Priorities in Operational Research to Improve Tuberculosis Care and Control

Stop TB Partnership

Investing in our future
The Global Fund
To Fight AIDS, Tuberculosis and Malaria

World Health Organization

WHO Library Cataloguing-in-Publication Data:

Priorities in operational research to improve tuberculosis care and control.

1.Tuberculosis - prevention and control. 2.Tuberculosis - therapy. 3.Tuberculosis, Pulmonary - prevention and control. 4.Operations research. I.World Health Organization. II.Stop TB Partnership. III.Global Fund to Fight AIDS, Tuberculosis and Malaria.

ISBN 978 92 4 154825 0 (NLM classification: WF 300)

Printed in France

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The Partnership acknowledges with gratitude everyone who has contributed to this document:

Overall coordination and finalization of the document
Christian Lienhardt

Coordination of the development of specific chapters of the document
Andrew Ramsay & Afranio Kristski (Improving Access, Screening and Diagnosis of TB), James Newell (Developing sustainable collaboration with all care providers for TB care and control); Christian Lienhardt (Treatment of Drug-susceptible and M/XDR-TB: optimal access, delivery and community participation; Methods); Jonathan Golub & Alwyn Mwinga (Prevention of TB in HIV-infected patients and joint treatment of TB and HIV); Anthony Harries & Rony Zachariah (Capacity building for operational research); Marieke J. van der Werf, Silke Heumann & Ellen M.H. Mitchell (Community involvement in TB research); Jacqueline Bataringaya & Wilson Lo (Global Fund investment in TB operational research); S. Bertel Squire, Brian Faragher & Kerry Millington (Methodology and protocol templates).

Core writing team
Jacqueline Bataringaya, Imelda Bates, Christian Lienhardt, Gillian Mann, Kerry Millington, S. Bertel Squire.

Scientific Editorial Committee
Jeremiah Chakaya, Frank Cobelens, Liz Corbett, Jonathan Golub, Anthony Harries, Afriano Kristki, Alwyn Mwinga, James Newell, Andrew Ramsay, Mukund Uplekar, Rony Zachariah, Alimuddin Zumla

Additional contributions
Rifat Atun, Jaap Broekmans, Jeremiah Chakaya, Frank Cobelens, Liz Corbett, Anne Detjen, Jennifer Furin, Giuliano Gargioni, Haileyesus Getahun, Jonathan Golub, Anthony Harries, Ian Harper, Shabbar Jaffar, Achilles Katamba, Jane Kengeya Kayondo, Salmaan Keshavjee, Afranio Kristki, Knut Lonnroth, Gillian Mann, Neil Martinson, Charles Mgone, Alwyn Mwinga, James Newell, Ikushi Onozaki, Philip Chukwuka Onyebujoh, Andrew Ramsay, Savanand Sahu, Felix M.L. Salaniponi, Delphine Sculier, Gurmit Singh, Soumya Swaminathan, Zafar Ullah, Mukund Uplekar, Maarten van Cleeff, Marieke van der Werf, Sanne Christine van Kampen, Jay Varma, Nguyen Viet Nhung, Diana Weil, Claire Wingfield, Rony Zachariah, Matteo Zignol, Alimuddin Zumla.

Technical editor
Tim France (www.iniscommunication.com)

