			
1.1. Diagnostic sensitivity	**Either:** **a)** (rule-in test) overall sensitivity >75%; including 90% of smear-positive patients ;or **b)** (rule-out test) overall sensitivity >95%, including 100% of smear-positive patients;	**Either:** **a)** (rule-in test) overall sensitivity >50%; including 75% of smear-positive patients; or **b)** (rule-out test) overall sensitivity >75%; including 95% of smear-positive patients	*Determined in microscopy and culture-confirmed TB cases*
2. Specificity			
2.1. Diagnostic specificity	**Either:** **a)** (rule-in test) > 98%; or **b)** (rule-out test) >75%	**Either:** **a)** (rule-in test) > 95%; or **b)** (rule-out test) > 50%	*Determined in TB suspects confirmed not to have TB by negative cultures (X2) plus negative microscopy (X2) AND either improvement on treatment other than for TB OR confirmation of an alternate cause of symptoms.*
			Must not require exclusion of patients with prior BCG-vaccination, TB exposure, NTM infection, cured TB or HIV infection in order to meet specificity requirements

Chapter 6

Aiming for the right targets

BASIC RESEARCH – WHAT EXACTLY DOES IT ENTAIL?

Basic research involves looking at nature and asking how it works. Applied research is when developed basic research is tested. Accordingly, basic and applied research feed into each other – as applied research may also provide answers about how things work while raising further questions. A good example of this is micro system technology and nanotechnology, where new electronic properties of molecular systems are discovered on an almost weekly basis. Serendipity, too, plays a large role in many discoveries in both basic and applied research, providing insights into how nature works.

BASIC RESEARCH: PROMISING TRENDS

In tuberculosis, the explosive growth in new techniques for research has the potential to revolutionize our approach to diagnosis, particularly through providing deeper insight into the structure, composition, metabolism and behaviour of the mycobacterium and the means by which TB becomes resistant to therapy. This section looks at the ways these new techniques will likely yield more accurate and rapid diagnosis of active infection or drug resistance. The focus is on promising trends in the search for better TB diagnostics.

A NEW ERA OF TECHNOLOGICAL INNOVATION

We live in an extraordinary time in scientific history. The convergence of multiple disciplines – such as physics, nanotechnology, nanobiotechnology or molecular biophysics, molecular biology, micro systems (or micro-electromechanical systems) and biomedical sciences – will revolutionize many research areas. The discovery of new scientific knowledge and the development of new tools for detection, treatment and prevention of disease will shift to the molecular level. This is equally true for new TB diagnostics.

Biological systems can provide guidance for the scientist wishing to develop new tools. For example, knowledge of the receptors, pathways and chemicals involved in the detection of pheromones via olfactory receptors, neurotransmitters and signal transduction may lead to methods that detect chemicals in the aqueous or vapour phase of relevant substrates in TB infection.

New techniques for the detection of minute amounts of material, likely involving single molecule detection, will become possible. Cantilever beams and nanotube paddle resonators that can detect single molecules are being developed [30]. Antibody functionalized silicon-nanowires will be used for the detection of analytes using changes in conductance [31]. DNA-decorated carbon nanotubes can be used for chemical sensing of gases [32], and dye-labelled DNA can help detect volatile organic compounds (VOCs) [33]. These new detection systems could be refined to produce "lab on chip" systems for the detection of relevant biomolecules.

Enormous advances in techniques for detecting substances present in body fluids at very low concentrations offer hope for future diagnosis and identification of active mycobacterial infection. However, the new level of precision brings its own problems, especially with regard to sample acquisition and preparation. The need for rapid and accurate diagnosis is most acute in the lowest-resourced countries and in areas where MDR-TB is a problem. Accordingly, our challenge is to refine today's elaborate and costly methodology into an inexpensive, sensitive point-of-care test that requires minimal technical expertise for its execution – while maintaining the high accuracy of the current models. For example, if the Ziehl-Neelsen test could be superseded, it is likely that the rapid progress in technology would lead to the identification of a group of biomarkers that could be used for a simple diagnostic test within five years.

A strong interdisciplinary approach to the development of these new diagnostics tests is recommended.

WHAT ARE WE AIMING FOR?

Selecting targets for diagnostic tools

Target selection is based on what is simple and appropriate for the setting. For example, it is possible to perform microscopy at health centre level, but it is not sensitive and requires a trained microscopist for reading. Culture and DST can be performed only in a tertiary level due the to costly infrastructure and equipment requirements and require highly trained laboratory staff. The trend is to develop point-of-care tests that can be applied near the patient through simple applications such as breath analysis, biomarker detection, lateral flow tests or immunochromatographic tests with sputum, urine and serum.

The currently used direct (detecting the organism itself) and indirect (detecting the organism's effects on the host) assays are described in greater detail below.

Direct assays
Microscopy

The enhanced sensitivity of fluorescence microscopy compared with ordinary light microscopy has long been recognized. Recently low-cost light-emitting diodes (LEDs) with very long lifespans have been used in fluorescence microscopy [34]. The method has been successfully validated in many highly TB-endemic countries [35, 36].

Further developments are expected in other areas:

- The use of automated microscopy using scanning microscopes or high-resolution CCD (charged couple device) camera and special lenses that can be developed into a robust, low-cost point-of-care diagnostic tool [37];
- Raman spectroscopy [38, 39];
- The use of bioconjugated nanoparticles in combination with the scanning devices described above [40, 41];
- Simple flow cytometric devices (Luminex);
- Concentration of bacteria in sputum using sieve techniques;
- Digestion (bleach);
- Sputum filtration;
- Vital fluorescent staining of sputum smears;
- Use of functionalized paramagnetic particles for cytometric detection [42];

- Ultrasensitive detection using the Young interferometer sensor [43]; and
- Utilization of the autofluorescence of *M. tuberculosis* [44].

For all these new methods, we need improved sputum sample treatment procedures. The difficulties presented by sputum's complexity are further discussed under antigen detection.

Culture

Culture is still the gold standard for the diagnosis of TB. The automatic liquid culture systems are much faster than solid culture media. However, liquid culture techniques are more prone to contamination. Their use has caused a dramatic increase in the isolation of non-tuberculous mycobacteria. This may have the unfortunate consequence that the patient is wrongly treated for TB, or even for MDR-TB.

A positive liquid culture absolutely requires an *M. tuberculosis* specific identification test, e.g. a lateral flow test or a standard biochemical testing panel. Specific amplification tests are commercially available and further developments are imminent. For instance, the direct detection of *M. tuberculosis* specific volatile organic compounds (VOCs) in the headspace of the liquid culture vials is possible with the use of a combination of GC analysis with a surface acoustic wave detector or ion mobility spectrometer (z-Nose; Femtoscan; Sionex) [45], and with multi-purpose sensing arrays on small chips [46]. A test based on the volatiles CO_2 and NH_3 and detection by a micro-electromechanical system (MEMS) has recently been described by Ren and coworkers [47].

For the microscopic-observation drug-susceptibility (MODS) assay, future research will be conducted in the field of scanning inverted light and fluorescence microscopes that can identify the characteristic cord formation of *M. tuberculosis*.

Amplification reactions DNA and RNA

The nucleic acid amplification test (NAAT) has become a standard technique for the diagnosis of TB. The test combines the rapidity of microscopy and the sensitivity of culture in smear positive specimens, but

culture is still the gold standard for the diagnosis of TB since it is more sensitive, for microscopy-negative sputum samples. The sample treatment and the amount of sample used in the amplification reaction are key elements for the sensitivity of the test. In commercial tests, this is often a compromise (a strong point for in-house PCR assays). In developing countries, the introduction of NAAT in routine settings is hampered by the relatively high workloads and skills that are needed for the test. There are three steps involved in the assay: **(a)** specimen processing (depending on the kind of specimen) and nucleic acid extraction, **(b)** nucleic acid amplification and **(c)** detection of amplified products. Research is needed to reduce technicians' workloads. More-automated systems and simpler amplification reactions such as loop mediated isothermal amplification (LAMP) are under development. With the emergence of MDR-TB and XDR-TB, the role of NAAT for rapid drug resistance will rapidly increase.

An important research area will be molecular assays to detect gene mutations. Existing tests are line probe assays. These tests should be extended to the detection of quinolone resistance and made more user-friendly. Other amplification tests, such as multiplex ligation-dependant probe amplification, seem promising both for the rapid typing of *M. tuberculosis* and for drug resistance measurement [48]. A new approach to DNA detection using recombinant proteins is the recombinase polymerase amplification (RPA), which couples isothermal recombinase-driven primer targeting of template material with strand-displacement DNA synthesis [49]. An important drawback of these techniques is that they do not allow to distinguish live from dead bacilli.

Antigen detection, point-of-care tests and 16S rRNA

Developments in antigen detection assays are promising. Two types of commercial tests based on antibodies are available: immunochromatographic tests for the identification of *M. tuberculosis* cultures (detection limit $1x\ 10^7$/ml) and the lipoarabinomannan (LAM) assay for the detection of LAM in urine.

The challenge is to develop a simple and inexpensive test with at least as good a detection limit as direct microscopy $1x\ 10^4$ bacteria/ml – indeed, this would reduce the workload of laboratory personnel. The development of such tests is hampered by the complexity of the sputum matrix sputum. Sputum is highly viscous and inhomogeneous. For antigen detection assays based on protein and non-protein antigens, sputum is a very difficult sample, as inhibition in immune assays often occurs. Research to improve sputum preparation is based on its physical and chemical properties. The available methods include chemical liquefaction, enzymatic degradation ultrasonication, shear forces, disruption with beads, heat, electroporation and filtration. So far it has been difficult to achieve sputum solubilization without destroying the mycobacterial protein targets.

Future detection systems

Surface plasmon resonance (SPR) sensing is one promising label-free technique that has found its way into practical applications, with more than 1000 publications annually and more than 20 companies developing SPR-based biosensors systems in sensitive detectors [50]. SPR requires sophisticated optical instrumentation; this has stimulated research in other sensor types such as complementary metal oxide semiconductor (CMOS) field effect transistors [51] and nanowire electrochemical sensors [52]. A fast, portable, ultrasensitive Young interferometer sensor also has been described [42]. It is expected that fully integrated point of care SPR imaging-based diagnostic systems will be introduced in the next five years [53, 54].

Use of ribosomal RNA as target

16S rRNA is another promising target for detecting mycobacteria. The advantage of rRNA detection is that this target is much more abundant (at least 1000 more copies) than DNA. This makes detection of rRNA possible without amplification. Isolation of rRNA is possible even after harsh preparation of the sample, such as treatment with NaOH to decontaminate the sputum. One problem is how to avoid the RNAse activity (highly thermoresistant) that is present on all surfaces. Sputum itself contains lots of RNAse activity that must be inactivated before isolation

of 16S RNA. Detection of rRNA is based on hybridization with a membrane-bound capture probe and a biotin-labelled detector probe followed by an enzymatic detection system (AHJ Kolk, personal communications).

Use of nanoparticles

A possibility for improvement would be the use of silicon [55, 56] or carbon nanoparticles [57, 58]. Quantum dots 1-6 nm nanoparticles can be functionalized with capture probe or antibodies [59] coated with capture probe. These nanoparticles coated with capture probe display a shift in photoluminescence upon hybridization, providing a label-free detection system. Mirkin and coworkers developed a promising nanotechnology-based platform for multiplex detection of protein and nucleic acids using a sandwich of magnetic and bar-code probes [60].

Electrochemical detection

Electrochemical methods are among the most popular electronic techniques for biosensing [61]. The sensitivity of electrochemical detection (ECD) is lower than conventional laser-based micro array fluorescence techniques. A variant of ECD is silicon nanowires for label-free biomeolecular detection [52]. A multi-walled carbon nanotube-based nanoelectrode array can be integrated into an electrochemical system for sensitive detection of DNA [62]. Carbon nanotube electrochemical sensors and "Kelvin probes" are label-free systems for biosensing [30; 63].

The challenge of miniaturization

One of the challenges for the development of a rapid test is the need for reactions in solutions with high analyte concentrations. Hybridization reactions and protein capture assays on membranes take a long time due to diffusion of the analyte to the membrane, and the signal depends on the analyte concentration. The problem using micro- and nano-scale detection is to go from the macroscopic world (e.g. a sputum sample of 2-5 ml) to the representative micro- or nano-litre sample of sputum. In order to achieve this, a highly efficient concentration step in the whole sample treatment process is required. A potential platform technology for this is the use of silica or functionalised paramagnetic beads [64, 65, 42].

HPLC and GC-MS to detect mycolic and fatty acids

Gas, thin-layer and high-pressure liquid chromatography have been used for decades for the species identification of isolated mycobacteria [66]. However, the use of chromatographic methods for the identification of mycobacteria has been restricted to larger reference laboratories and research settings. Reasons for this include the complexity of the procedure, as it requires hydrolysis, esterification and extraction, time to prepare samples and equipment cost and maintenance. Most methods also require a culture step prior to analysis of the clinical samples.

The group of Sandra and co-workers [67] used a method that avoids the extraction step usually done by hand. They used hydrolysis and esterification combined with stir bar solid-phase micro-extraction of decontaminated sputum prior to gas chromatography-mass spectrometry (GC-MS) analysis of tuberculostearic acid (TBSA). Kaal and coworkers [68] developed a fully automated procedure for the detection of TBSA and hexacosanoic acid in culture and sputum from TB patients. Brooks and coworkers [69] developed a quantitative chemical ionization gas chromatography MS and frequency pulsed electron capture GC (FPEC-GC) method to detect TBSA. Daikos and co-workers [70] used this method for the detection of TBSA in CSF and serum of patients. If these methods can be further simplified, especially sample treatment, this approach could become a reliable system for the detection of TBSA, mycolic acids or other organic compounds unique for mycobacteria. This is especially likely if dedicated instruments are developed, such as portable GC-MS and micro GCs with sensitive detectors; examples of these are the surface acoustic wave (SAW) detector z-Nose and the ion mobility detectors, e.g. Electronic Sensors Technology, Sionex and GAS.

Volatile organic compounds from clinical samples

New research is being done on using electronic noses (E-nose) and GC-MS, GC- ion mobility detection (GC-IMD) and GC- surface acoustic wave detection GC- SAW to identify and characterize VOCs in sputum and breath from TB patients.

An E-nose is the colloquial name for an instrument made up of chemical sensors (polymer, metallic oxide, bulk acoustic wave or surface acoustic wave sensors) with a pattern recognition system.

The reversible adsorption of VOCs to the sensor surface leads to a change in physical properties (change in resistance and frequency of the sensor) that can be measured. The E-nose mimics the human nose – it must be trained to recognize an odour pattern. However, unlike gas chromatography, it cannot separate the mixture into individual components. The E-nose has been applied mainly in the food industry. The potential of an E-nose to detect bacteria including *M. tuberculosis* in culture and sputum has been studied (Fend et al., 2006) [71]. So far, the existing E-noses do not fulfil the requirements needed to replace microscopy. To further improve the E-noses for this purpose, it is necessary to characterise and to quantify the VOCs present in the headspace (air above the sputum and cultures when sealed in a container) [72]. With this knowledge, dedicated VOC-specific E-nose sensors of polymer or metal oxide can be developed.

VOCs in breath might provide new biomarkers of active pulmonary tuberculosis [73]. The VOCs may be metabolites from the mycobacteria from the host's metabolism, including bacteria in the gut or the host's response to the mycobacterial infection. VOCs in breath have been captured onto sorbent traps and analysed by automated thermal desorption GC-MS, obtaining encouraging results after fuzzy logic and pattern recognition analysis [74]. Analysis of breath or exhaled breath condensate is a completely new method for the detection of disease. We can expect further research in the area of VOCs because a low-cost, accurate, rapid and non-invasive technique that is also useful for the TB diagnosis in children would immensely help tuberculosis control efforts.

Indirect assays

Antibody detection test

This would be the ideal point-of-care test using a simple ELISA or lateral flow format. Unfortunately, such a test does not yet exist despite the many academic and commercial researchers who have worked towards one. Existing commercial serological assays have inconsistent and poor accuracy, and it is not clear if such a test can be developed for clinical use. One may speculate that a better understanding of TB-specific biomarkers should help in the production of such tests for clinical use.

Tuberculin skin (Mantoux) test

The use of this test is limited by its lack of specificity, especially in highly endemic countries. PPD will be replaced in the future by specific antigens only found in *M. tuberculosis*. Recombinant dimer ESAT-6 (rdESAT-6) has been successfully tested for this purpose. Improvements of this new skin test continue to be developed using other antigens such as CFP-10 [75].

IFN-gamma release assays

These blood-based tests use specific antigens such as ESAT-6 and CFP10 and are useful for the detection of infection, but cannot distinguish between latent and active TB. Their use in low-incidence countries is growing rapidly, but their role in developing countries with a high burden of TB remains unclear. Perhaps these tests can help to establish a diagnosis of active TB among a selected group of people – such as young children, the immunocompromised and patients suspected of TB with a negative smear or extra-pulmonary disease. The test can be simplified for resource-poor settings with the use of the recently developed image cytometer [42].

Promising biological systems research

Proteomics and metabolomics are rapidly growing research fields. Proteins regulate chemical reactions in the body and the activities of cells, tissues, and organs. Understanding the proteome or composition of cells, tissues, organs or organisms is key to understanding disease.

Krishna and co-workers [76] recently described identification of diagnostic markers for tuberculosis by proteomic fingerprinting with serum profiling by mass spectrometry coupled with pattern recognition methods. They obtained proteomic profiles from patients with active tuberculosis and from controls by surface-enhanced laser

Chapter 6

desorption ionisation time of flight mass spectrometry (SELDI-TOF-MS). The support vector machine SVM classifier discriminated the proteomic profile of patients with active tuberculosis from non-TB patients with a diagnostic accuracy of 94%. Proteomics research could lead to the identification of a set of proteins or biomarkers specific for tuberculosis. This could lead to a serum-based antigen detection assay for the diagnosis of TB metabolomics, which involves the study of the metabolites of a cell or organism. Translating this metabolome data into biologically relevant information is a promising research area [77]. Genome sequence data can be used to predict possible metabolites, and this knowledge can be used to interfere in metabolic pathways. VOC analysis of cultures of mycobacteria and metabolomic data could help in developing a direct identification test for mycobacterial cultures based on mycobacterial-specific VOCs in the headspace of the mycobacteria container. Furthermore, a direct test with clinical material perhaps can be produced if metabolites or mycobacterial compounds are also found in TB patients' sputum, serum or urine.

Key points

* Tools that can function at molecular level show great potential for new TB diagnostics. However, a limitation of these techniques is their inability to distinguish live from dead bacilli.

* This precision can create difficulties in that an elaborate, expensive yet accurate methodology needs to be translated into tests for use in resource-poor settings.

* Whether through direct or indirect assays, each test has advantages and drawbacks depending on epidemiological context, geographical region, health care infrastructure and health policy. These considerations should not limit the scope of basic research, however.

* Opportunities are now present as never before for the development of innovative technological solutions.

Chapter

7

Feasibility –
A guarantee of strong foundations

GETTING THE CONCEPT RIGHT FROM THE START

The strong foundations which the feasibility phase is designed to lay down depend on the vision for the product that will have been scoped out in the concept definition document. This sets out a complete vision for the product: what it would do, who would be the target users and their needs, how users would act on the information it provides, in what tiers of the health care system the product will be used, what technologies might be employed, how it could be manufactured and strategies for marketing. The concept definition document should also include an initial assessment of the geographical regions where the product will be sold (i.e. high burden countries, any resource-limited country and/or established market economies).

For commercial development, there needs to be a business plan. This includes initial estimates of product manufacturing costs, selling prices, unit sales, cost-to-serve, customer support, profit projections and project timeline. These can be used to demonstrate how the project satisfies company or organizational objectives.

A CLEAR VISION FOR STAKEHOLDERS

The vision for the product must also be detailed enough to make a compelling case to donors and investors. This may include drawings of prototype devices or algorithms of how testing would affect patient care. Consideration should be given to whether the funding, development, manufacturing and/or marketing will be done internally or contracted to external contractors or partners, in which case there should be due diligence in selecting these. The selection of a contracted resource or collaborator is normally followed by a written agreement – this can be done early or later in the development process. Typically, a binding agreement is milestone-driven based on risk reduction as the project moves forward. Confidentiality agreements will also be required to protect ownership interest and any confidential and proprietary information exchanged in order to facilitate the partnership.

REDUCING RISKS IN TEST DEVELOPMENT WORK

The feasibility phase precedes any test development work. This phase determines whether a specific technical solution may feasibly be adapted for the specific diagnostic use required. The questions asked in the feasibility phase are driven by the user requirements document, or by a similar document defining user needs, as described in the needs assessment section.

The feasibility phase is meant to be a relatively inexpensive way to decrease risk prior to product development, to make choices between different technical solutions and to arrive at go/no-go decisions for full product development. However, success or failure in feasibility usually determines whether or not product development will ever take place, so definitions and targets must be clear, and studies must be robust enough to direct go/no-go decision making. Activities during the feasibility phase should be described in a *feasibility plan* describing the objectives, timelines and budget for this phase. The feasibility study may examine technical questions such as the ones that follow, or it may examine logistic questions related to the envisioned implementation of a final product.

SELECTING TARGETS

The first step in feasibility is the establishment of targets, which may be many or few, but in any case should include all the qualities of the selected technology that are intended to be assessed in this phase. As such, the feasibility target specification becomes the blueprint for the design of any trial in the feasibility phase. A design history file will have to be created to comply with international regulatory requirements for product development.

Feasibility studies should focus on those features of the technology that carry the highest risk for the realization of the

product characteristics, or that are the most critical to success of the final product. For example, feasibility studies of an established technology that is being used for the first time as a multiplexed assay should focus on the capacity and limitations surrounding its use in a multiplex format.

Some examples of questions that might routinely be addressed in the feasibility phase:

> Is it feasible to adapt an existing assay for use in a different target population, with a different specimen type, or in settings with limited infrastructure?

> Is it feasible to introduce a system of specimen transport from peripheral laboratories to central laboratories to implement a high-throughput testing method?

> Will it be practically safe to use in its intended setting?

> Is it feasible to lyophilize or otherwise stabilize the reagents needed for a specific technology to make it robust enough for implementation?

> For a technology that is functional but requires bulky equipment, is it feasible to miniaturize the equipment?

> For a test that might require ancillary clinical information (such as HIV status) in order to interpret results, is it feasible to perform this testing simultaneously?

> Does the manufacturing process represent some inherent risks which at the end would make the product non-manufacturable or manufacturable only at an unacceptable price?

Though these examples list single questions, most early technology assessments will address multiple questions at the same time. The feasibility targets may list anywhere between five and fifty parameters. Administrative and legal issues, including intellectual property considerations, also may need to be considered during the feasibility phase.

Legal assessment for intellectual property (IP) must be done to determine if there is freedom to operate or if licenses for IP will be required. There may be IP in areas besides

base technology, such as manufacturing processes or a physical design of the test device or instrument. The assessment should include IP status in the country of manufacturing and all the major regions or countries that product will be sold in.

If IP exists, plans need to be made to license the IP or purchase it from a licensed vendor. If new IP is generated, steps may be taken to protect the new technology. This ensures that access to key material and information is guaranteed for parties interested in the development and in production of tools for populations in need.

The plans described below are developed in the feasibility phase, although execution of those plans will occur in later phases. Each plan should consider whether the organization will resource these tasks itself or will pursue using external contractors or partner resources. External resources can be for funding or for personnel.

THE FEASIBILITY PLANS – MORE THAN R&D

A clinical trial plan outlines the studies needed for feasibility and optimization, and eventually for final product performance claims as well as additional studies for individual country registrations/approvals. The plan should cover the number of test results needed, potential site locations and compliance with international clinical trial standards. Additionally, if specimens are needed to perform studies in-house, where and how those specimens will be obtained should be considered.

A preliminary user support plan outlines how the product will be serviced, how users will be trained and technically supported, and how complaints will be handled. If a distribution partner will be sought for these steps, it should be indicated here.

For commercial tests

A regulatory plan needs to be established based on where the product will be manufactured and where it will be sold. At this phase, this is a high-level plan anticipating **(a)** what will be needed to

register the product or obtain regulatory clearance in target countries and **(b)** what regulations for development and manufacturing will apply and require compliance.

A marketing plan outlines where the product will be sold, how and who sells it (distribution) as well as the unit sales forecast (affects manufacturing strategy and cost). As this is a key component of the overall business plan, pricing and distribution also should be considered. If a marketing partner will be sought for these steps, this should be indicated here.

Beyond this, more than one technology may be assessed in feasibility. A typical activity in the feasibility phase is the screening and comparison of various technical alternatives to reach a specific aim (e.g. whole bacterial concentration/ detection in sputum). As the entry point to product development, the feasibility phase should refine to the degree possible the reagents and conditions needed in the final product design. Above all, the feasibility phase should allow the developer to assess the risk that user needs, as defined by the user requirements document, will not be met, or that the product will not reach or not serve the intended populations. The go/no-go decisions on entering the next phase and proceeding to the next technology choices will be made based on that risk profile,

the potential benefit of the technology if successful, the calculated time and cost required to get to final product and the strength of competitive technologies.

The size of clinical studies required during the feasibility phase varies with the question being asked, but such studies typically enrol 100 to 600 patients.

The outcomes of a feasibility phase are:
> A go/no-go decision to enter the proper development phase;
> A product requirement document which is based on the initial user requirements and the technology limitations;
> A design history file;
> A clinical trial plan;
> A preliminary user support plan.

For commercial tests:
> A marketing plan;
> A regulatory plan.

Key points

✱ The key to strong foundations for test development is the concept definition document which takes a holistic view, including such elements as manufacturing, pricing and distribution.

✱ Investors and donors need to identify strengths and hidden risks in progressing from prototype development to feasibility testing.

✱ Feasibility entails reducing risks prior to product development. This phase must produce clear answers and a conclusive go/no-go decision.

✱ Feasibility should be placed in the context of under-resourced settings where the finished products will be deployed.

✱ The feasibility plan will include a number of sub-plans and may include a comparative assessment of alternative technologies.

Chapter 8

Development and optimization: *additional hurdles*

PROOF OF PRINCIPLE

The development and optimization phase is when detailed design and development activity are most focused. The product is refined using rapid prototyping and user feedback. Iterations of the product are normal and can occur for a variety of reasons. This phase initiates the manufacturing processes, user support processes and a detailed launch plan. All plans from the feasibility phase are updated and execution of the plans begins. Estimates of product costs, selling prices, quality/reliability measures, unit sales, cost-to-serve and support users, and project timeline are updated and critically analyzed as part of an updated business plan. A fully validated design-locked product that meets the specifications outlined in the user requirements document will be approved before moving into the next phase.