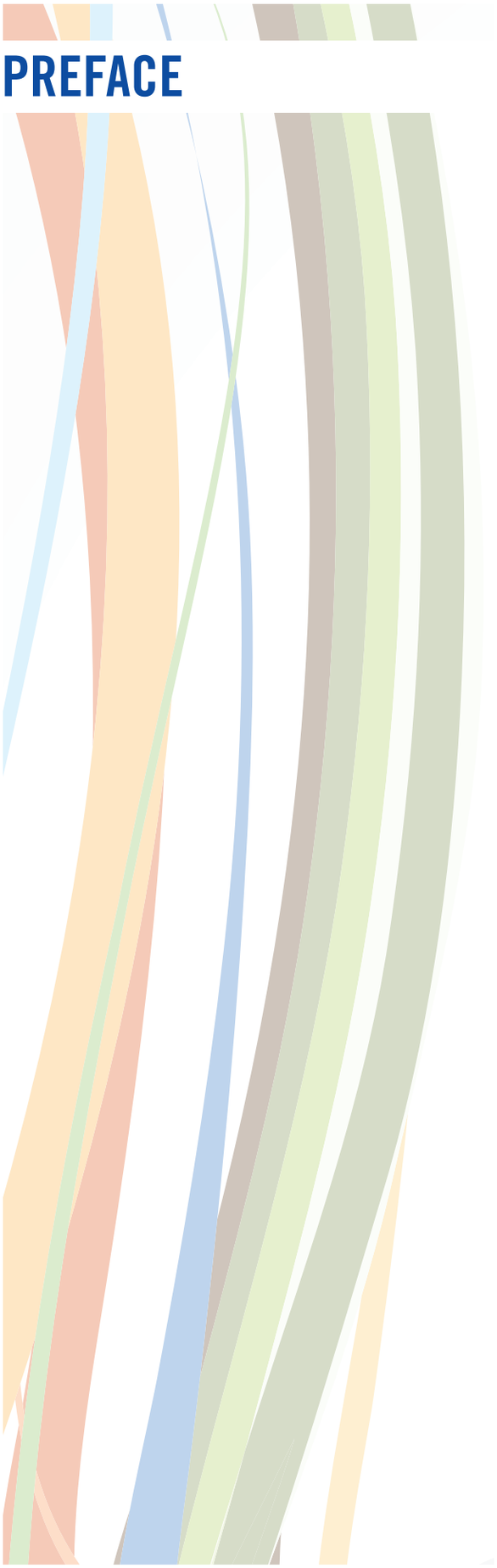
Administrative and secretarial support
Gloria Haselmann

Graphic design
Graphi 4

ABBREVIATIONS

ACF	active case-finding
ART	antiretroviral therapy
CAB	community advisory board
CBPR	community-based participatory research
CDC	Centers for Disease Control (US)
CDR	case detection rate
CI	confidence interval
DFID	Department for International Development (United Kingdom)
DR-TB	drug-resistant TB
DS-TB	drug-sensitive TB
DST	drug susceptibility testing
FBO	faith-based organization
Global Fund	Global Fund to Fight AIDS, Tuberculosis and Malaria
GIS	geographical information systems
GLC	Green Light Committee
HIV	human immunodeficiency virus
IAF	Impact Assessment Framework
IC	infection control
ICC	intra-cluster correlation
IDU	injection drug users
IGRA	interferon gamma release assay
IPT	isoniazid preventive therapy
JATA	Japan Anti-Tuberculosis Association
KNCV	Royal Dutch TB Association
LED	light-emitting diode
LPA	line probe assays
LTBI	latent TB infection
MDG	Millennium Development Goals
MDR-TB	multidrug-resistant TB
MODS	microscopic-observation drug susceptibility
MSF	Médecins sans Frontières
MSM	men who have sex with men
M/ XDR-TB	multi/extensively drug-resistant TB
NAAT	nucleic acid amplification technology
NGO	nongovernmental organization
NORAD	Norwegian Agency for Technical Cooperation and Development
NSA	national situation assessment
NTCP	national TB control programme
PAL	Practical Approach to Lung Health
PCR	polymerase chain reaction
PEPFAR	(United States) President's Emergency Plan for AIDS Relief
PLHIV	people living with HIV
PPM	public–private mix
PRCT	pragmatic randomized controlled trials
RCT	randomized controlled trial
SS+	sputum smear-positive
STAG-TB	Strategic and Technical Advisory Group for TB [WHO]
TAG	Treatment Action Group
TB	tuberculosis
TST	tuberculin skin test
USAID	United States Agency for International Development
WHO	World Health Organization
XDR-TB	extensively drug- resistant TB

PREFACE



Over the past two decades, we have seen impressive gains in the global fight against tuberculosis (TB). An estimated 41 million people have been successfully treated, and 6 million deaths have been averted. Yet there remain serious challenges to reach all people who need quality TB care. Every year as many as 4 million people with TB fail to receive such care, and their illness is never documented, and 400 000 MDRTB cases are not having access to proper diagnosis and treatment.

Operational research – which aims to develop interventions that result in improved policies; better design and implementation of health systems, and more efficient methods of service delivery – is critical to reaching the unreached people who need TB care. It produces evidence that lays the groundwork for improving current strategies and introducing new tools and new partners. For this reason, the *Stop TB Partnership has included a new section on operational research in the Global Plan to Stop TB 2011–2015.*

In 2010, the Stop TB Partnership, the World Health Organization Stop TB Department and the Global Fund to Fight AIDS, Tuberculosis and Malaria jointly organized an expert meeting and workshop on operational research, followed by international consultations. The goal was to identify priority areas in which knowledge gaps hamper optimal implementation of TB control activities. The outcome of these activities formed the basis for this publication.

We have identified five priority areas; and for each of these, we provide a list of the critical questions that must be addressed to improve TB care and control at the community, national, regional and international levels. In addition, for each of the critical questions, we provide a synopsis of a suitable study design and the methods required to identify and test suitable solutions.

In making this material widely available we hope to encourage national TB control programmes and research institutions in countries with a high burden of TB to lead operational research projects that will help them diminish the impact of TB on their populations. All concerned stakeholders, including civil society and affected communities, should participate in the development of operational research agendas. We also have high hopes that international funders will better understand the clear value of operational research and reinforce their commitments to supporting it.

Reaching every person who needs TB care is one of the primary goals of the Stop TB Partnership. There are multiple paths to reaching that point, and they must ultimately converge. Operational researchers should now have what they need to forge those paths. The present publication, driven by the Stop TB Partnership Research Movement, will be the appropriate vehicle for this.

Dr Lucica Ditiu
Executive Secretary
Stop TB Partnership
Geneva

Enabling and promoting research is a key component of the *Stop TB Strategy*, and it should be pursued vigorously. Despite dramatic progress on global TB control in recent decades, many challenges remain. The fight against this ancient scourge must now accelerate. Clearly we need new and better tools for the prevention, diagnosis, treatment and care of TB and its associated conditions and complications. However, that will not be enough. We also need to continue to develop innovative approaches to ensure equitable access to these tools for all who need them. And those approaches need to be adapted and fine-tuned based on local epidemiological and health system context.

Most innovations cannot be translated into effective local action without careful planning and adaptation. Well-planned and conducted operational research, in addition to routine surveillance, is required to assess the national and local epidemiological and health system situation, as well as to evaluate different implementation modalities for locally relevant interventions. There are, however, many barriers for this essential step in the chain of events from basic research to meaningful practice. Countries and national TB programmes often have limited capacity for operational research, and a strategically developed research agenda is often missing. There is, therefore, a need for guidance on which questions to address, how to do it, and how to strengthen the capacity for operational research.

Fostering better and more relevant operational research and ensuring careful evaluations of local experiences will not only help local implementation. It will also greatly assist the development of global policy. As we have learnt over the years, ideas for global solutions often stem from local innovations. Local experience was essential, for example, in the development of global policy on TB/HIV collaboration, engagement of all health-care providers, management of drug-resistant TB, community involvement, and currently on the adoption of new rapid tests for TB and MDR-TB. We learn from countries, and countries learn from each other. Time and again, essential knowledge on the effectiveness, cost-effectiveness, feasibility, unintended consequence, affordability and health systems requirements for the implementation of various tools and approaches have come from high quality operational research.

However, this only works when a good system for operational research prioritization, planning and implementation is in place. We know that many opportunities to test new approaches and to learn from successful local experiences have been missed because of a lack of strategic thinking. This publication will help national programmes, researchers and other stakeholders to identify key questions for operational research, do the needed planning and capacity strengthening, and raise the required resources. Through its dissemination and use we anticipate a further strengthening of the global TB research movement and, ultimately, to better, evidence-based global policy and local practice.

Dr Mario Raviglione
Director
Stop TB Department
World Health Organization
Geneva

Much progress has been made in global efforts to improve TB care and control over the past 20 years, and the incidence and mortality rates of tuberculosis are now declining globally. A substantial increase in funding from both domestic and international sources has enabled these successes. However, these trends need to be sustained and accelerated if the TB related Millennium Development Goals (MDGs) are to be achieved and the burgeoning problems of TB/HIV co-infection and multiple and extensively drug-resistant TB are to be effectively addressed.

2011 marks a critical point for TB control as we enter the last five years of the journey towards the MDGs. The new *Global Fund Strategy for 2012–2016* emphasizes a significant expansion of TB control, more strategic focus of investments where they are most needed, and ambitious scale up of actions to contain and treat MDR-TB.

The Global Fund is committed to increasing uptake of operational research to improve TB control in order to maximize the impact of its investments. In this regard, countries are encouraged to allocate up to 10% of total grant budgets to activities related to monitoring and evaluation including system strengthening, data management, operational research, and programme and impact evaluation.

Priorities in operational research to improve tuberculosis care and control is very timely for building the evidence base for effective implementation of TB programmes. It provides a clear road map of priorities in operational research to help countries improve implementation of TB control activities in critical areas. The priority operational research questions outlined in this report are also aligned with the Stop TB Partnership's Global Plan to Stop TB 2011–2015.