ACTIVITIES AND DOCUMENTATION

R&D activities during the development phase are broken down into a) detailed specifications, b) high-level design progressing to detailed design, and then c) optimization and verification/validation. A variety of appropriate statistical analysis methods need to be used throughout this phase.

R&D tasks and considerations include:

> Key product attributes for users that need to be addressed in the design are quality, cost and delivery. Among other qualities, products should be easy to manufacture/assemble. They should be designed to be effectively and efficiently tested prior to release to ensure quality, and comply with any hazardous material shipping criteria. Finally, they should reduce service costs by minimizing service calls, downtime and/or user training.

> In cases where target markets include different countries, the design should make allowance for international considerations such as varying regulatory standards, power requirements, safety and environmental standards, operating conditions and language requirements for labelling and user interface.

> Details of the technology protection strategy should be finalized.

> A prototype should be developed and put through preliminary testing.

> Efforts should focus on optimizing product workflow, performance and/or manufacturability, including overall performance as well as performance in key areas such as smear-negative samples or samples from HIV+ patients. Improvements to workflow could simplify or reduce the number of steps. Iteration should occur early in the development cycle, before significant investments have been made.

> Verification and validation (V&V) are executed to ensure that specifications were achieved and that any deviations from specifications are documented in the final V&V report. This includes user variability testing, to ensure that the product is robust in a variety of different users' hands, and also determining analytical performance, such as limit of detection, specificity and variability with critical parameters.

Clinical studies done in this phase are field evaluations using a partially validated prototype product and are part of the optimization process. These studies are typically smaller and less complex than those in the next phase. Each study requires planning, managing, staffing and data analysis. International Conference on Harmonization Good Clinical Practice (GCP) and other regulatory standards are followed in this phase just as in the evaluation phase *(Chapter 10)*. As optimization proceeds, the product is tested on different patient populations, such as smear+ and smear−, HIV+ and HIV−, and compared to a defined diagnostic algorithm including clinical and bacteriological diagnosis.

Labelling will be developed for field evaluations, for clinical trials and eventually for final product. This consists of labels on test devices, reagents, kit cartons and instruments as well as package inserts and user manuals. Consideration must be given to international standards as well as language needs.

There must be a successful transfer of product design to manufacturing. Typically, a *design transfer package* is compiled. This collection of documents is required to translate a product from R&D to manufacturing so the product can be routinely reproduced to a predefined set of specifications.

The phase also includes implementing procedures for development of expiration dating and storage conditions for raw materials, work in process and finished product. This includes test methods, sampling plans and acceptance criteria. Considering the effects of raw materials and process variation on product stability, documenting such variabilities in these areas can help reduce stability failure due to inherent inconsistencies.

Manufacturing tasks include developing manufacturing processes for reagents and instruments (if applicable) to sustainably manufacture quality products predictably supplied at the targeted cost. This includes procurement of raw materials and establishment of reliable suppliers and procedures. Manufacturing validations ensure a consistent product at the extremes of the process operating window as well as in the optimum window.

Quality assurance planning outlines an approach to evaluate the new processes and product. Specific activities include evaluating safety/risk; verifying systems and hardware; assuring verification of the design; documenting validation of the design and manufacturing processes, and of qualifying materials/components as well as compliance with the relevant quality systems used in development phase. Compliance to international product development requirements or standards is useful and often required. These include FDA registered (US), ISO 13845:2003 for medical devices or ISO 9001:2000 (Global), the European IVD directive, the EU Medical Device directive and PAL in Japan.

VALIDATION LEADS TO FINAL DESIGN-LOCK OF THE TEST

Validation of the final version of the prototype is a series of tests required to demonstrate the device or system meets all the criteria listed in the user requirements document. Once successfully completed, the design can be locked down and finalized.

Validation and verification of the product and its subsystems or components must be delineated in process, execution and results. This includes V&V plans and results, traceability matrix, design reviews and written summaries of these efforts.

The actual validation is guided by the user requirements document, and each key requirement should be tested. Appropriate statistical analysis and acceptance criteria are decided in advance and included in the protocol or testing plan. Testing is typically done in triplicate, using multiple different lots of materials. During validation, standardized samples are often used.

Validations and verification experiments establish expected boundaries for stability, reliability and tolerance characteristics. This includes:

> Interfering substances from sample matrix or sample processing;
> Specimen stability (both accelerated and real-time);
> Reagent stability;
> Testing environment where the test will be used (temperatures, humidity);
> Stressed temperature stability which simulates shipping conditions;
> Read times and tolerances;
> Inoculum or sample size tolerances;
> User variability testing;
> Limit of detection;
> Performance, including sensitivity, specificity and reproducibility; and
> Equivalency to any earlier versions of prototype as changes are made.

Additional validations are done for manufacturing and include:

> Laboratory standards and controls used by manufacturing;
> Manufacturing hold times for components during the actual manufacturing process;
> Stability of raw materials or components; and
> Testing of incoming raw materials.

Now that a validated, design-locked prototype is completed, the next phase is ready to begin. For a product to be deployed, it is critical to perform unbiased and accurate evaluation studies. These should be used for product performance claims in the labeling (package insert) and for product registrations and regulatory clearances. If a developer does not have the resources or expertise to perform and manage the evaluation phase as outlined in Chapter 10, it makes sense to have this phase done by a partner organization.

Key points

* The outcome of R&D during development and optimization will be a design transfer package which sets out how the product can move to manufacturing on the basis of a predefined set of specifications.

* Quality assurance planning enables evaluation of all aspects of the product and ultimately its conformity with established international standards.

* Validation of a prototype results in a design-locked device or system that is to be used in the next phase.

Chapter 9

Evluation, putting it through its paces: *does the test work?*

EVALUATION OF READY-TO-TEST PRODUCTS

Once a test is developed and optimized and proof-of-principle data are obtained on the prototype product, the design is locked and ready for the next phase: an evaluation study of accuracy, reliability and performance outcomes. In general, diagnostic research studies can be conceived as a series of designs with increasing sophistication, ranging from early pre-clinical studies in the laboratory to evaluation and impact/demonstration studies in clinical and field environments.

The major activities during the evaluation phase of a product is to plan and conduct an evaluation study to assess performance of the design-locked in vitro diagnostic (IVD), and to judge these results against target product specifications and gold-standard techniques. The high-quality data generated during this phase will be used as performance claims in product labeling and may be used for IVD registration with regulatory agencies. Expected outcomes are data on the performance of a test in the conditions of use, which then becomes critical information for the users of the test within laboratory services.

Important parameters to be taken in account during the evaluation phase are:

> Goals of the evaluation studies;
> Standardization of the study protocol;
> Targeted population for the use of the test;
> Intended use of the test requiring studies to be performed within a relevant context; and
> Bias due to conflict of interest between developers and evaluation team.

Evaluation studies are not meant to assess only commercial tests. Tests developed by academic institutions should be carefully assessed as well, but these may not have gone through the same development process. In any case, evaluation studies should be disconnected from the commercial goals and regulatory steps of an R&D project.

MAIN OBJECTIVES:

Step 1: Assess performance of a new test against the gold-standard technique on well-characterized samples or in a case-control setting where known TB cases and healthy controls are typically selected. This can be done retrospectively if clinical data and specimens are already collected and stored. However, these types of case-control comparisons often overestimate diagnostic accuracy and can provide a misleading impression of a test's accuracy. Thus, subsequent validation is needed in clinically indicated target populations. Unfortunately, commercial tests often are approved for clinical use solely on the basis of accuracy data from case-control studies; such tests often underperform when applied in routine clinical practice.

Step 2: Prospectively assess a new test's clinical performance in realistic settings in patient populations where the test is clinically indicated. This avoids the case-control situation, but instead is done in consecutively recruited individuals in whom the test would normally be indicated.

Step 3: Assess the reliability (reproducibility) of the test. This is important because a test that is accurate may not necessarily be reproducible, especially in field conditions. Reproducibility can be measured between operators or technicians (inter- and intra-observer reproducibility), between different test sites or laboratories, using different instruments, between different kit lots (lot-to-lot reproducibility) or on different days (run-to-run and within-run reproducibility), and within-person variations over time (i.e. variations in test results in the same person tested serially over time). Assessment of reliability can be as important as assessment of accuracy, especially for tests that have a subjective component or have complicated designs and steps. Reliability is likely to be a bigger concern with poorly standardized in-house (non-commercial) assays. For this reason, assessment of reliability must always be included in evaluation study protocols. Reliability assessment, where possible, should be done in a blinded fashion.

Step 4: Based on the data collected in Steps 1 through 3, plan further modification and improvement of the test.

Step 5: For commercial products: If accuracy and performance data meet specifications, prepare and submit data package for registration of the product at regulatory agencies based on

Chapter 9

a pre-defined regulatory strategy (objective for manufacturer). Approve a plan to launch and distribute the product and finalize actual data of product cost, selling price, quality/reliability measures, unit sales, cost-to-serve and support the user and validations included in an updated business plan.

CRITERIA FOR JUDGING SUCCESS

> Product performance specifications proven in target patients/populations;
> Technology proven suitable for the targeted setting;
> Data collected for publication of successful results and failures;
> Data collected for improvement of the test if needed;
> Data collected for demonstration and impact studies to be planned; and
> For commercial products: product registered with at least one regulatory agency (such as CE Mark) and ready for general sales.

EVALUATION: LAB STUDIES AND CLINICAL TRIALS

An evaluation study is not performed on a prototype or a beta version of a new test, but rather on a well-standardized product that has completed development, is design-locked and ready for scaled up manufacture or widespread use. It is conducted using a pilot production run of the product and final processes and documentation. Such studies are often intended to produce high-quality data for use in registering products with national regulatory agencies. This data must be collected in a manner conforming to relevant external quality standards (such as ISO 13485:2003 and the International Conference on Harmonization (ICH) Good Clinical Practice). Guidelines on how to conduct studies of infectious diseases diagnostics are available from TDR/WHO [78]. These Diagnostics Evaluation Expert Panel (DEEP) guidelines concern best practices for such research [78].

Among all IVD trials, evaluation studies are generally the most demanding for the trial sites with respect to clinical information needs (e.g. HIV/CD4 status, chest-X-ray, clinical follow-up). These studies require gold-standard resources (e.g. solid and liquid culture, blood culture, drug susceptibility testing and PCR, in addition to LJ and microscopy) and quality assurance (incoming quality control check, blinding and study monitoring). Data management needs include case report forms, documentation requirements and electronic double data entry systems. Sites must be adequately qualified and personnel trained to follow protocols and to provide data integrity and regulatory compliance. All shipments of product must follow appropriate regulations and have appropriate labelling. Statistical analysis is done for results analysis, and this clinical trial written report can be part of the overall validation report.

Results of evaluation studies are judged against targets detailed in the product specification document and gold standard techniques. Studies can be undertaken with (still manually manufactured) development lots, with the zero series of production or with already commercialized IVDs. Study goals, contents and validation criteria may vary depending on the product being evaluated, such as instruments, reagents, kits, culture media and techniques (i.e. improved microscopy).

PREREQUISITES FOR AN EVALUATION STUDY

For commercial and non-commercial tests:

> Availability of products necessary to perform the test in the settings or trial sites.
> Particular attention paid to customs clearance and expiry dates of products.
> The performance targets of the diagnostic product (product specifications) to be assessed during the study have been defined.
> Target populations are clearly defined, leading to a relevant selection of the control and patient sub-groups.
> Trial sites have been inspected and certified to conduct evaluation trials for specific IVDs according to pre-defined standards and requirements.
> An agreement with trial sites has been signed.
> Study protocol has been approved by all investigators and by national scientific review boards where required, as well as ethical approval by an institutional review

board (IRB) that is recognized by the legal authorities in the countries of origin of the investigators and of the patients providing samples or being tested.

> The trial sites are trained to use the medical device for clinical studies.

> All reference methods (gold standard) required for the study are available, have been implemented and are quality-assured.

> Developers of the test under evaluation have signed an agreement for the disclosure of the results through publication or reporting. They have received a courtesy draft of the report but they cannot modify any of the conclusions.

> Sufficient funding is available to accomplish goals.

For commercial products

> The standard operating procedures (SOPs) for outgoing quality control (OQC) at the manufacturing site and incoming quality control (IQC) at the clinical trial site following shipment of the IVD have been developed and implemented.

> The investigational registration of the test device has been completed, where applicable.

For in-house, non-commercial products

> A standard operating procedure (SOP) for quality assurance upon preparation of a new lot is as important for a non-commercial test as for a commercial test. Otherwise, it will be hard to distinguish between poor test performance and poor manufacturing.

THE EVALUATION STUDY PROTOCOL

The study protocol forms part of the agreement with the investigators or principal investigator(s) and describes the following points in detail.

Study objectives and endpoints include:

> Determine sensitivity of IVD, PPV and confidence intervals in smear-positive, culture-positive pulmonary TB (PTB) patients;

> Determine specificity of IVD, NPV and confidence intervals in smear-negative, culture-negative non-TB patients with clinical symptoms of PTB at enrolment.

To determine reliability (reproducibility) of the test:

> Study design (cross-sectional or case-control, prospective or retrospective);

> Setting, specimen types tested and patient and specimen flow;

> Study size (number of patients and controls, number of specimens);

> Study inclusion and exclusion criteria (e.g. age, sex, disease severity, co-infections such as HIV, etc);

> Diagnostic gold standard that will be used as comparison;

> Blinded evaluation of test and gold standard results;

> List of trial sites, partners and investigators with description of roles and responsibilities and trial site descriptions, evidence of conflict of interest;

> Ethical considerations and informed consent form;

> Requirements for import of trial materials and other regulatory considerations;

> Data management and analysis (statistical approach);

> Study budget;

> Timelines and deliverables.

THE EVALUATION STUDY REPORT

This report summarizes results from the evaluation study and compares them with performance targets listed in the product specification document. The report should contain the following information:

> Study outline to summarize endpoints, design, setting and patient/sample characteristics (geographical origin, WHO/TDR code, HIV status if relevant);

> Sample size calculation;

> Eligibility criteria and what proportion of eligible participants completed the testing process, and summary of withdrawals and drop-outs (e.g. in a flow chart);

> Blinding to reference standard results described;

> Performance review with a detailed comparison between performance targets and evaluation trial achievements;

> Summary of significant achievements and deviations, and conclusions drawn;

> Summary of indeterminate and missing results;

> Data on reproducibility and inter-reader deviations;

> Assessment of operational characteristics of the tests;
> Assessment of strengths and limitations of the study.
> Recommendation for product improvements or for additional investigations or studies if required;

If the evaluation study report is written up as a journal publication, then it should be formatted as per the STARD statement (*http://www.stard-statement.org/*) [79]. STARD is a widely endorsed reporting format for diagnostic research publications; its use ensures that critical design elements are transparently reported.

REDUCING BIAS IN EVALUATION STUDIES

Diagnostic accuracy studies are prone to several biases [80]. A methodologically rigorous trial would avoid or minimize such biases and produce valid and precise estimates of diagnostic accuracy (i.e. sensitivity and specificity). Despite the expanding TB diagnostics pipeline and the high output of research on TB diagnostics, there is a concern that trials on the accuracy of TB diagnostics lack methodological rigour [81; 82; 83; 88]. Consequently, there is a perception that new tests that reportedly perform well in evaluation studies may turn out to be less useful in routine clinical practice. Biased results from poorly designed studies can lead to premature adoption of tests that may have little or no clinical relevance and result in adverse consequences for patients and/or health care services.

Is there evidence that TB diagnostic studies lack methodological rigour? Pai and O'Brien used several published systematic reviews and meta-analyses of TB diagnostic tests, with over 800 primary studies, to assess the quality of primary TB diagnostic studies [82]. Diagnostics included in these meta-analyses covered a wide spectrum, including smear microscopy, phage-based tests, line probe assays, serological tests and nucleic acid amplification tests. The quality elements that were most frequently reported in many of the meta-analyses were: prospective versus retrospective data collection, consecutive or random sampling versus convenience sampling, cross-sectional

design (as opposed to case-control design), blinded interpretation of test and gold-standard results, and complete verification of index test results by reference standard. Their updated analysis showed that, on average, about 65% (range 16–100%) of the trials used a prospective data collection design. However, only 33% (range 0–95%) of the trials used a consecutive or random sampling method to recruit subjects. About 72% (range 43-100%) of the trials used a cross-sectional design, and the case-control approach was used in about 33% of the studies. Any form of blinding was used in only 34% (range 0–78%) of the trials. In most studies, (range 10–100%) of the index test results were verified by a reference standard test [82].

In another recent study, Fontela and colleagues identified diagnostic accuracy studies of commercial tests for TB, malaria and HIV through a systematic search of the literature using MEDLINE and EMBASE databases (2004-2006; [83]). Original studies on commercial tests that reported sensitivity and specificity were included, and two reviewers independently extracted data on study characteristics and diagnostic accuracy. QUADAS and STARD criteria were used to evaluate the quality of study methods and quality of reporting. QUADAS is a validated tool used to evaluate methodological quality of diagnostic research studies [85]. The analysis included 45 studies of various TB diagnostics. A vast majority (> 90%) of the TB studies used an adequate reference standard test, and did not generally suffer from incorporation and partial/differential verification biases. Only about half the studies included populations with adequate spectrum composition and provided a clear description of the eligibility criteria. Less than a third of the studies provided adequate descriptions of index test and reference standard execution. Blinded interpretation of index test and reference standard results was reported in less than 20% of the studies. Reporting of withdrawals and uninterpretable results was particularly poor (<20%; [83]).

Do design flaws actually lead to biased estimates of accuracy? There is empiric evidence that methodological flaws can produce misleading estimates of diagnostic accuracy [86, 87]. A recent, large empiric study of 31 meta-analyses (with 487 primary diagnostic studies on a variety of diseases) found significantly higher estimates of diagnostic accuracy in studies with nonconsecutive inclusion of patients and retrospective data collection [87]. The estimates were highest in studies that had severe cases and healthy controls (presumably because of spectrum bias).

HOW TO IMPROVE THE QUALITY OF EVALUATION STUDIES?

High-quality diagnostic evaluation studies are critical to evaluate new tools, to develop evidence-based policies on TB diagnostics and, ultimately, to effectively control the global TB epidemic [88]. Based on the results of recent studies, it is evident that TB diagnostic trials are poorly conducted and poorly reported. Lack of methodological rigour in TB trials is a cause for concern, as it may prove to be a major hurdle in effective application of diagnostics in TB care and control. Initiatives such as STARD and DEEP have given much-needed direction to improve methodological quality and reporting of diagnostic trials. TB researchers must be encouraged to design diagnostic trials using DEEP guidelines and report them using the STARD template. To improve quality of laboratory practices, researchers must be encouraged to follow widely accepted quality systems such as GCP, GLP and GCLP. All of these resources are available at a new NDWG website: *www.tbevidence.org.*

Key points

* ✱ Evaluation studies are separate from the commercial goals and regulatory steps of an R&D project, and the tests used in them must be design-locked and ready for scaled-up manufacture.

* ✱ Lack of methodological rigour in trials may prove to be a major hurdle for effective application of diagnostics in TB care and control.

* ✱ Careful design of protocols and reports is necessary to ensure data quality. TB researchers must be encouraged to design diagnostic trials using DEEP guidelines and report them using the STARD template.

* ✱ To improve quality of laboratory practices, researchers must be encouraged to follow widely accepted quality systems such as GCP, GLP and GCLP.

* ✱ The practice of test developers to evaluate their own products is flawed and needs to be addressed by encouraging independent evaluations.

Chapter 10

Demonstration, putting the test to the test: *is it worth it?*

DEMONSTRATION, PUTTING THE TEST TO THE TEST: IS IT WORTH IT?

Chapter 10

DEMONSTRATION STUDIES CHECK EFFECTIVENESS

It is generally accepted that more rapid, accurate and easily implemented TB diagnostic tests can improve patient care and reduce disease burden by decreasing detection delay in patients seeking care and by increasing access to diagnose otherwise-missed cases. This would result in decreased TB morbidity, mortality and disease transmission.

However, the ability of a new diagnostic test to deliver these benefits may not be evident solely on the basis of analytic data on test accuracy or controlled trials of test performance. For this reason, large-scale demonstration projects are required to provide evidence that new tests that perform well in controlled settings can also have public health impact in programmatic settings.

Demonstration studies seek to assess: feasibility and barriers to implementation, scaled-up test performance, patient and/or public health impact and, where possible, cost-effectiveness of using a new technology.

These projects are typically carried out in close cooperation with health ministries and national tuberculosis control programmes. Collaboration with other technical partners working in these countries is recommended.

PROJECT DESIGN

Demonstration projects are conducted in or with tuberculosis laboratories that provide diagnostic services for TB control programmes depending on the setting they are to be used in, such as a point-of-care test.

The first phase of a demonstration project usually involves validation of the new test in comparison with the reference standard. During this phase, patient specimens submitted for routine bacteriologic examination or stored specimens (e.g. well-characterized serum specimens, culture isolates) may be used. Validation is achieved by demonstrated proficiency against predefined test

performance characteristics such as test sensitivity and specificity, or by concordance with results of the reference test.

Once acceptable performance with an appropriate sample of specimens has been achieved, the demonstration phase begins. This is when results of the new test are provided to those making clinical treatment decisions and are used to inform patient management. In most cases, demonstration protocols should be submitted for human subjects review. However, in some cases, the ethical bodies may determine that a review is not necessary – the project may be deemed public health research rather than human subjects research because the data were being gathered on implementation of an approved test for the purpose of informing local policy.

In most cases, testing by reference standard will continue during demonstration, and performance of the new test will continue to be assessed. In such cases, the protocol should detail how discrepancies in results will be handled, as in what retesting might be done and how these results will be communicated to clinicians caring for the patients. For example, in FIND's demonstration studies of a line-probe assay (LPA) for MDR-TB, phenotypic drug susceptibility testing (DST) continued.

In some cases, continued testing with the reference standard might not be done. An example is FIND's demonstration studies of LED fluorescent microscopy (FM) where Ziehl-Neelson (ZN) testing is discontinued once LED FM proficiency has been achieved.

In addition to data on test performance, qualitative information about test implementation and technician acceptability should be collected. This would include information on logistics (shipping, clearance, storage), needs for changes in laboratory infrastructure (such as provision of continuous electricity), biosafety issues and technician training (on- and off-site, time to achieve proficiency).

Documenting turn-around times from specimen collection to reporting of results for both the new test and the reference standard

is also important. Recording additional time intervals, such as specimen transport, may also prove useful in identifying ways to maximize a new diagnostic assay's impact.

Patient-related data collected during demonstration should be based on how impact is to be measured. At a minimum, this would include time from specimen collection to initiation of treatment and/ or to a change in treatment. In most cases, in addition to data such as time to sputum smear and/or culture conversion and treatment outcomes (i.e. cure, relapse, mortality), some longitudinal data would also be collected. These impact-related data should be compared to historical data recorded prior to implementation of the new test in routine clinical practice.

IMPACT-RELATED OUTCOMES

> Does the test reduce diagnostic delays?
> Does the test change clinicians' diagnostic thinking?
> Does the test improve the number of patients who get treated and cured?
> Does the test reduce the wait time before TB suspects are seen?
> Does the test reduce lab technicians' workloads?

DEMONSTRATION STUDY OF A CULTURE/DST TOOL

Less than five years ago, liquid culture and DST were not widely available in low-resource, high-burden settings. Although numerous studies performed in higher-resource countries demonstrated that these systems are faster and more sensitive than solid culture and DST methods, concerns about the complexity of the systems, infrastructure requirements, cost and potential for contamination precluded their widespread implementation.

To determine whether this technology could be successfully implemented and achieve positive public health impacts in low-resource, high-burden countries, demonstration projects were undertaken. In these projects, two complementary assays were introduced – liquid culture and liquid DST. These demonstration projects consisted

of three phases, rather than the typical two phases previously described. In the first phase, liquid culture was implemented into the laboratory and standard operating procedures optimized until contamination rates in liquid culture were consistently maintained at acceptable levels (8-10%). Once these contamination rates were achieved, the second phase involved making liquid culture results available to clinicians while the laboratories validated liquid DST against panels of well-characterized isolates provided through the WHO/IUATLD Supranational Reference Laboratory Network. Liquid culture DST results were then made available to clinicians during the third phase of the demonstration project.

Follow-up clinical and bacteriologic data were collected on patients enrolled in all three phases throughout the demonstration project. Due to the length of time required to decrease liquid culture contamination rates (>6 months), patients enrolled in phase one of the project (only solid culture and DST results reported to clinicians) served as the comparison group for turn-around times, time to initiation of appropriate treatment and time to culture conversion.

DEMONSTRATION STUDY OF MDR-TB SCREENING

Line probe assays for the detection of rifampicin and isoniazid resistance directly from smear-positive sputum specimens have the potential to significantly decrease the time to detection – from months to days – of drug resistance and initiation of proper therapy. Although molecular technologies for the diagnosis of MDR-TB have been validated in controlled laboratory settings, data are scarce on their use in high-burden settings. Large-scale demonstration projects are under way to determine the feasibility, role and impact of LPAs' introduction in national TB programmes, where molecular techniques have not been routinely used.

LPA demonstration studies consist of validation and demonstration phases. In the validation phase, an adequate sample of smear-positive sputum specimens and culture isolates is tested with the LPA and culture-based DST in parallel, and only the phenotypic DST results are provided to

DEMONSTRATION, PUTTING THE TEST TO THE TEST: IS IT WORTH IT?

Chapter 10

clinicians. If the sensitivity and specificity of the LPA are consistent with levels reported in the literature, then the study moves to the demonstration phase. During this phase the LPA and culture-based DST are again performed in parallel. However, results of the LPA are provided to clinicians to inform treatment decisions. Culture-based DST results are likely to be available weeks after LPA results are reported, but these also are provided to clinicians. Tests yielding discrepant results are repeated and, in the event that the discrepancy is not resolved, sequencing is performed.

Prospective clinical and bacteriologic data are collected on patients for the duration of the demonstration project. In addition to assay performance characteristics over time, data are collected on the turn-around time, time to initiation of appropriate treatment and to culture conversion. These are compared with historical controls diagnosed by culture-based DST.