Through a consultative process led by the Stop TB Partnership Research Movement and funding from the Global Fund, five key areas were identified in which knowledge gaps hamper effective implementation of TB control: access, screening and diagnosis of MDR-TB; development of sustainable collaboration with all practitioners; prevention and treatment of TB in people living with HIV; optimal access and delivery of treatment for susceptible and drug-resistant TB; and operational research capacity building. New tools that offer prospects for substantially improving TB control are now available. Operational research is critical in generating evidence on the acceleration of the uptake of new interventions, diagnostic tools and technologies.

This publication is an important addition to the key resources for national TB control programmes to address the technical and structural challenges that impede optimal prevention, detection and treatment of all forms of TB. It will contribute greatly to the conduct of robust operational research, helping to identify major challenges and determine the solutions to address them.

Professor Rifat Atun
Director
Strategy, Performance and Evaluation Cluster
The Global Fund to Fight AIDS, Tuberculosis and Malaria

Chair, Stop TB Partnership Coordinating Board

EXECUTIVE SUMMARY

For the past 15 years, global tuberculosis (TB) control efforts have generated impressive results, and the implementation of the World Health Organization (WHO) *Stop TB Strategy (1995–2008)* is estimated to have cured 36 million people with TB and averted 6 million deaths worldwide. However, despite these important achievements, the burden of TB continues to rise, and current control efforts need to be greatly improved if elimination[a] is ever to be achieved. For this, new highly effective and widely accessible diagnostics, drugs and vaccines are needed to support TB control globally, and technical and structural challenges that impede optimal detection, treatment and prevention of all forms of TB must be overcome. Ensuring the smooth and wide uptake of control tools and the resolution of barriers to TB control is the role of operational research, which helps to identify the solutions that will have a significant impact on case detection and cure rates, and help improve the effectiveness of TB care services.

Operational research is increasingly recognized as an essential element of global TB control. In its broad sense, operational research covers a large spectrum of activities, from local setting-oriented research to improve TB control programme performance, to international policy-guiding research, including the field evaluation of new interventions to improve TB control. At the national level, an enabling environment for conducting operational research is key to achieving optimal TB control performance, and managers should be aware of the benefits of building research capacities through collaboration with research institutions, universities and nongovernmental organizations. At the international level, a robust evidence base is increasingly required for guiding policy-making (including the use of systematic reviews and GRADE[b] evaluation), often relying on operational research projects that give rise to changes in international policy.

New funding opportunities are emerging to improve the use of existing control technologies (e.g. increased case-finding, simplified treatment monitoring, etc.), and to evaluate the impact of introducing new tools in various health systems and epidemiological settings. In this context, the TB Research Movement of the Stop TB Partnership and the Global Fund to fight AIDS, Tuberculosis and Malaria jointly initiated a process to identify ways to promote rational operational research for TB control. This included an expert meeting and a workshop, accompanied by wide stakeholder consultations. The objective was to critically address the increasing need for improved and rational operational research in TB control, and identify key areas where evidence was lacking for proper implementation of new and existing technologies, as well as novel service delivery models.

As a result of this process, carried out in 2010, five key areas were identified in which knowledge gaps hamper optimal implementation of TB control activities. These are:

1. Access, screening and diagnosis of TB;
2. Sustainable collaboration with all care-providers for TB control;
3. Prevention of TB in people living with HIV, and joint treatment of HIV and TB;
4. Access to and delivery of treatment for drug-susceptible and M/XDR-TB;
5. Capacity-building for operational research.

In each of these five areas, *critical and outstanding questions* to improve TB care and control at the national and international levels were identified.

The **objective** of this publication is to help programme managers, consultants and researchers who intend to conduct TB-related operational research to identify the appropriate methods to be used according to the questions that are being addressed, and prepare for grant applications to donors (including the Global Fund) for operational research support.

- In the ***first part*** of the publication, the five above areas are reviewed, with identification of the *priority questions* to be addressed and their rationales.

[a] Defined in the Stop TB Partnership *Global Plan to Stop TB 2011–2015* as ≤ 1 TB case per million population per year.