KEY ECONOMIC CONCEPTS

The primary goal of an economic evaluation is to *identify, measure, value and compare the costs and consequences of alternatives being considered* and to establish the relative value (in monetary terms) of an intervention or programme. Economic evaluation focuses on *opportunity cost* and *incremental change*, concepts which relate to exchanges between people and the tradeoffs these exchanges involve. In health interventions, as with other types of settings, resources are the main constraint. This means that every available/possible intervention cannot be provided or implemented in every situation or setting for all of the population that can benefit from it. Therefore, choices must be made based on various underlying factors. The opportunity cost of deciding on one intervention can be seen as the health benefits that could have been derived from an alternative intervention under consideration. Furthermore, the proper decision relies on evaluating incremental changes in costs and effectiveness (or impacts) of interventions under consideration. However, over-evaluating or over-analyzing all of these parameters can become an obstacle to making a sound decision. A balanced analysis of costs and effectiveness of interventions should clarify the advantages and disadvantages of the interventions being compared.

ECONOMIC EVALUATIONS IN DEMONSTRATION STUDIES

A wide range of economic studies can be used to examine new diagnostic systems. However, unlike other types of health interventions, economic evaluation in a demonstration study of innovative diagnostic systems must first focus on detailed costs associated with the implementation, sustainability and affordability compared to available conventional diagnostic systems. Cost analysis is one of many economic evaluations that focus on assessing the cost of providing a service, programme or intervention. It is useful for assessing programme affordability, comparing costs with the available budget, guiding budget planning and determining the costs of expanding or contracting a service, programme or intervention.

To understand how cost is associated with an intervention's effectiveness, a cost-effectiveness analysis can provide more detailed and accurate information for the decision-making process. While cost analysis does not directly include the effectiveness aspect of an intervention or programme, its main advantage lies in the relatively easy comprehension of results (often calculated as cost per specified action of specified unit). Cost analysis also offers a way to directly compare costs of interventions being considered. Modeling should be simple and kept to a minimum, and researchers should not make assumptions that could significantly influence the analysis results. Cost analysis results can be used in the decision-making process when relative effectiveness data from demonstration studies of compared interventions also are available. Because of its simple yet thorough comparative economic approach, cost analysis provides a sound foundation for demonstrating prospective diagnostic tools' cost and cost-effectiveness.

DESIGNING COST ANALYSES FOR DEMONSTRATION STUDIES

Laboratory cost assessment describes performance of the new test compared with reference standards – including changes in time to detection, in staff workload and in direct and indirect costs – in terms of a monetary value, a unit cost in terms of the number of tests performed or specimen screened. Unit costs calculated for all methods compared should reflect economic, direct and indirect costs as well as monetary/financial costs associated with each method, as analyzed from a health services perspective. Table 2 and Figure 4 depict constituents of the overall unit cost calculated for each diagnostic method in cost analysis.

Table 2. Cost data elements for cost analysis of TB diagnostics and suggested data sources

Element	Cost items	Suggested sources of data
Physical Infrastructure	Construction	Construction contractors / Government Estates and building planning office / Recent laboratory construction budget
	Maintenance contracts for all laboratory equipment requiring periodic maintenance	Laboratory financial records; laboratory or hospital accounts offices
Chemicals and reagents and consumables	All types of chemicals and reagents utilized for diagnostic methods evaluated	Laboratory financial records / manufacturer catalogue (must include all costs associated with procurement, usually at 25% of the catalogue price)
Human resources	Laboratory staff salaries	Government salary scale / Laboratory or hospital accounts office
	Laboratory staff allowances and benefits	Government salary scale / Laboratory or hospital accounts office
	Staff training off-site	Laboratory records / interview
Training and quality assurance	Orientation training for new staff	Laboratory staff records
	No need for biosafety cabinet Disposal of clinical material without further treatment	No need for biosafety cabinet
	Internal QA/QC	Cost can be evaluated as part of the general cost analysis using 'ingredients' approach. The full list of internal QA/QC procedures can be found in the general SOP
Specimen transport	Cost of a vehicle used for specimen transport - evaluated as purchased 'new'	Accounting office/auto dealer
	Average distance traveled - annual figure	List of locations referring specimens to the laboratory
	Average driver salary	Accounting office
	Quantity of fuel used	Accounting office
	Fuel Price	General market research, Accounting office
	Insurance of vehicle	Accounting office
	Other consumables used in specimen transport	Accounting office

Abbreviations: QA: Quality Assurance, QC: Quality Control, SOP: Standard Operating Procedure

DEMONSTRATION, PUTTING THE TEST TO THE TEST: IS IT WORTH IT?

Chapter 10

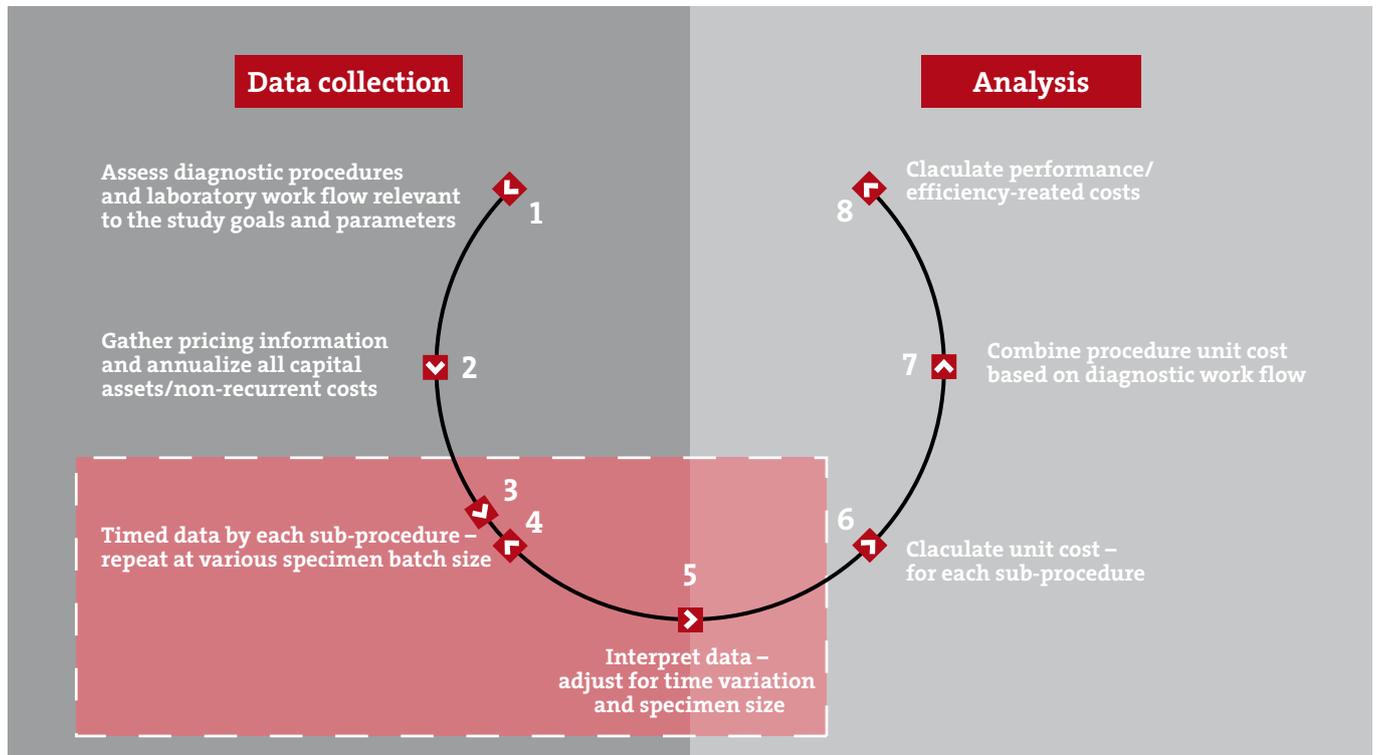## Figure 3. Components of unit cost of a laboratory procedure
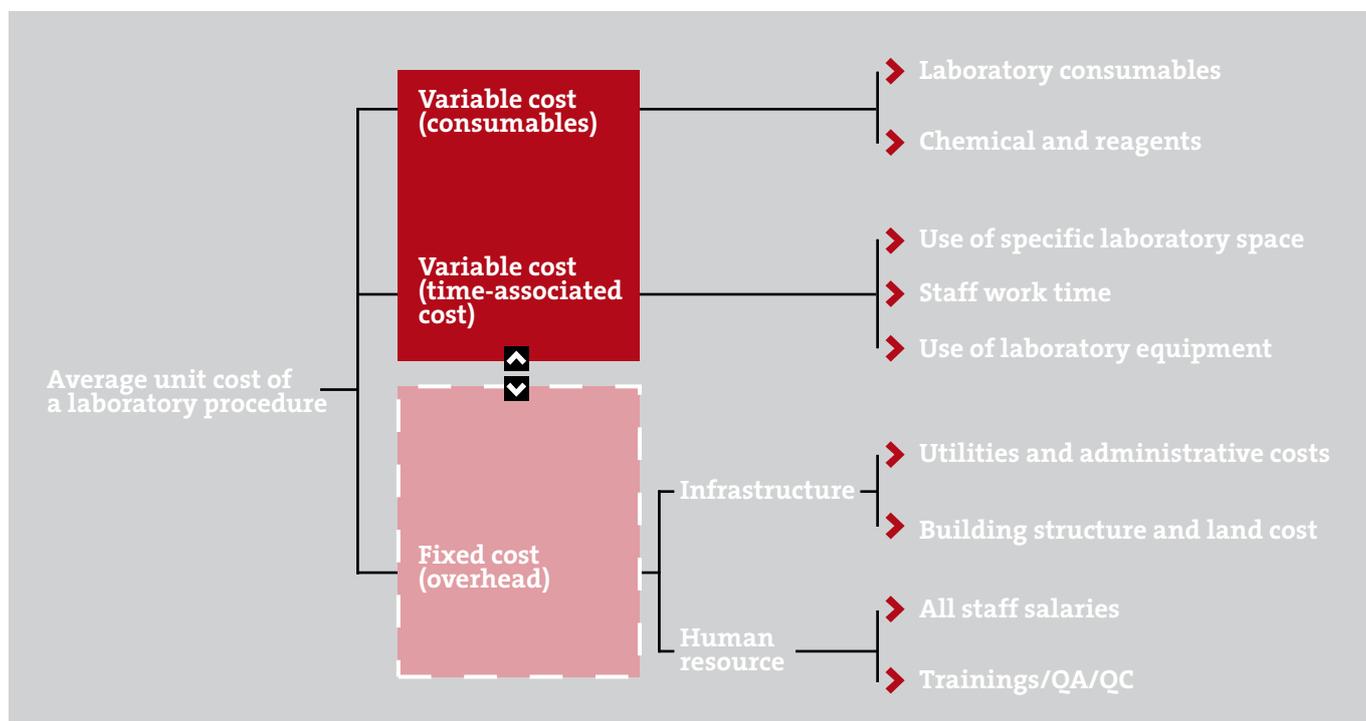
Source: Sohn H et al. (2009). TB diagnostic tests: how do we figure out their costs? *Expert Review of Anti-Infective Therapy*, 7(6) 723-733(11).

All capital and recurrent costs must be collated, from the time a specimen arrives at the laboratory until time of test results, to be analyzed for laboratory-only costs. Unit costs should be calculated using the 'ingredients' approach, multiplying the quantity of inputs used by price. All capital costs (mainly laboratory space, buildings and laboratory equipment) will need to be annualized based on their estimated expected life-years. Overhead costs are calculated by allocating overall laboratory expenditures to each test based on the number of staff and hours of staff time devoted, and based on those costs related to physical infrastructure and building space utilized by each diagnostic system.

All pricing on overhead costs, laboratory equipment, chemicals/reagents, consumables and staff salaries must be expressed and evaluated in the year the demonstration study is being evaluated; the currency can be either local or US dollars. However, for international comparison, local currencies should be converted into US dollar equivalents based on the last 12-month average local currency exchange rate for the US dollar.

In calculating costs of using capital assets (building space and equipment), data should be quantified as minute used per square meter of the space, and minutes used should be expressed based on overall capacity and the number of specimens processed in the evaluation period (point-evaluation of laboratory procedures). Laboratory consumables and chemicals should be quantified based on relevant units (units, pieces, metres, grams, millilitres). The effect of batching of tests (such as commonly occurs when processing sputum specimens for culture) should also be considered for its effects on staff time and consumable costs. Lastly, as it is difficult to quantify, measure and properly allocate staff time in between tests and lunch breaks, and similarly difficult to track wastage of consumables and unused equipment capacity, these items may not be captured in the cost analysis study. In general, the cost analysis study in demonstration studies should follow the study path described in Figure 5.

Figure 4. Path of cost analysis in demonstration studies

Source: Sohn H et al. (2009). TB diagnostic tests: how do we figure out their costs? *Expert Review of Anti-Infective Therapy*, 7(6) 723-733(11).

DEMONSTRATION, PUTTING THE TEST TO THE TEST: IS IT WORTH IT?

Chapter 10

Figure 5. Exemplary diagnostic workflow of a TB culture laboratory with molecular diagnostic capacity
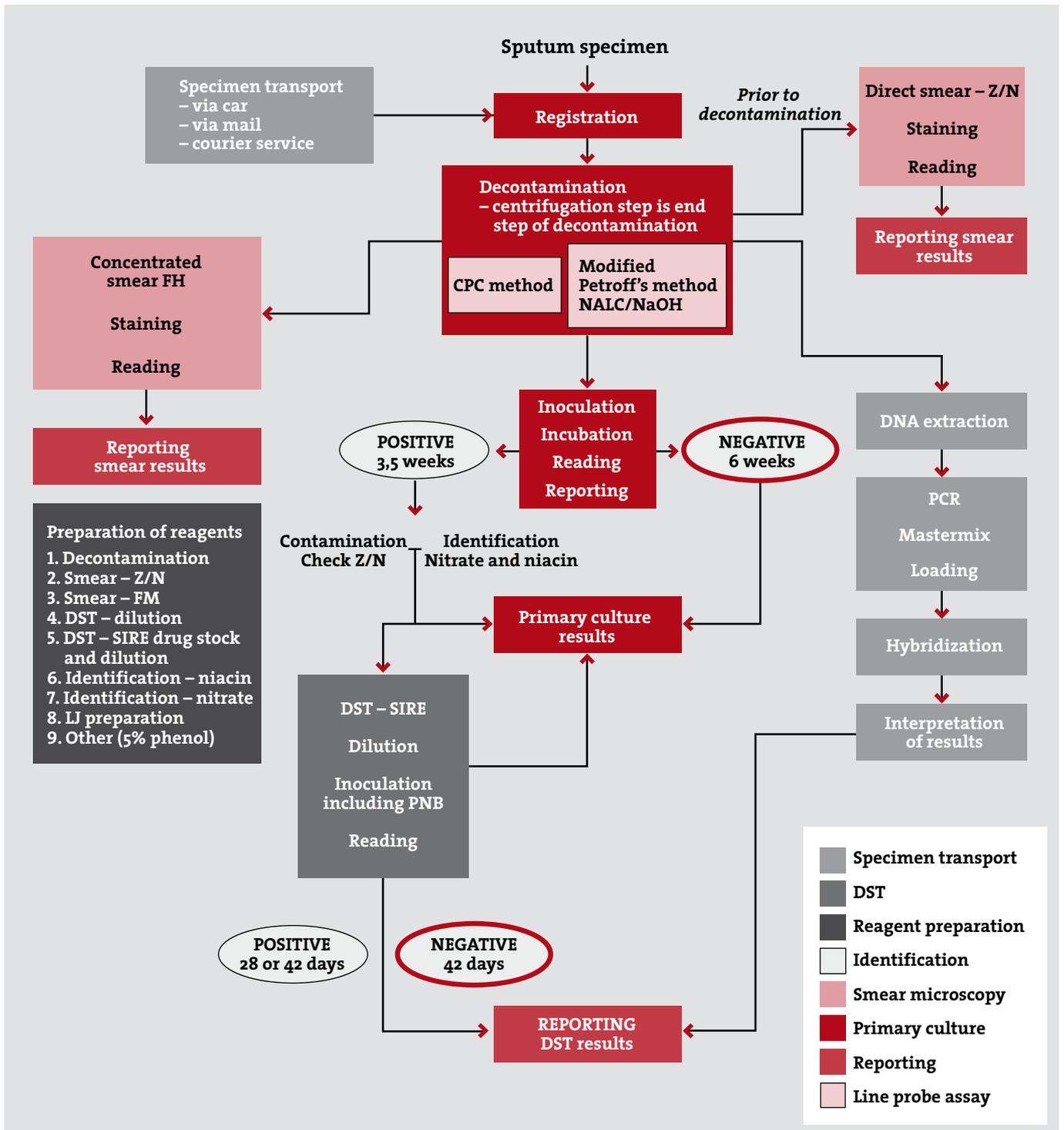
Source: Sohn H et al. (2009). TB diagnostic tests: how do we figure out their costs? *Expert Review of Anti-Infective Therapy,* 7(6) 723-733(11).

COST-EFFECTIVENESS ANALYSIS AND IMPACT EVALUATION

A cost-effectiveness analysis is appropriate when comparing alternative strategies that differ in both costs and effectiveness. Cost-effectiveness does not generally mean the lowest cost. The goal of a cost-effectiveness analysis is to identify interventions that bring the greatest effectiveness and health impact with the lowest cost per unit of output. The detailed cost analysis mentioned in a previous section provides a very good cost effectiveness perspective on laboratory services, as it evaluates time associated with each diagnostic intervention as a cost, ultimately providing costs associated with 'time-saved' in diagnostic procedure as part of the overall unit cost. However, the results would be incomplete if they did not include an evaluation of each diagnostic intervention's performance characteristics.

In a cost-effectiveness analysis, indicators can range from laboratory performance characteristics to health impacts, but these effectiveness indicators must remain the same for each type of intervention compared in the study. In interventions studying changes in diagnostic algorithms, detailed costs of all diagnostic alternatives as well as patient impact data must be evaluated simultaneously to accurately evaluate changes in costs associated with introduction of a diagnostic algorithm utilizing new diagnostic systems. In certain demonstration studies, the evaluation period may not provide adequate time to evaluate all types of health and societal impact indicators. In such cases, the study coordinator/evaluator can perform a comprehensive review of studies that evaluate comparative performance and effectiveness of an innovative diagnostic test with a 'gold-standard' or a conventional diagnostic test.

When evaluating effectiveness, one must consider the differences in the meaning and the use of efficacy and effectiveness in studies. Efficacy refers to the performance under a controlled setting, while effectiveness is described in the context of the real-world setting. Simple effectiveness indicators may include sensitivity and specificity, but effectiveness can further be evaluated in terms of increased speed of diagnosis, cases averted, reduction in transmission, etc. Effectiveness indicators will vary according to the perspective chosen for the study.

PERSPECTIVE

As noted earlier, cost analysis in demonstration studies should initially focus on health services/provider perspectives, because the target audiences often are concerned with the sustainability (i.e. laboratory operating and utilization costs) of new diagnostic tools in developing countries. As the study adopts a wider perspective, all types of costs should be reported separately where it is likely that they have an impact on the results of the analysis and are subject to sensitivity analysis. Where possible, patients should be disaggregated by socioeconomic status to help assess the equity of the test and identify whether the test is reaching people previously unable to access other diagnostic platforms. Table 3 shows perspectives in economic evaluation, types of costs considered relevant and examples.

DEMONSTRATION, PUTTING THE TEST TO THE TEST: IS IT WORTH IT?

Chapter 10

Table 3. Perspectives of economic evaluation in health interventions

Societal perspective	Types of cost	Examples
Provider perspectives	Direct costs associated with laboratory and other types of diagnostic services	Detailed costs of diagnostic services, laboratory staff costs, consumables costs and overhead costs (details presented in cost analysis section)
	Direct costs associated with medical services	Detailed costs of hospital admissions, treatment etc.
Patient perspectives	Direct costs incurred by patient	All types of out-of-pocket payments for drugs and other types of treatment following diagnosis, cost of travel, income transfer payments
	Time costs to patients and their families	Patient's time spent for travel and receiving treatment, patient and family's lost time/income from work (salary)
	Productivity	Lost productivity due to illness resulting in reduced working capacity, costs to employer to hire and train replacement worker, etc.

QALYS AND DALYS

When evaluating the health impacts from new diagnostic systems in a societal perspective, calculation of quality-adjusted life years (QALYs) and disability-adjusted life years (DALYs) can provide a generic health outcome measurement that captures the impact of an intervention on mortality and morbidity. These measurements promote the cost-effectiveness of health interventions targeting completely different types of health problems (e.g. heart surgery vs. cancer treatment, immunization vs. community health programmes). While this may not be relevant used for demonstration projects where single types of disease intervention are evaluated, QALYs and DALYs provide valuable information for the decision-making process.

Key points

* The demonstration study must provide evidence that what has been observed about a test in controlled settings can be replicated in terms of medical and public health benefits in the environment where it will be used.

* Demonstrations should in particular measure performance in line with the user requirements document.

* Demonstration studies independent from test developers and sponsors must be encouraged.

* Testing will normally be done against a reference standard or existing test; this will include economic evaluation to identify, measure, value and compare the costs and consequences of the alternatives being considered.

Chapter **11**

Measuring impact

DEFINITION OF IMPACT ASSESSMENT

The potential impact of any diagnostic test should be seen in terms of its ability to accurately diagnose TB in all its forms and in as many patients as possible. This entails summarizing evidence not only about a test's sensitivity and specificity, but also its affordability (to the health system), accessibility (to TB suspects) and impact on important patient outcomes. Each test should be considered as part of the diagnostic strategy for TB rather than in isolation. Articulating and communicating this

overall impact succinctly but with sufficient evidence enables national health systems to make rational decisions about which new diagnostic tests to implement, and when and how to implement them. Ideally, this involves collating and synthesizing data that already have been collected throughout the development process, and ensuring that gaps in the data are identified and filled.

The kinds of information that is required and the manner in which it can be obtained has previously been outlined, as shown in Table 4.

Table 4. Information required for impact assessment

Type of information	Relevant chapter
Target populations for the test, as articulated by medical needs and epidemiological context	> **Needs assessment, basic research**
Efficiency and feasibility of the test	> **Feasibility**
Clinical efficacy of the test (sensitivity, specificity, safety and impacts on clinical decisions and patient outcomes)	> **Evaluation and demonstration**
Costs to the health system	> **Demonstration**
Costs to patients	> **Demonstration**
Human resource requirements for implementation	> **Demonstration**

For purposes of this blueprint, pulling together all this information in order to help health policy decision-making is called impact assessment. The process essentially aims to assess a technology's performance under general conditions, as opposed to the carefully controlled environment in which trials are conducted, and it includes an explicit assessment of the likely societal impact. Impact assessment is similar to the emerging concept of health technology assessment (HTA) and shares most of its features [89]. Impact assessment explicitly focuses on new diagnostics' potential impact on health care

delivery to poor people – both in the global sense and within countries. This makes impact assessment particularly suited for tests for TB, an archetypal disease of poverty.

WHO UNDERTAKES IMPACT ASSESSMENT?

An international consensus around the full scope of impact assessment is in its infancy. Some forms have been carried out for decades by national regulatory authorities (such as the US Federal Drug Administration). These

assessments typically focus on tests' technical functionality and safety, rather than on issues of access to the tests among various population groups. Although such regulatory authorities are weak or absent in many poorer developing countries, global health care spending has been rising, and most remains financed by national governments. Health care decision-makers place increasing importance on comprehensive evaluation of health technologies (including diagnostic tests) with an eye towards using public resources most effectively. In countries where health care is managed predominantly through government funding, impact assessment is being undertaken by new authorities structured differently from standard regulatory agencies [90]. Examples of these include the Institute for Quality and Efficiency in Healthcare (Germany), the National Institute for Health and Clinical Excellence (England and Wales) and the Canadian Agency for Drugs and Technologies in Health. A recent comparative review of these agencies [89] notes the importance of the multi-disciplinary nature of the research required to complete assessments, along with the current prominence of health economics. This review also urged that HTA systems improve transparency, "allowing their decisions and decision-making processes to be open for all to see," which will require wider inclusion in these agencies of health professionals, industry representatives, lay people and patient groups.

At the international level, impact assessment is implied in the work of the WHO Strategic and Technical Advisory Group for TB (STAG-TB) and related bodies.

COLLATING EVIDENCE FOR IMPACT ASSESSMENT

Impact assessment can be seen as a five-layered framework (Table 2, page 64). This framework can be used to draw together all information about the actual or likely performance of a given new diagnostic test under general conditions. Where gaps in this information are identified, the impact assessment process would require the collection of additional data to fill these gaps.

The framework's layers are arranged in the order in which the data are likely to emerge during the development of new diagnostic tests, generally following steps already outlined in the blueprint. There is considerable interplay between the layers, with each layer informing how data in other layers are collected, analysed and interpreted. For example, Layer 5 is where a final decision is made regarding uptake of the new test into policy and practice. This requires a synthesis of information from all other layers of analysis as well as an assessment of the overall likelihood of the new test being quickly supplanted. For this reason, sometimes a Layer 5 scoping of new tests nearing market readiness should be the first step in the impact assessment framework, rather than the last. This can help reduce a waste of resources in completing a full impact assessment for a diagnostic test that may only have limited longevity in a particular country context.

Layer 1

Normally developers of new diagnostics universally provide data on sensitivity and specificity, which can be used in Layer 1 of the framework. However, estimations of the number of patients who might start appropriate treatment are typically calculated by extrapolating these parameters, rather than relying on evidence from field trials to provide estimates of actual numbers. All too often, diagnostic evaluations assess new tests solely in terms of their diagnostic potential, which may not always translate into appropriate clinic or public health management decisions for actual patients.

Layer 2

The quantitative and qualitative information required in Layer 2 is frequently not available unless explicitly requested. As the test's efficiency should be assessed on the time taken from a patient submitting the required sample (sputum, blood, urine) to starting treatment, both laboratory time and clinical time should be assessed. The reason for the latter is that a faster total turnaround is expected to result in more patients starting on treatment more quickly, thus reducing mortality.

Such data can be sourced from national TB programme records and from health facilities, which should record patient contact time and tracing time. Where possible, this should be verified by direct observation by a researcher and/or through patient exit interviews at time of diagnosis and starting treatment.

Important questions around key patient outcomes should be addressed by conducting randomized control trials or observation studies. If these cannot be conducted, patient outcomes can be inferred from other sources, such as the effect of positive and negative predictive values on patients being correctly put on treatment. These outcomes should ideally be weighted according to importance when making decisions on their suitability, for example using the Grading of Recommendations, Assessment, Development and Evaluation – GRADE system. [91]

Field evaluations frequently omit collection of qualitative data on test acceptability to patients and providers. High-quality information of this kind usually requires engagement of social science expertise that is difficult to source in many poorer countries.

Finally, only rarely are explicit data produced on the socio-economic profile of patients accessing new diagnostic tests in evaluation studies. There are political sensitivities around the designation of patients in terms of whether they are poor, less poor or not poor in the context of each country. Other problems in making these evaluations include inaccuracy around reported income and difficulties with collecting and analysing information on assets. However, unless such classifications are made through robust and transparent means [92] the risk remains that newer, better tests will follow the universal inverse care law and be preferentially available to the wealthier sectors of society.

Layer 3

The importance of human resource and infrastructure requirements for new diagnostics has been highlighted under evaluation and demonstration (chapters 10 & 11). In practice, such requirements are rarely reported at present.

Layer 4

The issue of economic evaluation of new diagnostic tests also has been discussed above. While some aspects of effectiveness are captured, the emphasis has been on the health system costs of implementing new diagnostic tests. While these are vital, the costs incurred by patients are also important and are given additional prominence in impact assessment. Details on how patient costs should be collected are provided on page 76.

Data collected within demonstration studies should provide policy-relevant information for local and national authorities to assess the benefit of using the proposed diagnostic test against those currently available and to assess the feasibility of scaling-up new tests. In order to inform international guidelines, however, the comparability of costs between sites must be taken into account. Health system costs will vary from location to location due to differences in staff salaries, transportation and distribution costs. Patient costs will also vary due to, for example, local income levels and average distances to health facilities. It is therefore necessary to analyse such variables in two ways:

1. Using local data to inform the TB programme of the expected costs and savings for the health system and for patients.
2. Converting local data to an internationally comparable standard.

For each of these, policy-makers need to know the costs of both setting up and sustaining the diagnostic processes. The cost data should therefore be analysed according to:

> Set-up costs (costs incurred prior to diagnostic activity for the purpose of establishing a working laboratory).
> Capacity-building costs (costs relating to training and mentoring laboratory staff).
> Recurrent costs (the ongoing costs of running the laboratory).

As noted above, the patient cost data should be disaggregated by poverty status wherever possible.

COST-EFFECTIVENESS

The costs and effectiveness analyses will be combined to give a cost per patient started on treatment, or a cost per DALY or QALY, which can be used to determine whether or not the new test offers measurable improvements over previous tests.

Patient and health system cost-effectiveness should first be calculated separately to see to which side benefits primarily accrue. These items can then be added together to establish a societal cost-effectiveness.

Layer 5

For new diagnostic tests to have relevant health impacts, they must be taken up in formal policy guidelines and implemented through re-tooled health systems. Ultimately, impact assessment seeks to facilitate uptake of new tests into formally agreed-upon policy guidelines and regulations at national and international levels. Yet as noted above, an important part of this layer of impact assessment is scoping the risk that a given new diagnostic test may be quickly supplanted by newer technology.

FURTHER CHALLENGES IN IMPACT ASSESSMENT

Gold standards

The central evidence around which impact assessment accrues is the clinical performance of a new diagnostic test. TB culture is the current gold standard required for such evidence in tests for the disease. However, this is an imperfect standard for many forms of disease, particularly in paediatric and extra-pulmonary TB. New diagnostic tests may replace the current gold standard, but explicit criteria by which future gold standards should be decided need further development. The need for such work is already demonstrated by uncertainty about what to use as the gold standard in diagnosing latent TB.

Costs of impact assessment

Because impact assessment itself has associated costs, there should be a further debate about who should bear them. In developed countries, HTA is publicly funded at varying levels. The National Institute for Clinical Excellence for England & Wales, for example, receives US$ 48.6 million annually and employs over 270 permanent staff [93].

Funding and human resources for impact assessment within developing countries will be a major challenge for years to come. It may be pragmatic for WHO to shoulder this burden through STAG, but with more explicit funding and more transparent methodology than now exists.

International comparability

Comparing costs faced in implementing new diagnostics across multiple countries is problematic. Not only do the employment costs differ significantly from country to country, but so do the costs of importing equipment and supplies. Staff data can be compared on the basis of Wellcan units, although even with these, the training that staff receives means that skill levels of laboratory technicians will often differ. Costs of supplies can be recorded at source cost, but it is important to clearly assess the costs of transportation and duty in the demonstration country.

NO TEST WITHOUT IMPACT

Impact assessment, if carried out in the comprehensive manner proposed here, should help national governments make rational decisions about when, whether and how to introduce new tests. As new tests are constantly in development, impact assessment should include a scoping of the potential of new tests coming through the pipeline. The costs to health systems of **(a)** completing a full impact assessment and **(b)** re-tooling to implement a particular test that may be superseded by a new test within a year or two may be high enough that some will wait for tests that are likely to have particular longevity for their circumstances. The grounds on which these sometimes difficult judgments are made need further work, but growing experience will build the evidence base for uptake of new diagnostic tests in policy and practice through impact assessment.

Table 5. A framework for impact assessment for new diagnostics

Layer of evaluation (methodological approach)	Examples of data required for evaluation
Layer 1: Primary outcome evaluation (quantitative)	> Sensitivity and specificity of test performance against relevant gold standard > Number of additional cases started on treatment (projected or actual
Layer 2: In-depth analysis of new diagnostic impact including: **a)** disaggregation of primary outcome data (quantitative) **b)** additional outcome evaluation (quantitative) **c)** additional outcome evaluation (qualitative)	> Number of additional cases started on treatment among the poor vs non-poor > Effect of test on clinical decision-making > Change in diagnostic delay and test performance > Change in diagnostic thinking (number of cases where the diagnosis was changed on the basis of the test result) > Improvement in patient outcomes (improvement in cure rates, proportion of patients completing therapy, reduction in treatment failures) > Acceptability of new diagnostic processes to patients and providers
Layer 3: Health systems evaluation (process evaluation – qualitative and quantitative)	> Staffing and infrastructure requirements for implementation of new diagnostic
Layer 4: Modelling potential impact of scale-up	Cost-effectiveness analyses, e.g. per additional case diagnosed and/or cured: **a)** cost savings to patients in relation to income **b)** cost savings to health providers **c)** effects of scale-up on primary outcome indicators
Layer 5: Decision on uptake into policy and practice	> Synthesis of all information in Layers 1 to 4 to help inform decision on whether test has added value over already existing tests > Extent of risk to health system of investing in a new diagnostic test that must be replaced before the benefits of the investment can be realized

COSTS TO PATIENTS

The costs that patients face between accessing the tests and starting on treatment should be calculated to estimate societal costs of the test, such as transport, accommodation and time costs. These costs are important since they impact case detection rates: the higher the cost, the higher the risk that poor patients (who are more vulnerable to TB disease) will not be able to access the test and be diagnosed.

DIRECT COSTS

On registration into a new diagnostic evaluation study, patients should be asked questions about the direct costs they have faced in accessing the facility (transport costs, fees paid at the facility and any food or accommodation costs). When they return to give their final sample (if required) they should be asked about additional costs faced in the intervening period that have resulted from them needing to stay overnight or make multiple journeys. Where they are accompanied by a 'guardian' (someone who has attended the facility with them to provide them with assistance), the costs of the guardian should be captured as well.

INDIRECT COSTS

Indirect or opportunity costs refer to the value of time lost during the care-seeking process. This should be calculated based on the time the patient (and guardian) spent accessing a diagnosis. The cost of time will be imputed to all adults: this is to reflect the fact that time has a value for everyone, regardless of whether or not they are in paid employment. Many women, for example, have household responsibilities that have to be undertaken later in the day or by someone else (hence a cost to them) while they are seeking care.

This time will then be multiplied by their estimated daily income. Rather than relying on reported income, which has proved to be unreliable in a number of settings, a figure for average income will be used, based on data from national or local studies such as integrated household surveys.

EQUITY

The equity of tests should also be assessed where possible. Poor patients are less likely to access a diagnosis since the costs of doing so, in relation to their income, are too high. It is therefore necessary to disaggregate patient cost data by poverty status and gender. Patient poverty status should ideally be assessed on the basis of asset ownership (through regression analysis of national data sets assessing poverty lines), although some proxy measures may also be used. The income attributed to them will be based on income levels reported in national or local studies for poor and non-poor populations.

Key points

* Impact assessment of a technology's performance under general conditions – as opposed to its results in a trial's controlled environment – is essential to enable national health systems make rational decisions on whether to implement a particular diagnostic test.

* Proper impact assessment should summarize not just evidence related to a test's sensitivity and specificity, but also its affordability to the health system and accessibility to patients.

* In the case of TB, the potential impact of new diagnostics on poor people is critically important.

* Developers of diagnostic tests should be mindful that health care decision-makers in most countries are placing increasing importance on using public resources in the most efficient and effective way possible.

* Decision-makers are also asking that impact assessment become more transparent by including health professionals, industry representatives, lay people and patient groups in the process.

Chapter 12

Access:
the final test of success

While the global target for case detection is to notify at least 70% of new smear-positive cases, the humanitarian aim for tuberculosis control is to rapidly and correctly identify *all* tuberculosis cases. With 61% of new smear-positive cases being detected in DOTS programmes in 2006, even the 70% case detection target has been difficult to attain. The multiple challenges to detecting all TB cases have been discussed in previous chapters and range from affordable technology to health system delivery constraints and patient socio-economic barriers. Underlying each of these challenges to our ability to detect all cases are limitations in access to appropriate technologies by patients, health providers, health systems and disease-endemic country governments. Recent modelling comparing the impact of hypothetical TB diagnostics with varying performance characteristics on TB deaths averted concluded that "larger gains (in numbers of lives saved) require either increased access or much better test performance" [94]. Important gains were achieved by taking a test that performed only as well as smear microscopy and making it rapid and widely available to all who needed it; i.e. increasing access such as through a point-of-care technology. In the model, this averted around 400,000 deaths due to TB annually. The modelling showed that the gains from new tests increased roughly in proportion with access to testing. Obviously, tests that perform better than smear microscopy are becoming available and diagnostics that can routinely address challenges such as extrapulmonary TB, TB in HIV-infected individuals, drug-resistant TB and TB in children are much needed. For all of the new technologies and diagnostic approaches that will emerge in coming years, the key to maximizing their public health impact is ensuring their ability to be accessed.

For the purposes of this Blueprint, it is important to consider the dimensions of access that can and should be considered as an integral part of evaluating new technologies. Exploring these from both patient and health system perspectives may be useful. Optimizing access by patients to a diagnostic technology is obviously the goal, with all other efforts around introducing or implementing a test supporting this end point. Therefore, it makes sense for all tests to be evaluated for some aspects of patient access as these will ultimately impact government decisions about the uptake or use of a test, and about how to best introduce and use a given technology. From the patient perspective, access to a diagnostic may be defined in geographical, financial and social terms. Increasing geographic access to a diagnostic technology means getting tests to patients. New tests that can be delivered at the point of care will certainly increase access. Short of this, there are many other ways to increase access. Some of these are bringing the diagnostic to where patients seek services (pharmacies, traditional healers, formal health care settings); considering the application of the diagnostic such that it limits the time or financial burden on patients, e.g. same-day sputum collection; and planning for and enabling specimen transport to the diagnostic, rather than moving patients to the diagnostic. During demonstration trials, an evaluation of alternative applications or delivery approaches for the diagnostic that facilitate patient access may be conducted. The proposed delivery approaches to maximize patient access should be part of the information on a new technology that is provided to countries for advance consideration of how this may fit within the current health care delivery system and where modifications will be needed, notably considering the links between clinical care and the laboratory network. In financial terms, demonstrations may consider approaches that will facilitate patient access by reducing the costs to acquiring a diagnosis, either in terms of time away from work/home or in direct expenditures for travel or the test itself. The costs to the patient must be documented such that policy decisions about appropriate public subsidies to promote access, particularly to the poor, through the health system or otherwise, can be considered as part of the decision-making process by governments. Finally, from a social perspective, the acceptability of the new diagnostic technology may influence the level of uptake by patients and communities. Efforts to identify patient concerns or misperceptions about the test may enable modifications to the test itself or highlight

issues for advocacy or patient education that need to be taken up to enhance access.

From the health system perspective, several tiers of access must be considered, namely **a)** supportive government policy to enable access to a technology by those in a country, **b)** appropriate infrastructure and delivery systems to make a technology accessible to providers, **c)** provider willingness and capacity to utilize the test, and **d)** sufficient supplies available where needed. The Retooling Task Force of the Stop TB Partnership described in *New technologies for tuberculosis control: a framework for their adoption, introduction and implementation* [95] a sequence of actions that need to be taken by national TB programmes to reduce the delay between availability of new technologies globally and their use in disease-endemic settings. The framework acknowledges the parallel and often simultaneous processes of policy development, stakeholder engagement and implementation preparedness that must occur at global and country levels to support access to new technologies. Importantly, the document highlights the need to reduce delays in accessing new technologies by considering the appropriateness of each new tool for a given country and planning for its use long before it is endorsed by WHO or other international technical agencies or becomes available for widespread distribution.

POLICY MUST ENABLE ACCESS TO TECHNOLOGY

Within disease-endemic countries, the most fundamental building block for access to a new technology is the decision by the government to introduce the diagnostic. Such a decision may be reflected in a change in the national TB control and/or laboratory guidelines to incorporate the use of the tool and the registration of the diagnostic technology / product for use in the country. This scientific blueprint will promote the availability of comparable and timely information from which decisions about the appropriateness of a diagnostic for global or country-specific use can be evaluated. The information needed by a country to adopt a new diagnostic will likely be more robust than scientific data alone, and may include details about both the required inputs and potential gains of the test. Sufficient information should be made available about the health system and human resource requirements to implement a new diagnostic, the costs involved to introduce a new technology and the running costs to use it, and about the test's availability and technical support.

The decision to change or introduce a new technology has enormous implications for a health system. Potential benefits that make it worth the change and which need to be documented may include efficiency gains, such as reduced delays to accurate diagnosis. Another benefit is the ability to provide better patient care. An example of this would be the ability to diagnose all forms of TB in HIV-infected individuals or to rapidly screen for drug resistance, or with a test that responds to other current operational challenges such as not being able to easily detect TB in children.

Once a country has decided to introduce a new diagnostic, it must be registered for use in the country.

Submission for WHO endorsement

A critical outcome of the demonstration phase is WHO endorsement (though WHO endorsement is not only based on demonstration studies) of the use of the test in the public health sector in low-income countries, and development of policy guidance on the test's appropriate use.

In 2007 the WHO Stop TB Department, through its consideration of and recommendations on TB liquid culture and drug susceptibility testing systems, established a procedure for considering new TB control technologies for adoption. In most cases, an outside entity would request that WHO review available data on a new tool. If WHO considers that these data are sufficient for review, it will assemble an expert advisory committee to review the data, make a recommendation to WHO on whether to use the tool and possibly develop draft guidance how the tool should be used. The data to be reviewed include published and unpublished performance data, optimally aided by systematic reviews and, importantly, data from demonstration projects.

If the committee believes there is sufficient generalizable data available to support a recommendation on its use, WHO will prepare a policy guidance document for presentation to the Strategic and Technical Advisory Group for TB (STAG TB) that meets each June.

Appropriate infrastructure and delivery systems are essential for implementation

Once national policy has been established to enable access to a diagnostic, the item must become accessible to and within the health system. Given current case detection rates, it must be assumed that challenges remain in accessing technologies that are delivered through existing laboratory networks and linked to existing formal health delivery centres. As it is likely that lab-based diagnostics will continue to be the cornerstone of many aspects of TB control for years to come, improving access to existing and new technologies will require some changes. An obvious place to start is to strengthen the capacity of laboratory networks in the public sector. Of the 22 highest-TB-burden countries, 7 have the recommended culture capacity per population and 10 have functional external quality assurance systems. Widespread acknowledgement of the need to enhance laboratory capacity has recently emerged, and efforts such as the Global Laboratory Initiative (GLI) promise to bring needed standards and resources to bear on this issue. Beyond the public laboratory network, private and often sophisticated reference laboratories exist in many countries and may be underutilized, particularly given the private sector's role in treating TB cases. Ensuring the acceptance and appropriate use of new diagnostics and including them in a comprehensive national laboratory network may enhance access by increasing access points for patients and referral points for other labs.

Patients do not tend to walk into a laboratory, self-directed, to request a test for TB. They are reliant on health providers to refer them for a diagnosis. This means that even a fully functioning laboratory network may be insufficient to make access universal. The role of national TB programmes (NTPs) in ensuring effective provider practice with respect to suspecting and testing for TB cannot be overestimated. New diagnostics may provide an incentive to providers to refer patients for a TB diagnosis, particularly where they promise more efficient and accurate results, and where they can enable better care practices. Product information that is targeted to providers may stimulate increased referrals for diagnosis, i.e. increased access. Several studies have suggested that many TB patients seek care with various types of health providers, such as traditional healers or pharmacies, before interacting with a formal provider who may be considered part of the DOTS network. If referrals for diagnosis only occur in formal health centres or hospitals, access to diagnostics may be delayed or unobtainable. Efforts to ensure that all providers seeing TB suspects can appropriately refer patients to a new diagnostic will similarly increase access.

The experiences of public-private mix DOTS (PPM-DOTS) would suggest that consideration for how private providers and the private sector gain access to and make accessible to patients the new technologies must be carefully negotiated to stimulate appropriate use of the diagnostics, encourage uniformity in the diagnostic strategies of the public and private sectors (or at least compliance with national diagnostic policies) and protect patients from barriers to accessing the new diagnostic, such as cost.

Within the public health network, it will become the joint responsibility of NTPs and national laboratory services (in the case of lab-based diagnostics), to ensure access to each new diagnostic by planning for its introduction, financing, implementation and monitoring. Among the information provided to countries by product developers and technical agencies should be guidance on how each test can be utilized in concert with other approved technologies, given differing health system realities and epidemiological priorities.

With each new tool, the NTP will have to revise its diagnostic algorithm, determining how the tool can best be used to overcome a specific operational challenge (e.g. to reduce delay to diagnosis, to rapidly screen for drug resistance) and how it will be used in concert with other existing technologies (e.g. the availability of line-probe assays to screen

for drug resistance may allow a country to rapidly test all re-treatment cases, while reducing the number of cultures that need to be done in a reference lab). Based on this, the NTP can define the TB suspects or patients who need to access the test and assess where these patients are being seen within the health system. To make the test accessible to the greatest proportion of patients, the NTP will need to carefully map how it will unite the patients with the diagnostic. This is done through a combination of capitalizing on the existing health centres and laboratory infrastructure, planning for enhancements to the lab network to increase its reach and capacity, considering how to draw in non-formal health facilities and providers, and planning for the transport of specimen samples and test results from and to patients. This roadmap to patient access will give the NTP several key aspects needed for planning for implementation and the financing required. Obviously, the change to the diagnostic algorithm will need to be reflected in the national TB control manual, diagnostic guidelines, standards of practice (SOPs), laboratory and clinical registers, training materials and monitoring systems. The introduction of new TB diagnostics should be an integral, budgeted part of all 5-year plans for TB and any medium-term plans for the laboratory network. For the tool to be accessible through/to all health providers and labs, considerable effort will go towards developing the on-the-job tools, manuals and training to support its appropriate use.

With national policy and an overall roadmap for how a new diagnostic will be introduced, financing needs can be calculated to cover the technology and its running expenses as well as operational costs to support training, infrastructure upgrades, dissemination of manuals/SOPs, increased supervision, specimen transport, etc. Without adequate financing for all aspects of the introduction and implementation of a new diagnostic, access will be limited. Securing financing through international donors may take considerable time. To reduce delays in the introduction of a new diagnostic, an NTP may need to predict the financing needs for new tools by forecasting the availability of certain technologies a year or more in advance, based on the information that is available through this scientific Blueprint.

With this forecast they may choose to apply for funding in accordance with their comprehensive roadmaps for strengthened diagnosis.

Providers must be willing and able to utilize the test

Providers, formal and informal, across the health system must understand the NTP's diagnostic algorithm and their roles in it. When a new technology is introduced, it must be understood and accepted, or access will be limited due to underutilization of the diagnostic. For providers, especially the laboratory technicians who will directly conduct a test, the need to be trained on technical aspects is evident. These providers may also benefit from understanding their role in patient care: how specimens are flowing to them and how results will return to patients. For providers who do not directly conduct the test, such as clinicians or peripheral laboratory technicians in the case of a reference lab-based test, the need for awareness and acceptance of the diagnostic is less obvious.

However, it is clinicians who may have the most important impact on access to the technology. They are the front line with patients, and are responsible for making appropriate referrals and for moving specimens and results within the laboratory network. If they are not aware of a technology's applications and how to access it, they may miss opportunities to draw on this strength of their health system. For example, if an overwhelmed peripheral lab is unaware of the high-volume capacity of the neighbouring laboratory with LED microscopes, it may fail to conduct all sputum smears in a timely manner rather than link with the LED site for alternative slide preparation/reading. Similarly, a provider not aware of the existence of line-probe assays may continue a re-treatment regimen without benefiting from a rapid screening for drug resistance. Even providers who do not directly care for TB patients, such as pharmacists, can promote access to TB diagnostics if they can communicate to patients the array of diagnostic possibilities and refer them appropriately.

Key points

✱ In disease-endemic countries, the most fundamental issue of access to a new diagnostic technology is the government's decision to introduce it.

✱ Since governments and health agencies seek to maximize public health impact, a critical factor in their evaluations of diagnostic tests will be the ease with which they can be delivered and used in impoverished areas.

✱ One key means to address access issues is to strengthen the capacity of public-sector laboratory networks.

✱ To avoid underutilization of a promising test, it is important that providers across the health care network understand it, accept it, become trained in administering it and promote its use.

Chapter 13

Barriers and challenges

While there is considerable diagnostic work going on in smaller biotechnology companies and academic research groups, the private sector diagnostic development community generally views this market as too difficult to enter to ensure a return on investment, thus presenting a significant barrier to innovation. Too often there is no real cohesion between the various actors in the development of new diagnostics, with many initiatives working in isolation and considerable duplication of work. Overcoming these and other challenges, as described in this section, is critical to bringing better TB diagnostics to market.

AN ABSENCE OF OVERSIGHT

A recent laboratory-based evaluation of 19 commercially available rapid diagnostic tests for tuberculosis demonstrated universal poor performance: none of the assays performed well enough to replace smear microscopy. This finding in many instances contrasted with the claims of excellent performance that featured in the tests' package inserts [96]. A recent analysis of systematic reviews showed that trials of TB diagnostics often lacked methodological rigor and were poorly reported [97]. Poorly designed and executed trials could result in inflated claims of performance. However, the absence of an active regulatory body with responsibility for diagnostics allows underperforming or inadequate tests to be marketed, which in turn can lead to either inappropriate treatment or delayed diagnosis. Clearly, the current practice for test developers to evaluate their own products is flawed and needs to be addressed by encouraging additional independent assessment and rigorous comparison to competing or alternative technologies. Regulatory control of diagnostics is also needed to sustain advances in diagnostics technology. Manufacturers of high-performing, high-quality products may not be able to compete with manufacturers of cheap, poorly performing, low-quality products unless quality and performance are valued and demanded. Previous chapters have outlined the process of TB diagnostics development, including key steps to ensure that a diagnostic tool or technique is

adequately validated, accepted and used. The goal should be to establish a systematic, standardized and independent validation mechanism.

PUTTING THE REAL NEEDS UP FRONT

Priority-setting is critical to ensure that TB test development is driven by the urgency of clinical need and not solely by the needs and interests of donors and/or individual test developers. Given that so many varying research agendas co-exist, the most urgent medical needs don't always attract the greatest attention or resources (as is the case with detection of active versus latent tuberculosis). Establishing specification requirements for future tests requires robust processes based on comprehensive needs assessments and a firm understanding of technological feasibility. Many of the barriers to implementing new tests are associated with issues such as acceptance by countries or donors, while the capacity of clinics or laboratories is directly linked to performance, ease-of-use and cost of the new tool. This situation reinforces the importance of establishing a rigorous and transparent process to determine needs and priorities – one that expands the range of actors involved and broadens the base of institutions working on tuberculosis across the public, private, academic and corporate sectors.

KNOWLEDGE SHARING

Access to knowledge has been identified as a particularly problematic area. The need, under the current structure, of academic and commercial partners to own intellectual property in order to leverage funding encourages secrecy and leads to duplication of work. Too often there is no real cohesion between the various actors in the development of new diagnostics, with many initiatives working in isolation. Accordingly, there is a need to find alternatives to the current intellectual property mechanisms that will encourage greater collaboration, avoid unnecessary duplication and accelerate scientific progress.

Chapter 13

INADEQUATE FUNDING

Funding levels are still insufficient to adequately support TB diagnostics development. Funding for TB R&D, including diagnostics, drugs, vaccines as well as basic science and operational research, totalled US$ 429 million in 2006 – a paltry amount when compared to the estimated annual funding need of US$ 2 billion. Diagnostics research remains the lowest-funded research category. The US$ 31.4 million invested in diagnostics in 2006 is just 7.3% of all 2006 TB R&D [98]. Yet analysis has shown that a significant market exists for TB diagnostics and it has been estimated that approximately US$ 1 billion is spent on the purchase of TB diagnostics each year [99]. However, in low- and middle-income countries where the majority of TB cases occur, spending is significantly lower than in developed countries. It is apparent that in financial terms the main market opportunity is for the detection of latent TB, while there continues to be only a modest commercial incentive to develop a test to detect active TB despite that being the main public health priority.

STIMULATING THE MARKET

A large global market exists for TB diagnostics. As noted above, a billion US dollars are spent annually on TB diagnostics, and an even larger potential market exists for more effective and affordable tools. Between 70% and 90% of the potential available market for new diagnostic tools is concentrated in the 22 countries with the highest burden of tuberculosis [100]. However, market incentive mechanisms, which rely on high prices to fund R&D, do not result in creating advanced diagnostics in the areas of highest need, and accordingly there is a need to expand financing approaches. The effort should begin with increasing philanthropic and government funding, which usually takes the form of academic grants or product development partnerships. In addition, the diagnostics community should explore alternative funding mechanisms aimed at encouraging and attracting new R&D partners, including commercial entities. One idea that has generated interest is establishing a prize fund to stimulate development of a point-of-care TB diagnostic, as the governments of Bolivia and Barbados have recently proposed. By rewarding the test developers or their sponsors and making the resulting knowledge publicly available, the fund would ensure both innovation *and* access. Ideas such as this require significant funding commitment and raise difficult issues of governance, independence and management of conflict of interests. Discussions furthering these initiatives should take place at a global level, and it is anticipated WHO's Intergovernmental Working Group on Public Health, Innovation and Intellectual Property would assist with this process.