[b] See: Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group: http://www.gradeworkinggroup.org

- In the ***second part*** (the annexes), after a broad description of the main study designs used in operational research, the methods of the suggested research to be conducted to address each question are succinctly described, using the following standard *synopsis*:
    - objective(s)
    - design outline
    - setting/study population
    - methods: recruitment of subjects (eligibility criteria); intervention (as appropriate)
    - expected endpoints
    - analysis
    - guidance for sample size calculation (and estimated number of participants)
    - expected duration/timeline
    - suitable scale
    - estimated budget range.

The main outstanding questions in the five priority areas are listed thereafter.

1. Improving access, screening and diagnosis of TB

TB diagnosis in most endemic countries still relies mainly on direct sputum smear microscopy, and the diagnostic services for drug-resistant TB (DR-TB) are based on complex technologies that require sophisticated, biosafe laboratories with highly-trained staff. New tools are becoming available and, since 2007, WHO has endorsed the use of over 10 new TB diagnostic tools (technologies or approaches) that, if used wisely, could improve TB control considerably. In 2010, WHO endorsed a new automated real-time nucleic acid amplification technology (NAAT) for rapid and simultaneous detection of TB and rifampicin resistance (the Xpert MTB/RIF system) that offers the prospects of drastically improving the diagnosis of active TB and MDR-TB. In general, there is insufficient evidence available to determine which package of current and newly-developed diagnostic tests would work best in a given set of circumstances, and there is as yet little guidance to countries on what new diagnostic tools, or combinations of tools, should be implemented in particular epidemiological/ health system settings or for different risk groups, and at what level of the health service it should be done. Operational research with collection of multiple country experiences should be the starting point to better guide diagnostic scale-up, and respond to the current outstanding questions for optimization of TB diagnosis, including:

i. *how to improve access to TB diagnosis?*
ii. *how to improve screening of patients and high-risk groups?*
iii. *how to use the introduction of new tools to improve service delivery practices?*
iv. *how to improve active TB case-finding?*
v. *how to build accessible, effective and efficient diagnostic services with new diagnostic tools?*

2. Developing sustainable collaboration with all care-providers for TB care and control

In many countries, a significant proportion of TB suspects and cases, including those from poor and vulnerable populations, present themselves to a range of public or private care-providers that are not linked to national TB control programmes. Evidence shows that TB diagnosis and treatment practices of many non-programme care-providers are inappropriate and that care-seeking from diverse care-providers hampers access to quality TB care, causes delays in TB diagnosis and imposes financial burden on patients. Several 'Public–Private Mix' (PPM) projects have demonstrated the feasibility, effectiveness, cost-effectiveness and scalability of engaging non-programme care-providers in TB care and control in diverse country settings. To be further scaled up, the following outstanding questions need to be addressed for optimization of collaboration with all practitioners:

i. *How to improve and scale up existing approaches to engaging all care-providers?*
ii. *How to measure the contribution of different provider groups to TB care and control?*
iii. *How to encourage involvement of as yet unengaged providers?*
iv. *How to encourage involvement of the non-public sector in MDR-TB management and TB/HIV collaborative activities?*

v. *How to develop and assess responses to changing involvement of diverse providers in TB care and control?*

vi. *How to encourage introduction of regulatory approaches to collaborating care-providers?*

3. Prevention of TB in people living with HIV (PLHIV) and joint treatment of TB and HIV

Optimal control of TB in high-HIV burden areas requires implementation of collaborative TB/HIV interventions through a sound policy and programme environment that gives due consideration to the local context, the respective epidemiology of TB and HIV, as well as the health system infrastructure that determines service delivery models. System-wide differences between HIV and TB care-providers and stakeholders – and operational difficulties for providing effective and appropriate interventions – have contributed to sub-optimal implementation and scale-up of collaborative activities. It is therefore important to identify measures that would facilitate wider implementation and scaling up of collaborative TB/HIV interventions through effective service delivery models, including community-based interventions. Collaborative TB/HIV interventions include the prevention of TB in PLHIV, the joint treatment of TB and HIV in people dually infected, and improved infection control and prevention. Operational research is needed to optimize prevention and treatment of TB in PLHIV and address the several barriers that may occur at the level of screening, diagnosis, treatment and prevention. Particularly in high-burden countries, the following outstanding operational questions need to be addressed for improved TB/HIV core group activities:

i. *What are the barriers to TB diagnosis, and how to overcome these barriers?*

ii. *What are the barriers to initiation of isoniazid preventive therapy?*

iii. *What are the barriers to optimal combined TB/HIV diagnosis and treatment, and what are the optimal models for joint TB and HIV care activities?*