There are other areas that defy pure market solutions and therefore need to be addressed. For example, access to specimens, patients and field test sites is limited by a lack of capacity and the high costs involved. Mechanisms to improve access are needed, particularly for new R&D entities entering the research field. Although a clinical specimen bank such as the one managed by WHO/TDR can play an important role in test validations, such specimen banks will need to be expanded to include more samples, a greater diversity of specimens and a wider array of sample types (e.g. blood, urine, sputum, saliva). Establishing transparent procedures for accessing samples and field sites also is important, particularly when increasing numbers of test developers will require such access.

A BROADER SCIENCE BASE

The lack of validated biomarkers for active disease in adults and children has been identified as a major barrier to test development, and greater efforts are needed in this area of research. New detection platforms currently under development offer the potential to utilize different types of markers, and such work should be expanded to include a broad range of biomarkers. Metabolic research should be encouraged alongside genomic and proteomic approaches. This is an area that would benefit from increased collaboration and better sharing of existing knowledge.

SUMMING UP THE CHALLENGE

The key challenges in the fight against tuberculosis are to find ways to adapt promising new diagnostic tools for use in high-burden settings and to open a pipeline for their development, marketing, distribution and widespread use in the places where they are needed most. Overcoming these challenges and making faster and more accurate diagnosis the norm even in the most remote settings will expand effective treatment in the developing world, vastly reducing the numbers of infections worldwide and saving millions of lives.

Fortunately, a great and growing interest in the multilateral community (e.g. the Global Fund to Fight AIDS, Tuberculosis and Malaria) has led to increased resources for action against TB, including the development and distribution of new and more effective diagnostics. Monies are becoming available through national research institutes and donor programmes for these efforts. Public-private partnerships are emerging for the purpose of improving diagnostic tools and finding ways to speed them to market, and the international infrastructure for accomplishing this aim is beginning to take shape.

Key points

✳ A significant barrier to TB diagnostics innovation is that many smaller biotechnology companies and academic research groups view the market as too difficult to enter to ensure a return on investment.

✳ Market incentive mechanisms, which rely on high prices to fund R&D, do not result in creating advanced diagnostics in the areas of highest need, and so financing approaches should be diversified and expanded.

✳ For a number of reasons, cohesion is lacking between various actors in the development of new diagnostics, and thus many initiatives work in isolation.

✳ The need for academic and commercial partners to own intellectual property in order to leverage funding encourages secrecy and leads to duplication of work.

References

1. Hess W (1992). Walther and Angelia Hesse: early contributors to bacteriology. D Gröschel, trans. *ASM News, 58*:425–428.

2. Koch R (1932). The aetiology of tuberculosis. Pinner M, Pinner B, trans. *American Review of Tuberculosis, 25*:285–323.

3. Koch R (1882). The etiology of tuberculosis. In: Brock TD, Milestones in microbiology: 1546 to 1940, Brock TD, trans. 1999. Washington, American Society for Microbiology.

4. Koch R (1882). Die aetiologie der tuberculose, Berliner. *Berliner Klin Wochenschr, 19*:221.

5. Baldwin E, Petroff S, Gardner L (1927). *Tuberculosis: bacteriology, pathology and laboratory diagnosis*. Philadelphia, Lea & Febiger.

6. Middlebrook G, Cohn ML (1958). Bacteriology of tuberculosis: laboratory methods. *American Journal of Public Health, 48*:844–853.

7. World Health Organization (1948). Report of the expert committee on tuberculosis, 30 July–2 August 1947 *First Session*. In: *Bulletin of the World Health Organization, 1*(2):205–211.

8. Chadwick HD, Pope AS (1942). *The modern attack on tuberculosis*. New York, Commonwealth Fund.

9. MacNabb AL (1941). The tubercle bacillus. In: *Diagnostic procedures and reagents-techniques for the laboratory diagnosis and control of communicable diseases*. New York, American Public Health Association, 281–292.

10. Abbott JN (1950). Technics and circumstances in the bacteriologic diagnosis of tuberculosis. *American Journal of Public Health, 40*:833–837.

11. Dufault P (1957). *The diagnosis and treatment of pulmonary tuberculosis*. Philadelphia, Lea & Febiger.

12. Communicable Disease Center (1960). Tuberculosis: laboratory methods in diagnosis. *Public Health Service Publications No. 770*.

13. Middlebrook G (1965). The mycobacteria. In: Dubos RJ, Hirsch JG, *Bacterial and mycotic infections of man*, 4th ed. Philadelphia, JB Lippincott, 490–529.

14. Runyon EH et al. (1970). Mycobacterium. In: Blair JE, Lennette EH, Truant JP, eds. *Manual of clinical microbiology*. Bethesda, American Society for Microbiology.

15. Mitchison DA et al. (1972). A selective oleic acid albumin agar medium for tubercle bacilli. *Journal of Medical Microbiology, 5*:165–75.

16. Middlebrook G, Reggiardo Z, Tigertt WD (1977). Automatable radiometric detection of growth of *Mycobacterium tuberculosis* in selective media. *American Review of Respiratory Disease, 115*:1066–9.

17. Tenover FC et al. (1993). The resurgence of tuberculosis: is your laboratory ready? *Journal of Clinical Microbiology, 31*:767–70.

18. Centers for Disease Control and Prevention (2009). Updated guidelines for the nucleic acid amplification (NAA) tests in the diagnosis of tuberculosis. *Morbidity and Mortality Weekly Report, 58*:7–10.

19. Piersimoni C, Scarparo C (2003). Relevance of commercial amplification methods for direct detection of mycobacterium tuberculosis complex in clinical samples. *Journal of Clinical Microbiology*, 41:5355-5365.

20. World Health Organization (2007). *Global Tuberculosis Control Report 2007*. WHO/HTM/TB/2007.376.

21. Maher D (1999). Smear-positive pulmonary tuberculosis: good clinical management is good public health. *African Health, 21*:6–9.

22. Hopewell P et al. (2006). International standards for tuberculosis care. *Lancet Infectious Diseases, 6*:710–725.

23. Rieder HL (1999). *Epidemiologic basis of tuberculosis control*, 1st ed. Paris, International Union Against Tuberculosis and Lung Disease.

24. Harries AD, Maher D (2004). *TB/HIV: a clinical manual*, 2nd ed. (WHO/ HTM/TB/2004.329). Geneva, World Health Organization.

25. World Health Organization (2003). Treatment of tuberculosis, *Guidelines for national programmes*, 3rd ed. Geneva, WHO (WHO/CDS/ TB/2003.313)

26. World Health Organization (2004). *Toman's tuberculosis: case detection, treatment, and monitoring*, 2nd ed. Geneva, World Health Organization, 1–332.

27. Stop TB Partnership, World Health Organization (2006). *Global Plan to Stop TB 2006-2015*. Geneva, World Health Organization.

28. *Guidance for national tuberculosis programmes on the management of tuberculosis in children*. (2006). Geneva, World Health Organization.

29. Mukadi YD, Maher D, Harries A (2001). Tuberculosis case fatality rates in high HIV prevalence populations in sub-Saharan Africa. *AIDS, 15*:143–152.

30. Witkamp B et al. (2008). Self-detecting gate tunable nanotube paddle resonators. *Applied Physical Letters, 93*:111909–1-3.

31. Zheng G (2005). Multiplexed electrical detection of cancer markers with nanowire sensor arrays. *Nature Biotechnology, 23*:1294–1301.

32. Staii C, Johnson AT (2005). DNA-decorated carbon nanotubes for chemical sensing. *Nano Letters, 5*:1774–1778.

33. White J et al. (2008). Solid-state dye-labeled DNA detects volatile compounds in the vapor phase. *PLoS Biology, 6*:30–36.

34. Anthony R et al. (2006). Light emitting diodes for auramine O fluorescence microscopic screening of *Mycobacterium tuberculosis. International Journal of Tuberculosis and Lung Disease, 10*:1060–1062.

35. Torrea G et al. (2008). Evaluation of the FluoreslenS and fluorescence microscopy blinded rechecking trial, Nairobi, Kenya. *International Journal of Tuberculosis and Lung Disease, 10*:1060–1062.

36. Van Deun A (2008). Performance and stability of the FluoLED easy module for tuberculosis fluorescence microscopy. *International Journal of Tuberculosis and Lung Disease, 12*:1009–1014.

37. Cui X et al. (2008). Lensless high-resolution on-chip optofluidic microscopes for *Caenorhabditis elegans* and cell imaging. *Proceedings of the National Academy of Sciences USA*, 105:10670–10675.

38. Buijtels P et al. (2008). Rapid identification of mycobacteria by raman spectroscopy. *Journal of Clinical Microbiology, 46*:961–965.

39. Rebuffo-Scheer C et al. (2007). Rapid species and strain differentiation of non-tubercoulous mycobacteria by fourier-transform infrared microspectroscopy. *Journal of Microbiological Methods, 68*:282–290.

40. Zhao X et al. (2004). A rapid bioassay for single bacterial cell quantitation using bioconjugated nanoparticles. *Proceedings of the National Academy of Sciences USA, 101*:15027–15032.

41. Edgar R et al. (2006). High-sensitivity bacterial detection using biotintagged phage and quantum-dot nanocomplexes. *Proceedings of the National Academy of Sciences USA*, 103:4841–4845.

42. Ymeti A et al. (2007). A single platform image cytometer for resourcepoor settings to monitor disease progression in HIV infection. *International Society for Analytical Cytology, 71A*:132–142.

43. Ymeti A et al. (2007). Fast, ultrasensitive virus detection using a young interferometer sensor. *Nanoletters*. 7,2:394–397.

44. Patiño S et al. (2008). Auto fluorescence of mycobacteria as a tool for the rapid detection of *Mycobacterium tuberculosis. Journal of Clinical Microbiology*, 46:3296–3302.

45. Bruno E et al. (2008). Can the electronic nose diagnose chronic rhinosinusitis? A new experimental study. *European Archives of Otorhinolaryngology, 265*:425–428.

46. Popplewell J et al. (2005). Quantification of the effects of melittin on liposome structure. *Biochemical Society Transactions, 3*:931–933.

47. Ren J et al. (2008). A new MSPQC for rapid growth and detection of *Mycobacterium tuberculosis. Biosensors and Bioelectronics, 24*:403–409.

48. Bergval IL et al. (2008). Development of multoplex assay for rapid characterization of *Mycobacterium tuberculosis. Journal of Clinical Microbiology, 46*:689–699.

49. Piepenburg O et al. (2006). DNA detection using recombination proteins. *PLOS Biology*, 4,7:e204. doi:10.1371/journal.pbio.0040204

50. Schasfoort B, Tudos A (2008). *Handbook of surface plasmon resonance*. Cambridge, Royal Society of Chemistry.

51. Stern E et al. (2007). Label-free immunodetection with CMOS-compatible semiconducting nanowires. *Nature, 445*:519–522.

52. Carlen ET, van den Berg A (2007). Nanowire electrochemical sensors: can we live without labels? *Lab on a Chip, 7*:19–23.

53. Yager P et al. (2006). Microfluidic diagnostic technologies for global public health. *Nature, 442*:412.

54. Fu E et al. (2008) Chapter 10. In: Schasfoort BM, Tudos AJ, eds. *Handbook of surface plasmon resonance*. Cambridge, Royal Society of Chemistry.

55. Heintz A, Fink M, Mitchell B (2007). Mechanochemical synthesis of blue luminescent alkyl/alkenyl-passivated silicon nanoparticles. *Advanced Materials, 19*:3984–3988.

References

56. Huisken F et al. (2002). Light-emitting silicon nanocrystals from laser pyrolysis. *Advanced Materials*, 14:1861–1865.

57. Liu H, Ye T, Mao C (2007). Flurescent carbon nanoparticles derived from candle soot. Angewandte Chemie *International Edition*, 46:6473–6475

58. Bottini M, Mustelin T (2007). Nanosynthesis by candlelight. *Nature Nanotechnology*, 2:599–600.

59. Resch-Genger U et al. (2008). Quantum dots versus dyes as fluorescent labels. *Nature Methods*, 5:763–775.

60. Cheng MM-C et al. (2006). Current opinion in chemical biology. *Nano Science and Technology Institute*, 10:11–19.

61. Kerman K, Kobayashi M, Tamiya E (2004). Recent trends in electrochemical DNA biosensor technology. *Measurement Science and Technology*, 15:R1–R11.

62. Koehne J et al. (2004). Miniaturized multiplex label-free electronic chip for rapid nucleic acid analysis based on carbon nanotube nanoelectrode arrays. *Clinical Chemistry*, 50:1886–1893.

63. Kinsella JM, Ivanisevic A (2007). Taking charge of biomolecules. *Nature Nanotechnology*, 2:596–597.

64. Petousis I et al. (2007). Transient behaviour of magnetic micro-bead chains rotating in a fluid by external fields. *Lab Chip*, 7:1746–1751.

65. Derks R et al. (2008). Self-organized twinning of actuated particles for microfluidic pumping. *Applied Physics Letters*, 92:024104/1–024104/3. doi:10.1063/1.2834851

66. Minnikin D, Hutchinson I, Caldicott A (1980). Thin-layer chromatography of methanolysates of mycolic acid-contaning bacteria. *Journal of Chromatography A*, 188:221–233.

67. Stopforth A et al. (2005). A rapid method of diagnosing pulmonary tuberculosis using stir bar sorptive extraction–thermal desorption–gas chromatography–mass spectrometry. *Journal of Chromatography A*, 1071:135–139.

68. Erwin Kaal, Arend H.J. Kolk Sjoukje Kuijper Hans-Gerd Janssen. 2009 A fast method for the identification of Mycobacterium tuberculosis in sputum and cultures based on thermally assisted hydrolysis and methylation followed by gas chromatography–mass spectrometry. J. Chromatogr. 1216: 6319-6325

69. Brooks J et al. (1998). Development of a quantitative chemical ionization gas chromatography-mass spectrometry method to detect tuberculostearic acid in body fluids. *Journal of Chromatography B: Biomedical Sciences and Applications*, 712:1–10.

70. Daikos G et al. (2004). Detection of tuberculostearic acid in serum and other biological fluids from patients with tuberculosis by electron capture-gas chromatography and chemical ionisation-mass spectrometry. *International Journal of Tuberculosis and Lung Disease*, 8:1027–1031.

71. Fend R et al. (2006). Prospects for clinical application of electronic-nose technology to early detection of *Mycobacterium tuberculosis* in culture and sputum. *Journal of Clinical Microbiology*, 44:2039–2045.

72. Syhre M, Chambers ST (2008). The scent of *Mycobacterium tuberculosis*. *Tuberculosis*, 88:317–323.

73. Phillips M et al. (2007). Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis*, 87:44–52.

74. Phillips M (1997). Method for the collection and assay of volatile organic compounds in breath. *Analytical Biochemistry*, 247:272.

75. Arend S et al. (2008). Double-blind randomized phase I study comparing rdESAT-6 to tuberculin as skin test reagent in the diagnosis of tuberculosis infection. *Tuberculosis*, 88:249–261.

76. Agranoff D et al. (2006). Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. *Lancet*, 368:1012–1021.

77. van der Werf M et al. (2007). Standard reporting requirements for biological samples in metabolomics experiments: microbial and in vitro biology experiments. *Metabolomics*, 3:189–194.

78. Banoo S et al. (2006). Evaluation of diagnostic tests for infectious diseases: general principles. *Nature Reviews Microbiology*, 4:S21–S31.

79. Bossuyt P et al. (2003). Towards complete and accurate reporting of studies of diagnostic accuracy: the STARD initiative. *Annals of Internal Medicine*, 138:40–44.

80. Whiting P et al. (2004). Sources of variation and bias in studies of diagnostic accuracy: a systematic review. *Annals of Internal Medicine*, 140:189–202.

81. Small P, Perkins M (2000). More rigour needed in trials of new diagnostic agents for tuberculosis. *Lancet*, 356:1048–1049.

82. Pai M, O'Brien R (2006). Tuberculosis diagnostics trials: do they lack methodological rigor? *Expert Review of Molecular Diagnostics*, 6:509–514.

83. Fontela PS, et al. Quality and Reporting of Diagnostic Accuracy Studies in TB, HIV and Malaria: Evaluation Using QUADAS and STARD Standards. *PLoS ONE* 2009;4(11): e7753.

84. Walsh A, McNerney R (2004). Guidelines for establishing trials of new tests to diagnose tuberculosis in endemic countries. *International Journal of Tuberculosis and Lung Disease*, 8:609–613.

85. Whiting P et al. (2003). The development of QUADAS: a tool for the quality assessment of studies of diagnostic accuracy included in systematic reviews. *BMC Medical Research Methodology*, 3:25.

86. Lijmer JG et al. (1999). Empirical evidence of design-related bias in studies of diagnostic tests. *Journal of the American Medical Association*, 282:1061–1066.

87. Rutjes A et al. (2006). Evidence of bias and variation in diagnostic accuracy studies. *Canadian Medical Association Journal*, 174:469–476.

88. Pai M, Ramsay A, O'Brien R (2008). Evidence-based tuberculosis diagnosis. *PLoS Medicine*, 5(7):e156.

89. Healy P, Pugatch M (2009). Theory versus practice: discussing the governance of health technology assessment systems. *The Stockholm Network*.

90. The international network of agencies for health technology assessment – HTA resources. http://www.inahta.org/ (accessed 23 June 2009)

91. Schünemann AHJ, Oxman AD, Brozek J, Glasziou P, Jaeschke R, Vist GE, Williams JW, Kunz R, Craig J, Montori VM, Bossuyt P, Guyatt GH. GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. BMJ 2008 336: 1106-1110

92. Nhlema-Simwaka B et al. (2007). Developing socio-economic measures to monitor access to tuberculosis services in urban Lilongwe, Malawi. *International Journal of Tuberculosis and Lung Disease*, 11:65–71.

93. Annual Report NICE. (2007/08). National Institute for Health and clinical Excellence. http://www.nice.org.uk/

94. Keeler M et al. (2006). Reducing the global burden of tuberculosis: the contribution of improved diagnostics. *Nature*, 22 Nov. 2006;49–57. doi:10.1038/nature05446

95. Stop TB Partnership (2007). *New technologies for tuberculosis control: a framework for their adoption, introduction and implementation*. Geneva, Stop TB Partnership. www.stoptb.org/retooling/

96. TDR/FIND (2008). Laboratory-based evaluation of 19 commercially available rapid diagnostic tests for tuberculosis. UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases. http://apps.who.int/tdr/publications/tdr-research-publications/diagnostics-evaluation-2/pdf/diagnostic-evaluation-2.pdf.

97. Pai M, Ramsay A, O'Brien R (2008). Evidence-based TB diagnosis. *PLoS Medicine*, 5;(7): e156. doi:10.1371/journal.pmed.0050156 http://www.tbevidence.org/documents/whatis/Pai_PLoS_Med_2008.pdf.

98. Diagnostics for Tuberculosis: Global Demand and market Potential. WHO Geneva, 2006

99. TDR/FIND (2006). Diagnostics for tuberculosis: global demand and market potential. UNICEF/UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases, Foundation for Innovative New Diagnostics. http://apps.who.int/tdr/svc/publications/tdr-research-publications/diagnostics-tuberculosis-global-demand

100. Feur C et al. (2008). An update. In: *Tuberculosis research and development: a critical analysis of funding trends, 2005-2006*. New York, Treatment Action Group.

Further reading

In addition to the numbered references in the text, several contributors have provided additional reading suggestions for those wishing to delve deeper into the subject.

With reference to "The current TB epidemic", chapter 3:

Stop TB Partnership, World Health Organization (2006). "*The Global Plan to Stop TB, 2006–2015.*" Geneva, WHO. (WHO/HTM/STB/2006.35).

World Health Organization (2009). Global Tuberculosis Control Report: Epidemiology, Strategy and Financing. Geneva, WHO. (WHO/HTM/TB/2009.411).

World Health Organization (2008). Global Tuberculosis Control Report: Epidemiology, Strategy and Financing. Geneva, WHO. (WHO/HTM/TB/2008.393).

With reference to "Assessing the Needs", chapter 5:

Websites for consultation:

http://www.who.int/tb/en/

http://www.who.int/whosis/en/index.html
(statistical information system)

http://www.stoptb.org/partners/joinus.asp

http://www.who.int/tb/dots/laboratory/policy/en/

http://www.who.int/tb/research/en/

With reference to "Aiming for the right targets", chapter 6:

Agranoff D et al. (2006). Identification of diagnostic markers for tuberculosis by proteomic fingerprinting of serum. *Lancet*, 368:1012–1021.

Anthony R et al. (2006). Light emitting diodes for auramine O fluorescence microscopic screening of *Mycobacterium tuberculosis*. *International Journal of Tuberculosis and Lung Disease*, 10:1060–1062.

Arend S et al. (2008). Double-blind randomized phase I study comparing rdESAT-6 to tuberculin as skin test reagent in the diagnosis of tuberculosis infection. *Tuberculosis*, 88:249–261.

Bergval I et al. (2008). Development of multiplex assay for rapid characterization of *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*, 46:689–699.

Buijtels P et al. (2008). Rapid identification of mycobacteria by raman spectroscopy. *Journal of Clinical Microbiology*, 46:961–965.

Bottini M, Mustelin T (2007). Nanosynthesis by candlelight. *Nature Nanotechnology*, 2:599–600.

Brooks J et al. (1998). Development of a quantitative chemical ionization gas chromatography-mass spectrometry method to detect tuberculostearic acid in body fluids. *Journal of Chromatography B: Biomedical Sciences and Applications*, 712:1–10.

Bruno E et al. (2008). Can the electronic nose diagnose chronic rhinosinusitis? A new experimental study. *European Archives of Otorhinolaryngology*, 265:425–428.

Carlen ET, van den Berg A (2007). Nanowire electrochemical sensors: can we live without labels? *Lab Chip*, 7:19–23

Cheng MMC et al. (2006). *Current Opinion in Chemical Biology*. 10:11–19.

Cui X et al. (2008). Lensless high-resolution on-chip optofluidic microscopes for *Caenorhabditis elegans* and cell imaging. *Proceedings of the National Academy of Sciences USA*, 105:10670–10675.

Daikos G et al. (2004). Detection of tuberculostearic acid in serum and other biological fluids from patients with tuberculosis by electron capture-gas chromatography and chemical innisation-mass spectrometry. *International Journal of Tuberculosis and Lung Disease*, 8:1027–1031.

Derks R et al. (2008). Self-organized twinning of actuated particles for microfluidic pumping. *Applied Physical Letters*, 92:024104/1–024104/3,

Edgar R et al. (2006). High-sensitivity bacterial detection using biotin-tagged phage and quantum-dot nanocomplexes. *Proceedings of the National Academy of Sciences (USA)*, 103:4841–4845.

Fend R et al. (2006). Prospects for clinical application of electronic-nose technology to early detection of *Mycobacterium tuberculosis* in culture and sputum. *Journal of Clinical Microbiology*, 44:2039–2045.

Fu E et al. (2008) Chapter 10. In: Schasfoort BM, Tudos AJ, eds. *Handbook of surface plasmon resonance*. Cambridge, Royal Society of Chemistry.

Heintz A, Fink M, Mitchell B (2007). Mechanochemical synthesis of blue luminescent alkyl/alkenyl-passivated silicon nanoparticles. *Advanced Materials*, 19:3984–3988.

Huisken F et al. (2002). Light-emitting silicon nanocrystals from laser pyrolysis. *Advanced Materials*, 14:1861–1865.

Kaal E et al. (2009). A fast method for the identification of *Mycobacterium tuberculosis* in cultures and sputum based on thermally assisted hydrolysis and methylation followed by gas chromatography-mass spectrometry. *Journal of Chromatography A*, 1216:6319–6325.

Kang X et al. (2005). Proteomic fingerprints for potential application to early diagnosis of severe acute respiratory syndrome. *Clinical Chemistry*, 51:56–64.

Kerman K, Kobayashi M, Tamiya E (2004). Recent trends in electrochemical DNA biosensor technology. *Measurement Science and Technology*, 15:R1–11.

Kinsella JM, Ivanisevic A (2007). Taking charge of biomolecules. *Nature Nanotechnology*, 2:596–597.

Koehne J et al. (2004). Miniaturized multiplex label-free electronic chip for rapid nucleic acid analysis based on carbon nanotube nanoelectrode arrays. *Clinical Chemistry*, 50:1886–1893.

Krebs M et al. (2006). Novel technology for rapid species-specific detection of Bacillus spores. *Biomolecular Engineering*, 23:119–127.

Liu H, Ye T, Mao C (2007). Fluorescent carbon nanoparticles derived from candle soot. *Angewandte Chemie International Edition*, 46:6473–6475

Minnikin D, Hutchinson I, Caldicott A (1980). Thin-layer chromatography of methanolysates of mycolic acid-contaning bacteria. *Journal of Chromatography A*, 188:221–233.

Pai M, O'Brien R (2008). New diagnostics for latent and active tuberculosis: state of the art and future prospects. *Seminars in Respiratory and Critical Care Medicine*, 29(5):560–568

Papadopoulos M et al. (2004). A novel and accurate diagnostic test for human *African trypanosomiasis*. *Lancet*, 363:1358–1363.

Patiño S et al. (2008). Auto fluorescence of mycobacteria as a tool for the rapid detection of *Mycobacterium tuberculosis*. *Journal of Clinical Microbiology*, accepted for publication.

Peeters R et al. (2000). Open-path trace gas detection of ammonia based on cavity-enhanced absorption spectroscopy. *Applied Physics B: Lasers and Optics*, 71:231–236.

Petousis I et al. (2007). Transient behaviour of magnetic micro-bead chains rotating in a fluid by external fields. *Lab Chip*, 7:1746–1751.

Phillips M (1997). Method for the collection and assay of volatile organic compounds in breath. *Analytical Biochemistry*, 247:272.

Phillips M et al. (2007). Volatile biomarkers of pulmonary tuberculosis in the breath. *Tuberculosis*, 87:44–52.Piepenburg O et al. (2006). DNA detection using recombinate proteins. *PLoS Biology 4*, 7:e204. 0001–0006.

Popplewell J et al. (2005). Quantification of the effects of melittin on liposome structure. *Biochemical Society Transactions*, 3:931–933.

Rebuffo-Scheer C et al. (2007). Rapid species and strain differentiation of non-tubercoulous mycobacteria by fourier-transform infrared microspectroscopy. *Journal of Microbiological Methods*, 68:282–290.

Ren J et al. (2008). A new MSPQC for rapid growth and detection of *Mycobacterium tuberculosis*. *Biosensors and Bioelectronics*, 24:403–409.

Resch-Genger U et al. (2008). Quantum dots versus dyes as fluorescent labels. *Nature Methods*, 5:763–775.

>Ruzsanyi V et al. (2005). Detection of human metabolites using multi-capillary columns coupled to ion mobility spectrometers. *Journal of Chromatography A*, 1084:145–151. >

Schasfoort B, Tudos A (2008). *Handbook of surface plasmon resonance*. Cambridge, Royal Society of Chemistry.

Staii C, Johnson AT (2005). DNA-decorated carbon nanotubes for chemical sensing. *Nano Letters*, 5:1774–1778.

Stern E et al. (2007). Label-free immunodetection with CMOS-compatible semiconducting nanowires. *Nature*, 445:519–522.

Stopforth A et al (2005). A rapid method of diagnosing pulmonary tuberculosis using stir bar sorptive extraction–thermal desorption–gas chromatography–mass spectrometry. *Journal of Chromatography A*, 1071:135–139.

Syhre M, Chambers S (2008). The scent of *Mycobacterium tuberculosis*. *Tuberculosis*, 88:317–323.

Torrea G et al. (2008). Evaluation of the FluoreslenS and fluorescence microscopy blinded rechecking trial, Nairobi, Kenya. *International Journal of Tuberculosis and Lung Disease*, 10:1060–1062.

van der Werf M et al. (2007). Standard reporting requirements for biological samples in metabolomics experiments: microbial and *in vitro* biology experiments. *Metabolomics*, 3:189–194.

Van Deun A et al. (2008). Performance and stability of the FluoLED easy module for tuberculosis fluorescence microscopy. *International Journal of Tuberculosis and Lung Disease*, 12:1009–1014.

White J et al. (2008). Solid-state dye-labeled DNA detects volatile compounds in the vapor phase. *PLoS Biology*, 6:30–36.

Witkamp B et al. (2008). Self-detecting gate tunable nanotube paddle resonators. *Applied Physical Letters*, 93:111909–1–3.

Yager P et al. (2006). Microfluidic diagnostic technologies for global public health. *Nature*, 442:412.

References

Ymeti A et al. (2007). Fast, ultrasensitive virus detection using a young interferometer sensor. *Nanoletters*. 7(2):394–397.

Ymeti A et al. (2007). A single platform image cytometer for resource-poor settings to monitor disease progression in HIV infection. *International Society for Analytical Cytology*, 71A:132–142.

Zhao X et al. (2004). A rapid bioassay for single bacterial cell quantitation using bioconjugated nanoparticles. *Proceedings of the National Academy of Sciences USA*, 101:15027–15032. Zheng G et al. (2005). Multiplex electrical detection of cancer markers with nanowire sensor arrays. *Nature Biotechnology*, 23:1294–1301.

With reference to "Development and optimization", chapter 8:

Creveling C (2003). *Design for six sigma*. New Jersey, Prentice Hall.

Creveling C, Slutsky J, Antis D Jr (2007). *Six sigma for marketing process*. New Jersey, Prentice Hall.

McGrath M, Anthony M, Shapiro A (1992). *Product development: success through product and cycle-time excellence*. Oxford, Butterworth-Heinemann.

McGrath M (1996). *Setting the PACE in product development*. Oxford, Butterworth-Heinemann.

Shillito M (2001). *Acquiring, processing, and deploying voice of the customer*. St. Lucie Press.

World Health Organization. *Diagnostics for tuberculosis: global demand and market potential*. (2006). Geneva, WHO/TDR.

With reference to "Access", chapter 12:

Schünemann A et al. (2008). GRADE: grading quality of evidence and strength of recommendations for diagnostic tests and strategies. *BMJ*, 336:1106–1110.

NOTES

NOTES

Annex 1
Optimizing smear microscopy

Optimizing smear microscopy

The diagnostic process for tuberculosis (TB) is an important bottleneck impeding access to treatment. Laboratory facilities are fewer in number and more centralized than the facilities where patients can obtain treatment, as treatment centres are located as close as possible to patients' home communities. Further, current diagnostics are insensitive and/or time-consuming, and require laboratory infrastructure considerably more advanced than those commonly found in areas of the world where the burden is high. Since the increased detection of individuals with active disease is central to the global strategy for TB control, simple, sensitive and specific diagnostic approaches are desperately needed. First they must be developed, but they also must be capable of being made widely available and of producing results within the timescale of a clinical diagnosis.

Despite its shortcomings, sputum-smear microscopy is the most accessible diagnostic method for poor populations. Thus, approaches that improve its low sensitivity and reduce the need to examine multiple samples may preferentially benefit disadvantaged populations while decreasing laboratory workloads. Modelling exercises to evaluate the potential role of better diagnostics in improving TB control in developing countries have highlighted that improving the performance of sputum-smear microscopy and reducing the number of patients abandoning the diagnostic process produce significant public health benefits (1;2). As improved smear microscopy could play a major role in the current strategies for TB control, its optimization should be prioritized.

What steps could be used to optimize smear microscopy?

Potential approaches and technical modifications that could be used to optimise smear microscopy are summarized below. These represent developments currently under evaluation as well as ideas that need further development. As in all endeavours to improve microscopy, it is important to safeguard the microscopy's key strength: simplicity.

Improved sputum quality: No technique can achieve its potential without a good-quality specimen. Poor specimens are often obtained if samples are collected unsupervised (3) and from patients unaware of the reasons why repeated examinations are needed. Interventions to improve sputum quality could significantly increase the yield of smear microscopy and reduce gender discrepancies (4). These interventions could include standardized training packages (4), supervised methods to obtain and induce sputum (3) and, in appropriate settings, more invasive methods such as broncho-alveolar lavage (5). Although most of these methods have been described, few have been evaluated outside research environments.

Timing of sample collection: Patients usually submit sputum samples over two or three days, and many don't complete the process. WHO recently recommended that, as the third sputum sample only identifies about 5% of smear-positive patients not identified by the first two specimens, busy laboratories with good quality assurance systems could examine two instead of the three samples previously required. Current schemes for the collection of sputum would need to change to realize the full potential of this reduction, as two samples collected as spot-morning would still require at least two visits. More patient-friendly approaches could be developed if samples were collected more rapidly (e.g. collecting two samples on the spot) (6; 7) or if a higher proportion of samples are "frontloaded" the first day (e.g. as spot-spot-morning instead of the standard spot-morning-spot) to identify most smear-positive patients with samples submitted the first day. These combinations would comply with the revised WHO guidelines and are being evaluated. Research must assess whether the gains obtained with these accelerated schemes are maintained under field conditions. Accelerated schemes should also be examined in terms of reducing laboratory workloads and costs, and in terms of their acceptance by patients and programmes.

Combination of specimens: Most screening approaches collect samples from one anatomical area over several days (e.g. children are often hospitalized to obtain gastric aspirates over three consecutive days). Collecting several specimens from different anatomical sites in a single visit may result in the same yield. For example, a study in Yemeni children observed that collecting one gastric aspirate plus one nasopharyngeal swab and one induced sputum in one visit had the same sensitivity as three consecutive gastric aspirates over three days. These approaches have rarely been described in the literature and should be explored, especially in children.

Improved sputum processing: The chemical digestion/liquefaction of sputum using bleach, NaOH and others is attractive as they are easily available and eliminate cell artefacts in the smears, providing a clearer field. Several studies have suggested that this process may increase sensitivity by as much as 10%-15% (8). Further studies are necessary, as previous studies used different bleach concentrations and presentations; the quality of the bleach varies at source and there are no user-friendly methods to measure bleach activity and shelf life. There are also minor mechanical issues regarding drying of more liquid specimens in settings with high environmental humidity. WHO is currently sponsoring multicentre studies to evaluate a standardized approach for the use of bleach in disease-endemic countries (9).

Concentration methods: Centrifugation is probably the most effective concentration method for bleach or NaOH liquefied samples. However, care must be taken to centrifuge the sample long and fast enough. In a number of studies, the sensitivity of bleach treatment only increased when the samples were centrifuged, whereas sedimentation was hardly effective to concentrate the bacilli (8). Laboratory studies on cultures suggest that centrifugation above 3200g for more than 22 minutes is necessary to optimize the recovery of bacilli (10) and similar results have been found for spiked sputum (11). A major disadvantage of centrifugation is that peripheral laboratories rarely have suitable centrifuges.

A new technology (TB-Beads) has been reported to manually concentrate bacilli directly onto the slide by using paramagnetic beads coated with a chemical ligand that selectively binds to mycobacteria in sputum. The ligand does not bind many other bacteria, is not affected by pH and works in alkali conditions, allowing the direct capture of mycobacteria from thinned decontaminated sputum. The procedure does not require centrifuges and provides a cleaner sample by leaving many bacterial contaminants in the supernatant. The current protocol for this technology could be integrated with existing microscopy protocols and is currently under evaluation.

Sputum filtration

In this method, sputum is liquefied and passed through a filter, which is then stained or cultured by standard techniques. Filtration considerably concentrates mycobacteria, increasing sensitivity. Another advantage of using concentrated sputum is reduced time spent on sputum examination. Some studies have been published on this method, but further validation studies are required.

Vital fluorescent staining of sputum smears: Unlike most fluorescent stains, fluorescent vital dye only stains living, cultivable organisms. While a positive finding provides a basis for initiating antibiotic treatment, the sensitivity of the direct smear is highly variable, and the dye requires a fluorescent microscope. This technique might be most appropriate for use with patients not responding to therapy.

Improved microscopes: Fluorescent microscopes (FMs) have a higher sensitivity than light microscopes and slides are examined at lower magnification resulting in significant time savings (12). FMs are expensive and require a dark room and frequent bulb replacement; these factors result in many FMs being abandoned in disease endemic countries (DECs). Light-emitting diode (LED) fluorescent microscopes are becoming available that are low-cost, do not require dark rooms and feature long bulb life (13; 14). LED-FMs could be more accessible to laboratories that currently use light microscopes. Studies are under way to evaluate their sensitivity, specificity and acceptability (15). More research will be needed to standardize quality control/assurance mechanisms and to develop a logistical supply and maintenance infrastructure.

Simple and robust automated microscopic reading systems: Automating analysis is another way of improving the quality of microscopy. Given the long time needed to read sufficient fields on a slide, the time and quality of grading is linked to technician workload (16). Automation of the analysis would be an interesting option that could significantly reduce this workload and make analysis independent of operators and thus standardized. This could be possible using a fluorescent microscope equipped with a charge-coupled device (CCD) camera and a computer. Initial prototypes currently being evaluated require well-shaped mycobacteria, as false positives could lead to over-diagnosis and treatment of individuals without TB.

Combination of approaches and tests: Potential improvements to optimize smear microscopy are often reported singly, yet optimization probably will require a combination of techniques and approaches. For example, a package consisting of two supervised specimens in one visit (e.g. spot-spot) after providing suitable training, followed by the digestion and concentration of sputum plus examination using LED-based fluorescent microscopes, may result in additive or multiplicative gains across the system. This possibility

needs to be investigated. More evidence also is needed on how to combine improved smear microscopy packages with more sensitive but less easily accessible tests that require further visits. The goal is a streamlined diagnostic system that integrates emerging technologies and tests into a logical framework that is rapid, that facilitates access to treatment and that is cost-effective.

Annex 2
Rapid solid and liquid culture

Rapid solid and liquid culture

Although they share the common requirements of all diagnostic tests (quality assurance, training, timely result delivery), phenotypic diagnostic tests that use solid and liquid culture and drug susceptibility testing (DST) have special requirements related to sample delivery, sample processing and biosafety.

WHO recommends quality-assured, safe laboratory facilities for tuberculosis culture and DST. The basic laboratory requirements for smear microscopy, media preparation, sterilization, TB culture and DST using liquid culture include a negative-pressure containment facility. Standard operating procedures for culture and DST and a standard recording and reporting system should be in place. A national quality assurance system for the culture and DST programme, including panel proficiency testing where possible, should be in place before reporting for patient care is offered.

Infrastructure

Biosafety level (BSL) 3 laboratories are recommended for handling highly concentrated cultures for indirect susceptibility testing and species identification, though BSL2 could be used for methods requiring only specimen processing, culture inoculations and reading. Although new developments in "front-end" sample preparation are under evaluation, centrifugation is needed for most methods. Refrigerated centrifuges have been recommended to avoid high centrifugation temperatures reducing TB viability, though this has been questioned; if few centrifugations are performed per day, a non-refrigerated centrifuge is adequate and offers lower cost and maintenance. Only centrifuges with sealable lids and aerosol-containing canisters can be used, and these should be opened inside the biological safety cabinet. Space in the laboratory should be sufficient to house equipment and permit comfortable staff working conditions. Adequate staff training is mandatory to ensure acceptable isolation and DST testing results, and to ensure safety in the workplace.

Sample collection and transport

Specimens should be kept refrigerated and arrive at the laboratory within four days. Cetyl-pyridium chloride preservation is not compatible with liquid culture. Specimens from non sterile body sites should be processed according to a standard digestion-decontamination procedure such as the *N*-acetyl-L-cysteine-NaOH method. Sterile specimens should be centrifuged without decontamination.

Rationale for liquid culture systems

The Stop TB Partnership Laboratory Strengthening Subgroup and the World Health Organization (WHO) Strategic and Technical Advisory Group for Tuberculosis endorse the use of liquid TB culture and drug susceptibility testing in low-resource settings. Liquid medium is the best resuscitation and growth-support medium for quick recovery of mycobacteria from processed specimens. Its disadvantage is that liquid culture is much more sensitive to inoculum quality, i.e. remaining toxins and contaminants from the decontamination procedure. Liquid systems are more sensitive for detecting mycobacteria and may increase the case yield by 10% over solid media as well as reduce time to result from weeks to days. For drug susceptibility testing (DST), the delay may be reduced to as little as 10 days, compared to 28–42 days with conventional solid media. With increased sensitivity and reduced delays, liquid systems may significantly improve patient management.

Rapid solid and liquid culture

Solid media culture can only be deemed rapid if it also offers direct DST (as in the Griess method) or a detection method that does not depend upon the naked eye visualizing colonies (as in the thin-layer agar [TLA] method) and if this results in increased capacity to manage large quantities of samples at lower cost. Liquid media culture is inherently more rapid than solid media, and thus the automated liquid culture methods, all of which are validated for indirect (delayed) DST, are rapid culture methods. Automated methods detect growth in specially designed tubes or bottles produced and distributed under quality-assured conditions: An alternative to commercial products is microscopic observation drug susceptibility (MODS), a manual liquid culturing technique.

Time to detection

In addition to overall recovery, the time to detection (TTD) is an important performance characteristic of mycobacterial detection systems.

Contamination

Contamination with fungal or non-mycobacterial organisms most commonly derives from oral flora present in the sputum sample, but may also arise from contamination of media reagents within the laboratory. Highly enriched liquid culture is generally associated with higher rates of contamination than solid media. In experienced laboratories, approximately 5%–10% of specimens do not yield results because of contamination. Laboratories that have very low contamination rates (<1%) may well be over-decontaminating specimens thus reducing the yield of mycobacteria.

Cross-contamination

Laboratories must strictly follow procedures to prevent cross-contamination (due to carryover of MTB bacilli from positive to negative specimens), especially in high-incidence countries where laboratories process many positive specimens. In addition, laboratory staff should process specimens and perform culture inoculation in biological safety cabinets that have uninterrupted power supply systems to minimize their risk of exposure to possibly harmful aerosols.

Species identification

Many culture methods require a rapid and affordable method, either biochemical or molecular, of species identification to differentiate isolates of *M. tuberculosis* complex from nontuberculous mycobacteria.

Rapid phenotypic drug susceptibility testing (DST)

Indirect DST using the automated liquid culture platforms described above is validated and technically straightforward for first- and second-line DST, though this is highly costly. More simple and low-cost tests such as colorimetric redox methods and nitrate reductase assays perform very well; primary isolation is still required, and thus manipulation of cultures, so BSL3 laboratory facilities might be necessary. Yet DST takes less than two weeks after primary culture isolation, and emerging data suggest second-line DST is also possible. Direct DST, in which the sputum sample is inoculated onto cultures containing TB drugs and drug-free control cultures, underlie the Griess and MODS methodologies (on solid LJ and liquid

modified Middlebrook 7H9, respectively). These deliver a more rapid DST result effectively concurrent with primary isolation, and have the advantage of not requiring potentially hazardous manipulation of concentrated suspensions of mycobacteria for secondary inoculation. Drawbacks are that performance for first-line drugs other than rifampicin and isoniazid is hampered by the uncontrolled inoculum concentration, and data on second-line DST are not yet available.

Annex 3
Antigen detection tests
to diagnose active TB

Antigen detection tests to diagnose active TB

The Antigen detection assay, in contrast to conventional serological tests, detects disease status and not the host antibody response to the disease etiological agent; therefore it can be used for both diagnosis and treatment follow-up. Antigen detection assay has for a long time been successfully used to diagnose infectious diseases, including sore throat caused by *Streptococcus pyogenes*, pneumonia caused by *Streptococcus pneumoniae* or *Legionella pneumophilla*, hepatitis and amoebiasis. Although antigen detection assay has the potential to discriminate latent from adult or pulmonary TB, this test has only recently been under development. Early attempts to develop these tests met with variable success, mostly because the target antigens used in these initial studies comprised a mixture of native *M. tuberculosis* plasma membrane proteins (17; 18). An antigen capture immunoassay using rabbit antibodies raised against these antigens and evaluated with serum samples from patients with active TB revealed a sensitivity of 45%. More recently, antigen-detection assays were described with somewhat more abundant antigens like lipoarabinomannans (LAM). LAM are phosphorylated lipopolysaccharides, which are a major cell wall component of all bacteria of the genus *Mycobacterium* (19; 20; 21). Most of these tests have used a capture ELISA format to detect LAM in the sputum (22) or urine (22; 24; 25; 26) of tuberculosis patients. Overall, the test sensitivity was 74% in smear-positive samples and 57% in AFB smear-negative samples. Unfortunately, 10% of healthy controls were also positive. In addition to LAM, defined *M. tuberculosis* protein antigens have also been target of antigen detection assay development. These proteins include a 32kDa protein and a 45-47kDa coded for by the gene alanin-prolinerich antigen complex (APA), both present in the *M. tuberculosis* culture filtrate. Capture ELISA assembled with specific antibodies to these antigens and sputum as patient's sample had sensitivities of ~56% and <40% respectively (27; 28). Finally, antigen detection assay in sputum has also been tested to follow tuberculosis treatment efficacy (29). This approach used the antigen 85 complex because this molecule is critical in cell wall biosynthesis and is induced by isoniazid in vitro, making it an interesting candidate molecule to achieve this purpose. The results were promising, indicating that patients with levels of Ag85 <60 pg/mL of sputum responded rapidly to treatment and were cured.

Thus far, only LAM-based tests have moved to further development. Whether the performance of LAM tests will allow them to eventually be useful for TB diagnosis in the field or hospital setting remains to be seen. The conflicting results thus far observed with the

LAM detection tests can perhaps be explained by the lack of specificity of the anti-LAM antibodies that are used as capture molecules. Anti-*M. tuberculosis* LAM antibodies, even monoclonals, are not specific for *M. tuberculosis* and cross-react with the majority of *Mycobacterium* species. In particular, tests based on LAM as antigen would likely be positive in patients who are infected with *M. avium*, a common opportunistic infection in patients with HIV/AIDS, as well as in those infected with *M. leprae* (20), the agent of leprosy. Moreover, preparation of LAM involves the growth of *M. tuberculosis* followed by several purification steps, and the final product may differ from laboratory to laboratory, particularly in terms of the molecule's mannose content (27). However, recent studies from India, South Africa and Zimbabwe show that the problem with urine LAM is poor sensitivity; specificity is less of an issue but still a concern. In general, data suggest that LAM sensitivity is better in HIV-infected than HIV-negative TB patients.

Alternatively, *M. tuberculosis* proteins might be useful candidate molecules for the development of a specific and sensitive antigen detection assay for tuberculosis. Supporting this possibility are the recent identification by mass spectroscopy of four *M. tuberculosis* proteins shed during active disease and eliminated in the patient's urine (28). One of these molecules is present only in the organisms of the *M. tuberculosis* complex. In addition, a capture ELISA formatted with rabbit antibodies specific for one of these proteins indicated the potential of these antigens as targets for antigen detection assay development (29).

Antigen detection assay and specimen diversity

An antigen detection assay for tuberculosis is uniquely attractive because tests can be performed using a variety of patient specimens (e.g. sputum, blood, faeces, urine and spinal fluid). This is highly applicable to tuberculosis because the disease can affect multiple organs and systems. Therefore, antigens that are shed from *M. tuberculosis* present in infected tissues can be present in the bodily fluids surrounding these tissues (e.g., respiratory tract mucus, faeces, spinal fluid). Moreover, shed antigens can reach the blood circulation and be eliminated in urine, a highly practical specimen for diagnostic tests.

Antigen detection assay using urine

Urine is one of the easiest and least invasive samples that can be collected from a patient. This is particularly important for children as they generally do not produce sputum, which is the most common sample currently used to perform tuberculosis diagnostic tests in adults. In children, gastric lavage is used to investigate the presence of pulmonary tuberculosis. This sample collection is very invasive and dangerous.

In contrast, urine is a readily available sample for point-of-care diagnosis of tuberculosis. The presence of *M. tuberculosis* antigens excreted in urine may be detectable sooner than the presence of mycobacteria in sputum. This is important for earlier therapeutic intervention and better cure efficacy. In addition, urine may facilitate TB diagnosis in HIV-co-infected patients, since they often have reduced mycobacterial loads in their sputum.

In patients with extra-pulmonary TB (e.g. bone, lymph nodes, meninges), the antigens produced by *M. tuberculosis* are likely to leak into the lymphatic/blood circulation to be eliminated in the patient's urine, not in sputum.

Antigen detection assay using breath

A technology using bio-optical sensors is being developed to detect tuberculosis antigens in patient breath. This technology is under evaluation at different test sites under the trade name of RBS TB Breathalyser. This unit comprises of a single-use sample tube containing a bio-chemically coated glass sensor which the patent coughs onto. This tube is placed into a reader which analyses the sample using a displacement assay. A diode laser reads the glass sensor containing fluorescent-coated analogues. The TB antigens displace the fluorescent-coated analogues, causing a reduction in the fluorescent signal generated when the diode laser passes over the glass sensor.

Current stage of ADA for tuberculosis

As of this document's writing, there was no US FDA-approved antigen detection assay for tuberculosis. However, a LAM-based capture ELISA antigen detection assay has been commercially available through Inverness Medical Innovations. This test targets urine samples as the specimen of choice and is primarily aimed at HIV-infected persons. Moreover, FIND is also actively involved in the validation of other LAM-based antigen detection assay formats, including a dipstick test. Final results, proof of utility and earliest date for availability for adoption in the public sector is 2010. Another product, an *M. tuberculosis* protein-based antigen detection assay, is under development at the Forsyth Institute in conjunction with InBios Inc. Clinical validation studies are to begin in 2009 and 2010 in Brazil and Peru.

Annex 4
Antibody detection

Antibody detection

Current status of antibody detection-based diagnostic tests

Although often referred to as serodiagnostic tests (SD), antibody-based tests can in theory be used on any fluid from an infected individual (30; 31). Unlike CMI-based tests, SD tests do not require live cells or culture, improving the feasibility of adaptation to multiple simple formats (ELISA, lateral flow, dipstick) that can be integrated in laboratories at different levels of health care systems. There are many accurate, simple SD tests for other diseases, so the technology is clearly mature. Efforts to devise a SD for TB date back decades (32-35) and include a vast amount of confusing and contradictory literature. Approximately 20 SD assays are currently sold in developing countries (36), yet recent systematic analysis of commercially available tests demonstrated that their performance does not warrant any role in diagnosis of pulmonary or extrapulmonary TB (37; 38).

Shortcomings of serodiagnostic tests

Most current SD tests are based on a small number of antigens, such as the 38 kDa protein, HSP60 and LAM (35). These antigens were chosen largely on the basis of abundance/availability or recognition by sera from limited patient pools (for example, the 38 kDa is strongly recognized only in patients with advanced cavitary disease). There have been few attempts to systematically identify antigens strongly and broadly recognized by different clinical cohorts (39-41). None of the commercial tests has been evaluated in rigorous clinical trials (42; 43) nor are any approved by regulatory agencies (e.g. WHO, US FDA). The claims in package inserts, frequently described as inflated, are generally based on small, poorly designed and poorly conducted trials (37; 38).

Challenges in developing SD antibody tests for TB

A diagnostic test that aims to specifically detect the immune responses of the host against a given pathogen must be based on bacterial products that the host expresses during the disease process. Studies in the last decade have identified several factors to be considered in devising a TB SD test.

TB is a chronic disease that presents with a spectrum of pathological, immunological and bacteriological stages. Since the *in situ* environment of the bacteria differs at different stages of TB, *M. tb* regulates the expression profile of genes/proteins to adapt to its environment (42; 43). The repertoire of antigens that elicit antibodies changes as the disease progresses (39; 40). Thus either the antigens selected for SD must be those expressed at all stages of active TB, or multiple antigens that encompass all stages of active TB need to be included.

There is incomplete overlap between the antigen repertoires that elicit antibodies in HIV-/TB+ and HIV+/TB+ patients. The selected antigens must be recognized in both types of patients (39; 40).

Smear-negative patients are among the more difficult patient classes to positively identify and have lower antibody titers, presumably due to lower bacterial load and therefore lower antigenic load. This shows a need for multiple antigens to boost sensitivity (44).

Most SD tests detect free antibodies. The titers of antibodies to each antigen varies in patients, and antibodies form immune complexes. Multiple antigens will be required to ensure detection of free antibodies directed against some of the antigens used in SD tests.

Several *M. tb* antigens have post-translational modifications that affect their recognition by antibodies (45; 46). Considering that a SD test based on native *M. tb* antigens is unlikely to be commercially feasible, only recombinant antigens that can bind with antibodies against the native protein can be focused upon.

M. tb is a slow-growing pathogen; it can take months or years before the infection or reactivation manifests as bacteriologically or radiologically detectable or as symptomatic clinical TB. Scant information is available on immune responses at this incipient, sub-clinical stage of active infection. Studies with HIV+/TB+ patients have shown that antibodies to *M. tb* antigens can be detected in patients for months prior to clinical manifestation of TB (47-50). Inclusion of subjects at high risk for TB (household contacts of smear-positive patients, HIV+ patients from TB-endemic countries) in evaluation of candidate antigens/SD tests will provide falsely low specificity.

Infection with *M. tb* typically results in latent infection with a relatively low risk of reactivation. Unless based on carefully selected antigens that are not recognized during latent TB, SD tests may merely confirm infection rather than be clinically diagnosed despite being highly sensitive and specific.

The current status of SD antibody tests

Antibodies to several antigens (including malate synthase and TBF6) are present in HIV+ sputum smear-positive patients. Interestingly, there is some evidence that serological responses in HIV+/TB+ patients are typically stronger than those in HIV-/TB+ patients (41;44, author's unpublished data), perhaps because of HIV-induced hypergammaglobulinemia and/or higher bacterial load.

Among lipid antigens, only cord factor provided high specificity (91%, 95% CI 78-97).

Assays detecting IgG and/or IgA antibodies provided higher sensitivities than assays detecting IgM antibodies. The inclusion of IgM antibody detection offered no advantage compared to the detection of the other two immunoglobulin classes.

Recommendations for SD antibody test development

Although progress has been made in antigen discovery for TB SD, the process needs to be accelerated and improved on several fronts.

Additional targets are needed that can surpass the performance of the antigens identified so far or add to their value. Combinatorial studies can identify the best balance of antigens to maximize test performance versus test cost.

Sensitive yet simple technologies are needed to detect low titers of antibodies, such as screening against multiple targets (51).

Peptides or polyproteins derived from selected highly immunogenic antigens have been used in devising SD tests for other infectious diseases (HIV), and proof of principle has been shown for TB (49; 52); this option could be further explored.

Most critically, new SD products resulting from such research need to be tested rigorously using GCP methodology using effectively-powered studies in relevant test populations that also employ cost-benefit and impact analyses.

Annex 5
T-cell-based interferon-gamma release assays

T-cell-based interferon-gamma release assays

Background

Until recently, the diagnosis of latent tuberculosis infection (LTBI) depended solely on the tuberculin skin test (TST), an imperfect test with known limitations (53). A major advance in recent times has been the development of T-cell-based interferon-gamma release assays (IGRAs). IGRAs are in-vitro blood tests that are based on interferon-gamma (IFN-g) release after stimulation by antigens (such as early secreted antigenic target 6 [ESAT-6] and culture filtrate protein 10 [CFP-10]) that are more specific to *M. tuberculosis* than the purified protein derivative (PPD) used in tuberculin skin tests. ESAT-6 and CFP10 are encoded by genes located within the region of difference 1 (RD1) segment of the *M. tuberculosis* genome; they are more specific than PPD because they are not shared with any of the BCG vaccine strains or certain species of non-tuberculosis mycobacteria (NTM) (e.g. *M. avium*). The use of such specific antigens in an *ex-vivo* assay format is a distinguishing feature of IGRAs, which are essentially tests of cell-mediated immune response.

Test principles

Two IGRAs are currently available in many countries: the QuantiFERON-TB Gold In-Tube (QFT) assay (Cellestis Ltd, Carnegie, Australia) and the two-stage T-SPOT.TB assay (Oxford Immunotec, Oxford, UK). Both tests are US FDA-approved and CE-marked for use in Europe.

The QFT assay is an ELISA-based whole-blood test that uses peptides from three TB antigens (ESAT-6, CFP-10 and TB7.7) in a simple in-tube format (54). The QFT assay is performed in two stages. In the first stage, three 1.0 mL aliquots of heparinized whole blood are incubated in blood collection tubes with TB-specific antigens, mitogen control or nil control. Following a 16–24 hour incubation period, the blood collection tubes are centrifuged, the plasma removed and the amount of IFN-γ is quantified by ELISA. A person is considered positive for *M. tuberculosis* infection if they have an IFN-γ response to TB antigens above the test cut-off. The mitogen-stimulated plasma sample serves as a positive control for each individual tested. The nil sample adjusts for background and nonspecific IFN-γ in blood samples.

T-SPOT.TB is an ELISPOT (enzyme-linked immunospot) assay that uses ESAT-6 and CFP-10 peptides (55). In Stage 1, blood samples are collected into sodium citrate, sodium heparin or lithium heparin blood collection tubes with peripheral blood mononuclear cells (PBMCs) subsequently separated using standard cell separation techniques. PBMCs are then washed and counted, and specified numbers are added to each well in a microtiter plate coated with monoclonal antibodies to IFN-γ. TB-specific RD1 antigens are then added and the plate is incubated for 16–20 hours, causing the release of IFN-γ from sensitized T-cells. This is captured by specific antibodies in the microtiter plate. A conjugated secondary antibody is added to visualize the secreted IFN-γ, which is seen as spots on the microtiter plate well. Each spot represents one activated T cell. A person is considered positive for *M. tuberculosis* infection if the spot counts in the TB antigen wells exceed a specific threshold relative to the control wells.

While QFT and T-SPOT.TB are both considered commercial short-term incubation IGRAs that use RD1 antigens, they have different features. While the QFT uses whole blood in an ELISA platform, the T-SPOT.TB requires separated PBMCs for the ELISPOT platform. The read-out in the QFT assay is quantification of the amount of IFN-g as international units (IU) per mL, while the read-out in the T-SPOT.TB assay is number of IFN-g-producing T cells (spot-forming cells).

Technical issues and challenges

IGRA implementation presents several technical issues.

The blood draw might pose challenges in very small children, because the currently available assays require venupuncture and 3 mL of blood as a minimum.

Once the blood is drawn, it must be processed within a specified time and incubated, lest test results be affected by poor lymphocyte viability. Until incubation in the laboratory, a portable incubator can be used with the In-Tube QFT version, but not with the T-SPOT.TB test. Furthermore, the T-SPOT.TB assay requires separation of PBMCs and counting of cells, making it technically more challenging than the QFT assay.

Indeterminate results are likely in immunocompromised individuals with low CD4+ cell counts, especially with the QFT assay. Thus it is important to run the assays with positive and negative controls.

IGRAs are dynamic tests and results vary over time within the same individuals, causing conversions and reversions; this has implications for serial (repeated) testing. Finally,

cost considerations and the need for a well-equipped laboratory and trained technicians may influence this technology's use in high-burden countries.

Evidence on accuracy and reliability

Table 1 summarizes the current evidence on IGRAs based on recent systematic reviews and meta-analyses (56; 57; 58). There is strong evidence that IGRAs, especially QFT, have excellent specificity unaffected by BCG vaccination. TST specificity is high in non-BCG-vaccinated populations but low and highly variable in BCG-vaccinated populations. In low-incidence settings, IGRAs appear to correlate well with markers of TB exposure. Overall, the high specificity of IGRAs might prove to be useful in BCG-vaccinated individuals, particularly in settings where TST specificity is compromised by BCG vaccination after infancy or by multiple BCG vaccinations (53; 59). Sensitivity of IGRAs and TST is not consistent across tests and populations, but T-SPOT.TB appears to be more sensitive than QFT and TST. This could be partly because the cut-off for T-SPOT.TB is designed to maximize sensitivity, while the cut-off for QFT is designed to maximize specificity. The higher sensitivity of T-SPOT.TB may be clinically useful in evaluating high-risk populations with immunosuppressive conditions. However, the diagnosis of active TB rests on microbiological detection of *M. tuberculosis.* Immune-based tests, such as IGRAs and TST, do not directly detect *M. tuberculosis;* they merely indicate a cellular immune response to recent or remote sensitization with *M. tuberculosis.* Because IGRAs cannot distinguish between LTBI and active TB, a positive IGRA result may not necessarily indicate active TB. Furthermore, a negative IGRA result would not conclusively rule out active disease in an individual suspected to have TB; this also applies to the tuberculin skin test.

Despite the substantial body of literature on IGRAs, almost all the available studies have limitations, namely that they lack a gold standard for LTBI, cross-sectional design, and use of sensitivity and specificity as surrogates for patient-important outcomes. These tests also lack adequate data on outcomes such as accuracy of diagnostic algorithms (rather than single tests), incremental or added value of IGRAs, impact of IGRAs on clinical decision-making and therapeutic choices, and IGRAs' prognostic ability to accurately identify individuals with LTBI at highest risk of progressing to active TB, and therefore most likely to benefit from preventive therapy. Available evidence on IGRAs cannot be considered high-quality, and further research will influence future recommendations and guidelines. Ongoing studies should resolve these issues within the next few years and inform evidence-based guidelines on how to implement IGRAs in clinical practice (60-62).

Table 1. A comparison of tuberculin skin test (TST) with interferon-gamma release assays (IGRA)

Performance and Operational Characteristics	TST	IGRA	
		QFT-Gold In Tube	T-SPOT.TB
Estimated sensitivity (in patients with active TB)	70 – 90%	75 – 85% (sensitivity higher in low incidence compared to high incidence countries)	90 – 95% (sensitivity higher in low incidence compared to high incidence countries)
Sensitivity in immunocompromised individuals (predominantly HIV-infected)	Low (50 - 70%)	Lower sensitivity (60 - 70%) in immuno-compromised populations; indeterminate results more likely if CD4+ counts are low	Lower sensitivity (70 - 80%) in immuno-compromised populations
Estimated specificity (in healthy persons with no known TB disease/exposure)	60 – 100% (variable; high in BCG non-vaccinated; low when BCG given after infancy or if BCG is repeated)	95 – 100% (not affected by BCG vaccination)	90 – 95% (not affected by BCG vaccination)
Cross-reactivity with BCG	Yes	No	No
Cross-reactivity with nontuberculous mycobacteria	Yes	Less likely (but data are limited)	Less likely (but data are limited)
Association between test-positivity and subsequent risk of active TB during follow-up	Modest positive association	Limited evidence, but appears similar to TST	Limited evidence, but appears similar to TST
Correlation with M. tuberculosis exposure	Yes	Yes (correlated better with exposure than TST, mostly low incidence settings)	Yes (correlated better with exposure than TST, mostly low incidence settings)
Benefits of treating test-positives (based on randomized controlled trials)	Yes	No evidence	No evidence
Reliability (reproducibility)	Moderate; within-subject variations occur.	Limited evidence; within-subject variations occur	Limited evidence; within-subject variations occur
Inter-reader variability	Yes (inter and intra-reader variations occur)	No	Variation in counting spots (if done manually)
Boosting phenomenon (initial test influencing the subsequent test)	Yes	No	No
Impact of initial TST on subsequent IGRA result	Not applicable	Likely	Likely
Potential for conversions and reversions	Yes	Yes	Yes
Patient visits to complete testing protocol	Two	One	One
Material costs	Low	Moderate	High

Sample processing and assay complexity	Not applicable	Moderate	High
Laboratory infrastructure required	No	Yes	Yes
Time to obtain a result	48-72 hours	24-48 hours (longer if assays batched)	24-48 hours (longer if assays batched)
Trained personnel required	Yes	Yes	Yes

Adapted from:
Pai M et al. Ann Intern Med 2008;149(3):177-84
Pai M et al. Clin Chest Med 2009;30(4):701-716
Dheda K et al. Curr Opin Pulm Med 2009;15(3):188-200

Annex 6
Nucleic acid amplification test (NAAT)

Nucleic acid amplification test (NAAT)

The goal of detecting and identifying mycobacteria directly from clinical specimens, and thereby avoiding the time required for culturing, has been difficult to achieve because clinical specimens usually contain small numbers of bacilli. Direct detection requires either an extremely sensitive and specific assay or a process by which a diagnostically useful mycobacterial component can be "amplified" to a detectable level. Commonly used laboratory methods for amplifying nucleic acids include polymerase chain reaction (PCR) amplification, transcription-mediated amplification (TMA), strand displacement amplification (SDA) and ligase chain reaction (LCR) amplification.

Polymerase chain reaction (PCR) uses oligonucleotide primers to direct the amplification of a target sequence via repeated rounds of denaturation, primer annealing and primer extension. Many PCR assays for the detection of *M. tuberculosis* have been described in recent years. General features of these are:

a) These assays require a thermocycler and a thermostable DNA polymerase, usually Taq polymerase;

b) PCR can generate $>10^7$ copies of a target molecule in about two hours;

c) The specificity of the amplification process lies in the choice of primers for directing DNA synthesis or in the choice of hybridization probes to detect the amplified product;

d) Internal controls can be included to assess amplification efficiency and the presence of inhibitors; and

e) The assay can provide semi-quantitative results.

The Amplicor Mycobacterium Tuberculosis Test, Cobas Amplicor test (Roche Diagnostics) and *Mycobacterium tuberculosis* Amplified Direct Test (Gen Probe, San Diego California) are commercially available PCR tests to detect *M. tuberculosis* complex bacteria in respiratory specimens. Other manufacturers are also entering the market with PCR-based tests to detect *M. tuberculosis*. These tests use PCR to amplify a portion of the 16S rRNA gene that contains a sequence that hybridizes with an oligonucleotide probe specific for the *M. tuberculosis* complex. In addition to the commercially available PCR test kits, many PCR tests have been developed in laboratories during the past decade that use different targets for

detecting *M. tuberculosis* bacteria in clinical specimens. Such tests usually are referred to as in-house or "home-brew" tests.

Transcription-mediated amplification (TMA) uses a bifunctional oligonucleotide primer containing a portion that serves as a transcription promoter and a portion that hybridizes specifically to the target, allowing direct copying of target RNA into a double-stranded molecule which can be transcribed to produce more RNA molecules. The new transcripts are then templates for reverse transcription and further amplification. Performance characteristics of this method include: a) it is an isothermal reaction requiring reverse transcriptase and DNA-directed RNA polymerase, b) $>10^9$ copies of the target molecule can be generated in about two hours, and c) the specificity of the assay lies in the specificity of the primers used for the reverse transcription step and in the probes used in the hybridization assay. The Enhanced Amplified *Mycobacterium Tuberculosis* Direct Test (E-MTD, Gen-Probe) is a commercially available test which combines TMA of a portion of the 16S rRNA with a detection method that uses a hybridization probe specific for the *M. tuberculosis* complex to detect bacteria in respiratory specimens.

Strand displacement amplification (SDA) takes advantage of the ability of DNA polymerase to start at the site of a nick in double-stranded DNA, extend one strand from the 3'-end and displace the downstream strand of DNA. Repeated rounds of DNA synthesis, nicking and strand-displacement amplify the target molecule. This method involves an isothermal reaction requiring the Klenow fragment of *E. coli* DNA polymerase and a restriction enzyme such as HinCII. It also can generate $>10^8$ copies of the target molecule in about two hours. Its specificity depends on the primers chosen to direct DNA synthesis or on the hybridization probes selected to detect the amplified product. This method also allows for inclusion of internal controls to assess amplification efficiency and the presence of inhibitors, and it can provide semi-quantitative results. The BD-ProbeTec and BD-ProbeTec-ET (Becton Dickinson) are commercially available tests which use SDA of a portion of the IS*6110* element (an *M. tuberculosis* complex-specific insertion element) or the 16s rRNA to detect *M. tuberculosis* complex bacteria in respiratory specimens.

The ligase chain reaction (LCR) assay is based upon the ability of the target DNA to direct the covalent joining of two oligonucleotides by DNA ligase. Repeated rounds of denaturation, annealing and ligation are used to amplify the ligated products. LCR assay requires a thermocycler and a thermostable DNA ligase, and it can can generate $>10^7$ copies of the target molecule in about two hours. The assay's specificity lies in the specific hybridization of oligonucleotides to the target DNA. The commercially available Abbott LCx

test uses LCR to amplify a portion of the gene encoding protein antigen b to detect *M. tuberculosis* complex bacteria in respiratory specimens.

Loop-mediated isothermal amplification (LAMP) employs primers complementary to various regions of the target gene. The LAMP primers anneal to the complementary sequence of double-stranded target DNA, initiate DNA synthesis in a strand displacement activity mode and release a single-stranded stem-loop structured DNA. Loop primers can bind the single-stranded DNA and provide additional starting points for DNA synthesis, allowing for a 10^9-10^{10} fold increase in target DNA within 30 minutes. A LAMP-based assay (MTB-LAMP, Eiken Chemical) has been developed with a set of six specific primers that target the *M. tuberculosis* 16S rRNA gene to detect *M. tuberculosis* complex bacteria in respiratory specimens.

Fully automated NAAT platforms such as the GeneXpert (Cepheid) are currently under evaluation. This platform combines sample preparation with real-time PCR amplification and detection of MTB and rifampicin resistance. These simple-to-use systems have turnaround times of only minutes using sputum samples, but they are costly and require evaluation in field settings to assess their usefulness.

A number of meta-analyses have reviewed the accuracy and reliability of NAA tests for the detection and identification of *M. tuberculosis* complex bacteria in clinical specimens. In general, NAA tests have high specificity and positive predictive value but modest and highly variable sensitivity for laboratory confirmation of tuberculosis. Performance is more accurate and reliable with AFB smear-positive respiratory specimens and more variable with AFB smear-negative and non-respiratory specimens (i.e. extrapulmonary TB). For example, one meta-analysis found that the pooled sensitivity of the commercially available NAA tests was 0.85 (range 0.36–1.00) and the pooled specificity was 0.97 (range 0.54–1.00) for detection of *M. tuberculosis* complex bacteria in respiratory specimens. The performance of in-house NAA tests has been more variable than that of the commercially available tests. Issues around NAAT quality assurance have also been raised (64). The US CDC released guidelines in 2009 on the use of NAAT in TB diagnosis.

Improvement in the laboratory aspects of NAA testing for TB will require investment in research, including a) research to develop and drive the implementation of new and better technologies, including improved DNA extraction and concentration methods; b) cost-benefit analysis of services and technologies; and c) operational research to support science- and experience-based recommendations. An ideal test would:

• be an affordable point-of-care test with a turnaround time of <2 hours;

• be automatable with minimal hands-on time;

• require minimal specimen processing;

• employ a closed system to minimize end-product contamination concerns;

• include internal controls for inhibitors and test performance; and

• have a lower limit of detection for *M. tuberculosis* bacteria in a specimen (i.e. improved analytical sensitivity) and higher clinical sensitivity for detecting active TB than currently available tests.

Annex 7
Molecular drug resistance testing

Molecular drug resistance testing

Rapid molecular drug resistance tests are a pressing public health and diagnostic need given the sharp increases in MDR-TB and XDR-TB globally. Drug resistance in *M. tuberculosis* bacteria arises mainly through chromosomal sequence mutations that:

a) Block the activity of a drug (mutations in *rpoB* prevent binding of rifampin to RNA polymerase and inhibition of transcription);

b) Block activation of a prodrug (mutations in *katG* lead to loss of the ability of catalase to activate the prodrug isoniazid to its active form); or

c) Produce an activity that binds or destroys the drug, as mutations in *inhA* increase the amount of *InhA* protein, interfering with isoniazid activity by binding sufficient isoniazid to reduce its effective concentration in the bacterium to below an inhibitory level.

The mutations associated with resistance to many of the antituberculosis drugs have been identified, though much work remains to be done to identify the molecular basis of resistance for some drugs and to determine the predictive value of particular mutations in a strain of *M. tuberculosis*. For example, approximately 95% of rifampin-resistant *M. tuberculosis* strains carry mutations within the RIF resistance-determining region (RRDR), an 81-bp region encoding codons 507 through 533 of the *rpoB* gene

Molecular tests for detecting drug-resistant TB are in general variations of NAA tests. PCR typically is used to amplify a target sequence, followed by a second assay to determine if the sequence contains a mutation associated with resistance. Methods for the latter include DNA sequencing, pyrosequencing, electrophoretic detection methods (e.g. single-strand conformation polymorphism), methods for detecting mismatches in heteroduplexes (e.g. temperature gradient HPLC analysis or branch migration inhibition) and hybridization assays (e.g. molecular beacons, microarrays, membrane hybridization and line-probe assays). Commercially available kits include line-probe assays (INNO-LiPA Rif.TB from Innogenetics, GenoType MTBDR Plus from Hain LifeScience) and microarray assays (CombiChip Mycobacteria DR, GENE IN) for detecting mutations associated with rifampin resistance. Some also detect mutations associated with isoniazid resistance. In-house PCR-based tests using molecular beacons also have been used for diagnostic purposes in a few clinical laboratories.

For the hybridization assays, the region of the target gene associated with resistance is PCR amplified, and the labelled PCR products hybridized with oligonucleotide probes immobilized on a strip (line-probe assays) or in a microarray. Mutations are detected by lack of binding to wild-type probes and/or by binding to specific probes for the most commonly occurring mutations. Meta-analyses have evaluated the INNO-LiPA Rif.TB and GenoType MTBDR Plus assays' performance relative to conventional culture-based drug susceptibility tests. For the INNO-LiPA Rif.TB assay, the pooled sensitivity was 0.97 (95% CI 0.95-0.98) and the pooled specificity 0.99 (95% CI 0.98-1.00) for detecting rifampin resistance in isolates of *M. tuberculosis*. Overall discriminatory ability of the assay was 99% and overall accuracy 97%, with all studies yielding consistently high test performance values. In four studies, the INNO-LiPA Rif.TB showed 100% specificity and sensitivity ranging from 80% to 100% for detecting rifampin resistance in *M. tuberculosis* bacteria directly from clinical specimens.

For the MTBDR and MTBDR Plus assays, the pooled sensitivity was 0.98 (95% CI 0.96-0.99) and the pooled specificity was 0.99 (95% CI 0.97-0.99) for detecting rifampin resistance in isolates of *M. tuberculosis* or directly from clinical specimens. Overall discriminatory ability of the assay was 99% and overall accuracy 97%, with all studies yielding consistently high performance values.

Molecular beacons are hybridization probes which emit fluorescence only when hybridized to their target. Molecular beacons can discriminate between targets differing only by a single nucleotide. Because molecular beacons can be coupled to different-coloured fluorophores, PCR assays can be designed in which different DNA fragments or mutations can be amplified and detected simultaneously in the same tube. For example, a single-well molecular beacon assay has been developed that uses five molecular beacons to detect mutations associated with rifampin resistance in *M. tuberculosis* bacteria. In laboratory-based studies, the performance of this in-house test for detecting rifampin resistance in isolates of *M. tuberculosis* or directly from clinical specimens appears to be similar to that of the line-probe assays. An automated PCR-based molecular-based system developed by Cepheid also is entering clinical trials.

Molecular drug resistance tests for the other antituberculosis drugs are much less developed and studied than the tests for rifampin resistance. A meta-analysis of the performance of the MTBDR Plus assay for detecting isoniazid revealed a pooled sensitivity of 0.85 (95% CI 0.77-0.90) ranging from 57%-100% and a pooled specificity of 0.99 (95% CI 0.98-1.00) which was fairly consistent across studies.

Tests for the other key resistances, especially XDR-TB-defining resistances, are in various development stages, from discovery of the mutations associated with resistance through development of prototype assays and laboratory-based evaluations. Databases of mutations associated with drug resistance, such as www.tbdreamdb.com, are likely to help in developing and improving diagnostic tests.

Annex 8
Phage-based tests

Phage-based tests

Mycobacteriophages have been used to detect the presence of *M. tuberculosis* complex (MTB) and determine drug susceptibility from clinical specimens and isolates. Two main phage-based approaches have been used. In phage amplification, phages are amplified in rapid-growing mycobacterial "helper" cells (*M. smegmatis*) following phage replication and lysis of viable MTB. The other approach is detection of light produced by luciferase reporter phages (LRP) after infection of viable MTB. Drug susceptibility testing involves determination of MTB viability by phage replication following incubation of samples with and without the test drug.

Phage amplification

Phage amplification is a manual method requiring no specialized equipment besides facilities similar to those needed by solid culture; results are available in two days. Commercial (FASTPlaque) and in-house methods are available using tube or microwell plate test formats and several end-point detection methods, most commonly plaque formation on a lawn of rapid-growing mycobacterial cells (65). Detection of MTB in decontaminated sputum by phage amplification has high specificity but moderate and variable sensitivity; a recent meta-analysis showed accuracy is similar to that of smear microscopy. Sensitivity ranged from 21% to 88% and specificity was between 83% and 100% (66). The assay's sensitivity is affected by several factors influencing MTB viability in a specimen, including treatment of patients with anti-tuberculous medication, storage and transport of the specimen and the decontamination method used. The use of milder sputum digestion and decontamination methods that maintain MTB viability could potentially improve the assay's sensitivity, although higher levels of contamination may result.

Phage amplification has been applied to drug susceptibility testing for both first- and second-line drugs (65; 67). However, most data are available for detection of rifampicin resistance from TB isolates; these tests show high sensitivity but slightly lower specificity (68). Data are limited on use of phage amplification for drug susceptibility testing directly from clinical specimens. Sensitivity and specificity varied from 86–100% and 75–100% respectively when applied to smear-positive sputum specimens (66; 67; 70). A substantial proportion of specimens (up to 25%) do not yield an interpretable result when testing is performed directly (70) due to too few plaques in the drug-free control and contamination. An as yet unpublished South African trial suggests that the FASTPlaque assay lacks sufficient

robustness for application in a routine setting. While recent studies discuss development of an effective antimicrobial supplement for contamination control (71), the issue of inadequate plaque numbers on the positive control needs to be addressed. Further studies on other drugs are required to determine optimal drug concentration and incubation times, and to assess performance directly from patient specimens.

Luciferase reporter phages

Luciferase reporter phages (LRPs) are genetically-modified phages containing the *fflux* gene encoding firefly luciferase. This catalyses a reaction producing light in the presence of the luciferin substrate and ATP; light is only produced in the presence of viable mycobacteria. Detection of light released from viable mycobacteria can be achieved by a luminometer or photographic film (72) within two to four days from culture. The luminometer readout is more sensitive and enables quantification of results, while the use of Polaroid photographic film offers a lower-tech approach with lower sensitivity.

Data on use of LRP assays is restricted to detection and drug susceptibility testing from clinical isolates using in-house protocols (76). Differentiation of MTB complex and non-tuberculous mycobacteria (NTM) has been reported by use of p-nitro-α-acetylamino-β-hydroxy propiophenone (NAP), a selective inhibitor of MTB complex, in conjunction with the LRP assay (73).

For detection of resistance to rifampicin and isoniazid from clinical isolates, LRP assays had high sensitivity and specificity ranging from 92-100% and 89–100% for rifampicin, and 86-100% and 95-100% for isoniazid. Lower sensitivity and specificity of LRP assays for detection of streptomycin and ethambutol resistance has been reported (72; 74; 75). Expression of the *fflux* gene in a temperate phage is reported to result in increased light output (and thus increased sensitivity) compared with the conventional lytic phage constructs (76). The performance of LRP assays directly from sputum specimens requires further study.

In conclusion, although phage-based tests have the potential for being low-cost and relatively low-tech methods, test performance issues remain particularly when applied directly to clinical specimens. In their current format, phage-based tests are suitable for referral laboratory settings since they require biosafety cabinets, centrifuges and skilled laboratory technicians. The mycobacteriophages used in the assays are not specific for MTB complex and can also replicate in NTM. This may not be a problem in high–TB-incidence settings, but may lead to reduced specificity in settings with a higher incidence of NTMs. The need for fresh clinical specimens to enable efficient phage infection and minimize contamination is a

disadvantage of these methods, particularly in settings with poor specimen transport infrastructure. While a number of other promising detection methods have been reported, evidence has yet to be provided on their application to clinical specimens (77).

**Annex 9
Nose technologies**

Nose technologies

The first documented reports that tuberculosis could be diagnosed by odours emitted from specimens originate from ancient Greece, where Hippocrates suggested that *"persons [are] affected with phthisis, if the sputa which they cough up have a heavy smell when poured upon coals."* This technique was also favoured in Roman times and by traditional Chinese healers. Use of the human nose to detect TB is now discouraged, but recent reports suggest that African pouch rats can be trained to identify TB cases by smelling their sputum (78). The animals are conditioned to cultured bacteria using a system of food rewards. For the diagnostic test, sputum samples are heat-treated prior to exposure to an individual animal. The response to a positive signal is scratching by the rat of the surface close to the specimen, which is observed by a technician. Preliminary studies with insects such as bees and wasps suggest that they may also be trained; however no data is available on the clinical utility of such an approach. Concerns regarding the transferability of biological olfactory sensing systems and the perceived difficulty of quality-assured training of animals have limited investment in this area.

A number of instrument-based technologies have been developed to detect volatile organic compounds (VOCs), an area of R&D stimulated in recent years by the need to detect volatile bioterrorism agents. Manufactured devices to detect volatile compounds may be divided into electronic noses (E-noses) that attempt to mimic biological olfactory systems and those of a more analytical nature, sometimes referred to as artificial noses to differentiate them from the E-noses. E-noses vary widely in design and format but are based on the electrical signalling resulting from the interaction of VOCs with polymers, usually coated onto a probe or a series of panels. An E-nose that has been applied to TB is the Bloodhound Sensor (Insensive Ltd) that presents 14 conducting polymers based on polyaniline on a series of panels. When volatile compounds interact with these polymer surfaces, their electrical resistance changes. Following exposure to a series of known TB-positive and -negative samples, data collected are analysed via neural networking to establish a pattern that is predictive of disease. This is then programmed into the instrument. The device is portable and, once the test has been established, may be used to screen samples in a few minutes. Although very sensitive, the instrument is vulnerable to interference from water vapour and background odours in the environment. Other drawbacks of E-nose technology are the polymers' instability over time and variation between instruments requiring calibration.

Two approaches may be used during the design of a VOC-based diagnostic test. The first is whole gas analysis, where the differences between TB-positive and TB-negative samples are established. Proof of principle for TB diagnosis from breath has been reported using pattern recognition data and fuzzy logic analysis by Menssana Research. The second approach is identification and measurement of specific volatile biomarkers that are predictive of tuberculosis. Identification and validation of such markers is being undertaken in several laboratories; as yet no published data support the clinical utility of specific VOCs for TB diagnosis.

Analytical instruments that may be applied to VOC analysis range from traditional mass spectrometers (MS; often linked to gas chromatographic instruments) to portable devices designed for military settings or for detection of contraband. Samples tested may be breath or the gases above clinical samples (head space vapours). Direct testing of VOCs can be extremely rapid as sample processing is not required. Gases such as helium are often needed, and may be supplied in small cassettes to provide portability. Traditional non-portable MS instruments are costly and require high maintenance as well as transport of the patient samples to a specialist laboratory. Portability may be achieved by pumping gaseous samples through tubes containing absorptive resins; once in the laboratory the VOCs are desorbed by heating, enabling their analysis. This approach may also be used as a means of concentrating a particular VOC.

There are a number of portable or semi-portable devices available for VOC analysis. One such technology under investigation is differential mobility spectrometry (DMS) including the MicroDMxTM sensor developed by Sionex. This device exposes samples to a field of gaseous ions which assist separation of compounds at the molecular level. Gas chromatography (GC) is also being investigated using miniaturized columns to allow separation of compounds in minutes. The choice of detector is critical to enable high sensitivity. Surface acoustic wave (SAW) sensors have demonstrated high sensitivities where molecules hitting the crystal sensor depress the frequency of the acoustic wave. An alternative approach is to combine the analytic power of GC with specific biosensing technology. However, there are as yet no reports of the utility of portable field instruments for diagnosing TB from breath or head space of clinical specimens.

Annex 10
References and glossary

References and glossary

1. Keeler E et al. (2006). Reducing the global burden of tuberculosis: the contribution of improved diagnostics. *Nature,* 444 suppl 1:49–57. PMID:17159894 doi:10.1038/nature05446

2. Mase S et al. (2007). Yield of serial sputum specimen examinations in the diagnosis of pulmonary tuberculosis: a systematic review. *International Journal of Tuberculosis and Lung Disease,* 11:485–495. PMID:17439669

3. Chang KC et al. (2008). Supervised and induced sputum among patients with smear-negative pulmonary tuberculosis. *European Respiratory Journal*, 31:1085–1090.

4. Khan M et al. (2007). Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: a pragmatic randomised controlled trial. *Lancet,* 369:1955–1960. PMID:17560448 doi:10.1016/S0140-6736(07)60916-7

5. Mohan A, Sharma S (2008). Fibreoptic bronchoscopy in the diagnosis of sputum smear-negative pulmonary tuberculosis: current status. *Indian Journal of Chest Diseases and Allied Sciences,* 50:67–78. PMID:18610691

6. Hirao S et al. (2007). Same-day smears in the diagnosis of tuberculosis. *Tropical Medicine and International Health,* 12:1459–1463. PMID:18076552

7. Cambanis A et al. (2006). A one-day method for the diagnosis of pulmonary tuberculosis in rural Ethiopia. *International Journal of Tuberculosis and Lung Disease,* 10:230–232. PMID:16499267

8. Steingart K et al. (2006). Sputum processing methods to improve the sensitivity of smear microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases,* 6:664–674. PMID:17008175 doi:10.1016/S1473-3099(06)70602-8

9. Bonnet M et al. (2008). Bleach sedimentation: an opportunity to optimize smear microscopy for tuberculosis diagnosis in settings of high prevalence of HIV. *Clinical Infectious Diseases,* 46:1710–1716. PMID:18444789 doi:10.1086/587891

10. den Hertog et al.(2009) "Buoyant density of Mycobacterium tuberculosis: implications for sputum processing" International Journal of Tuberculosis Lung Disease 13(4):466-71.

11. Ratnam S, March S (1986). Effect of relative centrifugal force and centrifugation time on sedimentation of mycobacteria in clinical specimens. *Journal of Clinical Microbiology,* 23:582–585. PMID:3082926

12. Steingart K et al. (2006). Fluorescence versus conventional sputum smear

microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases,* 6:570–581. PMID:16931408 doi:10.1016/S1473-3099(06)70578-3

13. Anthony R et al. (2006). Light emitting diodesfor auramine O fluorescence microscopic screening of *Mycobacterium tuberculosis*. *International Journal of Tuberculosis and Lung Disease,* 10:1060–1062. PMID:16964802

14. Hanscheid T (2008). The future looks bright: low-cost fluorescent microscopes for detection of *Mycobacterium tuberculosis* and *Coccidiae*. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 102:520–521. PMID:18405928 doi:10.1016/j.trstmh.2008.02.020

15. Van Deun A et al. (2008). Performance and acceptability of the FluoLED easy module for tuberculosis fluorescence microscopy. *International Journal of Tuberculosis and Lung Disease,* 12:1009–1014. PMID:18713497

16. Cambanis A et al. (2007). Investing time in microscopy: an opportunity to optimise smear-based case detection of tuberculosis. *International Journal of Tuberculosis and Lung Disease,* 11:40–45. PMID:17217128

17. Krambovitis E et al. (1984). Rapid diagnosis of tuberculous meningitis by latex particle agglutination. *Lancet,* 324:1229–1231. PMID:6150274 doi:10.1016/S0140-6736(84)92792-2

18. Krambovitis E, Harris M, Hughes D (1986). Improved serodiagnosis of tuberculosis using two assay test. *Journal of Clinical Pathology,* 39:779–785. PMID:3734115 doi:10.1136/jcp.39.7.779

19. Chan J et al. (1991). Lipoarabinomannan, a possible virulence factor involved in persistence of *Mycobacterium tuberculosis* within macrophages. *Infection and Immunity,* 59:1755–1761. PMID:1850379

20. Hunter S, Gaylord H, Brennan P (1986). Structure and antigenicity of the phosphorylated lipopolysaccharide antigens from the leprosy and tubercle bacilli. *Journal of Biological Chemistry,* 261:12345–12351. PMID:3091602

21. Hunter S et al. (1999). Isolation and characterization of the highly immunogenic cell wall-associated protein of *Mycobacterium leprae*. *Journal of Immunology,* 142:2864–2872.

22. Boehme C et al. (2005). Detection of mycobacterial lipoarabinomannan with an antigen-capture ELISA in unprocessed urine of Tanzanian patients with suspected tuberculosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 99:893–900. PMID:16139316 doi:10.1016/j.trstmh.2005.04.014

23. Hamasur B et al. (2001). Rapid diagnosis of tuberculosis by detection of mycobacterial lipoarabinomannan in urine. *Journal of Microbiological Methods,* 45:41–52. PMID:11295196 doi:10.1016/S0167-7012(01)00239-1

24. Pereira Arias-Bouda L et al. (2000). Development of antigen detection assay for

diagnosis of tuberculosis using sputum samples. *Journal of Clinical Microbiology,* 38:2278–2283. PMID:10834989

25. Tessema TA et al. (2001). Diagnostic evaluation of urinary lipoarabinomannan at an Ethiopian tuberculosis centre. *Scandinavian Journal of Infectious Diseases,* 33:279–284. PMID:11345220 doi:10.1080/003655401300077306

26. Tessema TA et al. (2002). Clinical and radiological features in relation to urinary excretion of lipoarabinomannan in Ethiopian tuberculosis patients. *Scandinavian Journal of Infectious Diseases,* 34:167–171. PMID:12030387 doi:10.1080/00365540110077254

27. Khoo K, Tang J, Chatterjee D (2001). Variation in mannose-capped terminal arabinan motifs of lipoarabinomannans from clinical isolates of *Mycobacterium tuberculosis* and *Mycobacterium avium* complex. *Journal of Biological Chemistry,* 276:3863–3871. PMID:11073941 doi:10.1074/jbc.M004010200

28. Kashino S et al. (2008). Identification and characterization of *Mycobacterium tuberculosis* antigens in urine of patients with active pulmonary tuberculosis: an innovative and alternative approach of antigen discovery of useful microbial molecules. *Clinical and Experimental Immunology,* 153:56–62. PMID:18460016 doi:10.1111/j.1365-2249.2008.03672.x

29. Napolitano D et al. (2008). Identification of *Mycobacterium tuberculosis* ornithine carboamyltransferase in urine as a possible molecular marker of active pulmonary tuberculosis. *Clinical and Vaccine Immunology,* 15:638–643. PMID:18305107 doi:10.1128/CVI.00010-08

30. Boehme C et al. (2005). Detection of mycobacterial lipoarabinomannan with an antigen-capture ELISA in unprocessed urine of Tanzanian patients with suspected tuberculosis. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 99:893–900. PMID:16139316 doi:10.1016/j.trstmh.2005.04.014

31. Caminero J et al. (1993). Diagnosis of pleural tuberculosis by detection of specific IgG anti-antigen 60 in serum and pleural fluid. *Respiration: International Review of Thoracic Diseases,* 60:58–62. PMID:8469821 doi:10.1159/000196174

32. Bothamley G (1995). Serological diagnosis of tuberculosis. *European Respiratory Journal,* 8:676–688.

33. Gennaro M (2000). Immunologic diagnosis of tuberculosis. *Clinical Infectious Diseases,* 30 suppl 3:S243–246. PMID:10875791 doi:10.1086/313868

34. Daniels T (1996). Immunodiagnosis of tuberculosis. In: Rom W, Garay S, eds. *Tuberculosis*. Boston, Little, Brown, 223–231.

35. Pottumarthy S, Weels V, Morris A (2000). A comparison of seven tests for serological diagnosis of tuberculosis. *Journal of Clinical Microbiology,* 38:2227–2231. PMID:10834981

36. World Health Organization. (2008). Laboratory-based evaluation of 19 commercially available rapid diagnostic tests for tuberculosis. *Diagnostics Evaluation Series*. Geneva, WHO Press,1–70.</bok>

37. Steingart K, et al. (2007). Commercial serological antibody detection tests for the diagnosis of pulmonary tuberculosis: a systematic review. *PLoS Medicine,* 4:1–20.

38. Steingart K et al. (2007). A systematic review of commercial serological antibody detection tests for the diagnosis of extrapulmonary tuberculosis. *Thorax,* 62:911–918. PMID:17675320

39. Samanich K, Belisle JT, Laal S (2001). Homogeneity of antibody responses in tuberculosis patients. *Infection and Immunity,* 69:4600–4609. PMID:11402004 doi:10.1128/IAI.69.7.4600-4609.2001

40. Sartain M et al. (2006). Disease state differentiation and identification of tuberculosis biomarkers via native antigen profiling. *Molecular and Cellular Proteomics*, 5:2102–2113. PMID:16899542 doi:10.1074/mcp.M600089-MCP200

41. Weldingh K et al. (2005). Assessing the serodiagnostic potential of 35 *Mycobacterium tuberculosis* proteins and identification of four novel serological antigens. *Journal of Clinical Microbiology,* 43:57–65. PMID:15634951 doi:10.1128/JCM.43.1.57-65.2005.

42. Bentrup K, Russell D (2001). Mycobacterial persistence: adaptation to a changing environment. *Trends in Microbiology,* 9:597–605. PMID:11728873 doi:10.1016/S0966-842X(01)02238-7

43. Mehrotra J, Bishai W (2001). Regulation of virulence genes in *Mycobacterium tuberculosis*. *International Journal of Medical Microbiology,* 291:171–182. PMID:11437339 doi:10.1078/1438-4221-00113

44. Wanchu A et al. (2008). Biomarkers for clinical and incipient tuberculosis: performance in a TB-endemic country. *PLoS One,* 3:e2071–2079. PMID:18446210 doi:10.1371/journal.pone.0002071

45. Samanich K et al. (2000). Serodiagnostic potential of culture filtrate antigens of *Mycobacterium tuberculosis. Clinical and Diagnostic Laboratory Immunology,* 7:662–668. PMID:10882669

46. Zanetti S et al. (2005). Patients with pulmonary tuberculosis develop a strong humoral response against methylated heparin-binding hemagglutinin. *Clinical and Diagnostic Laboratory Immunology,* 12:1135–1138. PMID:16148186

47. Laal S et al. (1997). Surrogate marker of preclinical tuberculosis in human immunodeficiency virus infection: antibodies to an 88 kDa secreted antigen of *Mycobacterium tuberculosis. Journal of Infectious Diseases,* 176:133–143. PMID:9207359 doi:10.1086/514015

48. Singh K et al. (2005). Antigens of *Mycobacterium tuberculosis* recognized by antibodies during incipient, subclinical tuberculosis. *Clinical and Diagnostic Laboratory Immunology,* 12:354–358. PMID:15699433

49. Singh K et al. (2005). Immunogenicity of *Mycobacterium tuberculosis* PPE55 (Rv3347c) protein during incipient and clinical tuberculosis. *Infection and Immunity,* 73:5004–5014. PMID:16041015 doi:10.1128/IAI.73.8.5004-5014.2005

50. Gennaro M et al. (2007). Antibody markers of incident tuberculosis among HIV-infected adults in the USA: a historical prospective study. *International Journal of Tuberculosis and Lung Disease,* 11:624–631. PMID:17519093

51. Khan I et al. (2008). Profiling antibodies to *Mycobacterium tuberculosis* by multiplex microbead suspension arrays for serodiagnosis of tuberculosis. *Clinical and Vaccine Immunology,* 15:433–438. PMID:18077619 doi:10.1128/CVI.00354-07

52. Houghton R et al. (2002). Use of multiepitope polyproteins in serodiagnosis of active tuberculosis. *Clinical and Diagnostic Laboratory Immunology,* 9:883–891. PMID:12093690

53. Farhat M (2006). False-positive tuberculin skin tests: what is the absolute effect of BCG and non-tuberculous mycobacteria? *International Journal of Tuberculosis and Lung Disease,* 10:1192–1204. PMID:17131776

54. Rothel JS, Andersen P (2005). Diagnosis of latent *Mycobacterium tuberculosis* infection: is the demise of the mantoux test imminent? *Expert Review of Anti-Infective Therapy,* 3:981–993. PMID:16307510 doi:10.1586/14787210.3.6.981

55. Lalvani A (2007). Diagnosing tuberculosis infection in the 21st century: new tools to tackle an old enemy. *Chest,* 131:1898–1906. PMID:17565023 doi:10.1378/chest.06-2471

56. Pai M, Zwerling A, Menzies D (2008). Systematic review: T-cell-based assays for the diagnosis of latent tuberculosis infection: an update. *Annals of Internal Medicine,* 149:177–184. PMID:18593687

57. 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. *Annals of Internal Medicine,* 146:340–354. PMID:17339619

58. Dinnes J et al. (2007). A systematic review of rapid diagnostic tests for the detection of tuberculosis infection. *Health Technology Assessment,* 11(3):1–196. http://www.ncchta.org/pdfexecs/summ1103.pdf PMID:17266837

59. Menzies D et al. (2008). Thinking in three dimensions: a web-based algorithm to aid the interpretation of tuberculin skin test results. *International Journal of Tuberculosis and Lung Disease,* 12:498–505. PMID:18419884

60. Pai M et al. (2007). T-cell assays for the diagnosis of latent tuberculosis infection: moving the research agenda forward. *Lancet Infectious Diseases,* 7:428–438. PMID:17521596 doi:10.1016/S1473-3099(07)70086-5

61. Pai M, Ramsay A, O'Brien R (2008). Evidence-based tuberculosis diagnosis. *PLoS Medicine,* 5:e156. PMID:18651788 doi:10.1371/journal.pmed.0050156

62. Andersen P et al. (2007). The prognosis of latent tuberculosis: can disease be predicted? *Trends in Molecular Medicine,* 13:175–182. PMID:17418641 doi:10.1016/j.molmed.2007.03.004

63. Nahid P, Pai M, Hopewell P (2006). Advances in the diagnosis and treatment of tuberculosis. *Proceedings American Thoracic Society,* 3:103–110. PMID:16493157 doi:10.1513/pats.200511-119JH

64. Noordhoek GT, van Ende JDA, Kolk AHJ (1993). Questionable reliability of the polymerase chain reaction in the detection of *Mycobacterium tuberculosis*. *New England Journal of Medicine,* 329:2036. PMID:8247086 doi:10.1056/NEJM199312303292713

65. Wilson S et al. (1997). Evaluation of a new rapid bacteriophage-based method for the drug susceptibility testing of *Mycobacterium tuberculosis. Nature Medicine,* 3:465–468. PMID:9095184 doi:10.1038/nm0497-465

66. Kalantri S et al. (2005). Bacteriophage-based tests for the detection of *Mycobacterium tuberculosis* in clinical specimens: a systematic review and meta-analysis. *BMC Infectious Diseases,* 5:59. PMID:16022735 doi:10.1186/1471-2334-5-59

67. Eltringham I et al. (1999). Evaluation of reverse transcription-PCR and a bacteriophage-based assay for rapid phenotypic detection of rifampin resistance in clinical isolates of *Mycobacterium tuberculosis. Journal of Clinical Microbiology,* 37:3524–3527. PMID:10523546

68. Pai M et al. (2005). Bacteriophage-based assays for the rapid detection of rifampicin resistance in *Mycobacterium tuberculosis*: a meta-analysis. *Journal of Infection,* 51:175–187. PMID:16002146 doi:10.1016/j.jinf.2005.05.017

69. Albert H et al. (2007). Evaluation of a rapid screening test for rifampicin resistance in re-treatment tuberculosis patients in the Eastern Cape. *South African Medical Journal,* 97:858–863. PMID:17985057

70. Acuna-Villaorduna C et al. (2008). Cost-effectiveness analysis of introduction of rapid, alternative methods to identify multidrug-resistant tuberculosis in middle-income countries. *Clinical Infectious Diseases,* 47:487–495 PMID:18636955 doi:10.1086/590010

71. Albert H et al. (2007) Development of an antimicrobial formulation for control of specimen-related contamination in phage-based diagnostic testing for tuberculosis. Journal of Applied Microbiology,103: 892-899.

72. Hazbon M et al. (2003). Photographic and luminometric detection of luciferase reporter phages for drug susceptibility testing of clinical *Mycobacterium tuberculosis* isolates. *Journal of Clinical Microbiology,* 41:4865–4869. PMID:14532245 doi:10.1128/JCM.41.10.4865-4869.2003

73. Riska P et al. (1997). Specific identification of *Mycobacterium tuberculosis* with the luciferase reporter mycobacteriophage: use of p-nitro-α-acetylamino-β-hydroxy propiophenone. *Journal of Clinical Microbiology,* 35:3225–3231. PMID:9399524

74. Banaiee N et al. (2001). Luciferase reporter mycobacteriophages for detection and antibiotic susceptibility testing of *Mycobacterium tuberculosis* in Mexico. *Journal of Clinical Microbiology,* 39:3883–3888. PMID:11682502 doi:10.1128/JCM.39.11.3883-3888.2001

75. Banaiee N et al. (2008). Evaluation of a semi-automated reported phage assay for susceptibility testing of *Mycobacterium tuberculosis* isolates in South Africa. *Tuberculosis*, 88:64–68. PMID:17980664 doi:10.1016/j.tube.2007.08.006

76. Dusthackeer A et al. (2008). Construction and evaluation of luciferase reporter phages for the detection of active and non- replicating tubercle bacilli. *Journal of Microbiological Methods,* 73:18–25. PMID:18272245 doi:10.1016/j.mimet.2008.01.005

77. Banaiee N et al. (2003). Rapid identification and susceptibility testing of *Mycobacterium tuberculosis* from MGIT cultures with luciferase reporter mycobacteriophages. *Journal of Medical Microbiology,* 52:557–561. PMID:12808076 doi:10.1099/jmm.0.05149-0

78. Weetjens BJ et al. (2009). African pouched rats for the detection of pulmonary tuberculosis in sputum samples. *International Journal of Tuberculosis and Lung Disease*, 13(6):737–743.

Glossary

Active tuberculosis is clinically active disease caused by *M. tuberculosis*.

Clinical trial plans outline the studies required for feasibility, development, final product performance claims and individual country registration/approval. These plans should include the number of tests needed, potential trial sites, compliance with international clinical trial standards and sourcing of specimens.

A concept definition document indicates the technologies required to meet the customer needs detailed in the user requirements document. It should include initial assessment of geographical region, manufacturing cost estimates, selling price, unit sales, cost to serve and support customers, profit projections and project timeline.

The user requirements document covers the requirements for the product such as workflow, performance and design. Requirements are typically tiered (desired and minimum). This document may include type of specimen, laboratory infrastructure and biosafety needed, acceptable level of "uninterpretables" or invalid test results, shelf life from date of manufacture, environmental conditions for storage and shipment, packaging material, kit size, disposal of used test devices/reagents, number of workflow steps and standard manufacturing cost targets. If a reader or instrument is used, additional documentation requirements include reliability, electrical requirements, weight, environmental operating conditions and instructions for use of controls/calibrators.

Customer support plan should include: procedures for clinical trials; product training material and personnel to provide training; instrument installation, maintenance and repair (including a service training programme); technical inquiries; customer complaints; and lot-specific change information if required.

Design history files document the activities related to the development process of a device. This can be a collection of the actual documents generated in the product development process or an index of documents and their storage locations. These documents should include product labelling as well as reports on risk management, reliability, clinical trials, stability and validation.

The feasibility target specification should consist of all the qualities of the selected technology that are intended to be assessed through feasibility studies. This specification is the blueprint for the design of feasibility-phase trials.

GMP-IVD due diligence inspection reports describe whether skills, expertise, technical equipment and facility of the manufacturing partner comply with good manufacturing practice standards.

Infection is the spread of *M. tuberculosis* bacilli through the host.

Incoming quality control (IQC) is a procedure developed by manufacturers that describes the quality control to be performed by trial sites upon receipt of new reagent shipments or before starting to use new lots. Release documents should be signed off by the manufacturer and the study coordinator before reagents can be used.

Latency is where living (but inactive) *M. tuberculosis* bacilli are present in the body but the disease is not clinically active.

Marketing (launch) plans detail where, how and by whom the product will be sold. The plan should include a unit sales forecast, marketing partners, advantages over existing products and pricing/distribution considerations.

Outgoing quality control (OQC) is a procedure developed by the manufacturer that describes the quality control to be performed by the manufacturer upon production of new lots. Release documents should be signed off by the manufacturer and the study coordinator before reagents can be used.

Product requirement document is designed to demonstrate what a product should do and how it should work. It should include the following elements: purpose and scope, market assessment and target demographics, product overview, requirements (functional, usability, technical, environmental, manufacturing, support and interaction), constraints, workflow and evaluation plan.

Product specification documents define product performance claims and reflect the compromise between the desired features in the user requirements document and the achievable features within the chosen technology, time frame and budget constraints. These documents should describe intended use, product positioning, minimum and optimum performance characteristics of the final product, precise description of the product with components and eventual modifications from a previous version.

Progression is the development of active tuberculosis disease from a state of latency.

A regulatory plan details where the product will be manufactured and where it will be sold. It should include documents necessary to register or obtain regulatory clearance of a product in target countries and should list relevant regulations for manufacturing and development to be complied with.