4. Treatment of drug-susceptible and M/XDR-TB: optimal access, delivery and community participation

Access to health care is the cornerstone of TB control programmes, as all diagnosed cases must receive a full course of treatment. This includes establishing effective treatment as well as effective strategies to support the process of care from detection of disease through to the completion of appropriate treatment. Limited access and poor adherence to treatment remain major obstacles in the global fight against TB. This is particularly true for MDR-TB, as only a small fraction of the tens of thousands of diagnosed patients are receiving appropriate care. Operational research is thus critically needed to improve access to treatment for drug-sensitive (DS-) and drug-resistant (DR-TB) patients, and ensuring that appropriate support is being offered to patients to ensure adherence and address adverse treatment effects, taking into consideration the particular needs of combined treatment whenever needed (antiretrovirals, diabetes, etc). The following outstanding questions need to be addressed for optimization of treatment of DS- and DR-TB:

i. *What are the reporting gaps and deficiencies in first-line management of TB cases?*

ii. *How to address these deficiencies and improve management of drug-sensitive TB?*

iii. *What are the drivers of drug-resistant TB at individual and programmatic levels?*

iv. *What are the potential strategies for integration/scale up of drug-resistant TB management within TB control programmes?*

v. *How to reinforce PPM collaboration for treatment of DS- and DR-TB?*

vi. *How to improve decentralized and fully integrated access to TB and antiretroviral treatment ?*

5. Capacity building for operational research

Despite international interest in operational research, very little research is conducted or

published from resource-limited settings where the greatest burden of TB occurs. Building and sustaining the necessary capacity to conduct operational research at country level is a primary aspect to be considered while projects are being developed. National TB control programmes may lack essential expertise, infrastructure, staff, funds, policy cycle, and/or professional culture, and there may be only weak linkages between programmers and researchers. The key questions to be addressed in this area are:

i. *What are the existing models of operational health research capacity?*

ii. *What is the impact of existing training models in terms of products, outputs and outcomes?*

iii. *How to ensure sustainable operational research capacity at the national level?*

For each of the questions being addressed in the five priority areas above, suitable methods for operational research are described in the **second part of the publication**, together with a summary of the main operational research methods, statistics and definitions used in operational research.

The Stop TB Partnership, the WHO Stop TB Department and the Global Fund to fight AIDS, Tuberculosis and Malaria are jointly committed to the promotion of operational research as one of the keys to improving TB control and developing appropriate policies on implementation of new tools for TB control. In this respect, *Priorities in operational research to improve tuberculosis care and control* bridges the information gap on priorities for TB operational research for reference during grant application and implementation.

I. INTRODUCTION

A rational framework for operational research in tuberculosis

For the past 15 years, global tuberculosis (TB) control efforts have generated impressive results. The implementation of the World Health Organization (WHO) *Stop TB Strategy (1995-2008)* is estimated to have cured 41 million people with TB and averted 6 million deaths worldwide *(1)*. The global case detection rate (CDR) of sputum smear-positive TB rose from 15% to 61% over the period of the strategy, the treatment success rate increased from 77% to 87%, and the Millennium Development Goal (MDG) related to TB incidence is on target. However, despite these important achievements, the burden of TB continues to rise, and current control efforts need to be greatly improved if elimination (defined as ≤ 1 TB case per million population per year) is ever to be achieved. While the estimated incidence of TB has been declining globally since 2004, the present rate of decline (less than 1% per year) is insufficient to reach the elimination of TB by 2050, and the absolute numbers of TB cases continue to rise: in 2009, there were an estimated 9.4 million TB cases globally *(1)*. Reported data indicate huge gaps in the performance of national TB programmes (NTCP). Thus, it is estimated that 3.7 million TB cases, including 1.6 million with sputum smear-positive disease (39% of incident cases), were not reported by DOTS-based programmes; of an estimated 440 000 cases of multidrug-resistant TB (MDR-TB), only 30 000 (7%) were diagnosed and few of these had access to optimal treatment; lastly, of the estimated 1.1 million people co-infected with human immunodeficiency virus (HIV) and *Mycobacterium tuberculosis*, only 140 000 (12.7%) received antiretroviral therapy (ART) *(1)*.

Highly effective and widely accessible diagnostics, drugs and vaccines are needed to support TB control globally. There are technical and structural challenges, however, that impede optimal detection and treatment of all forms of TB, as well as prevention. Nevertheless, such barriers can be overcome by robust operational research to identify the solutions that have a significant impact on case detection and cure rates, and that help improve the availability and effectiveness of TB care services.

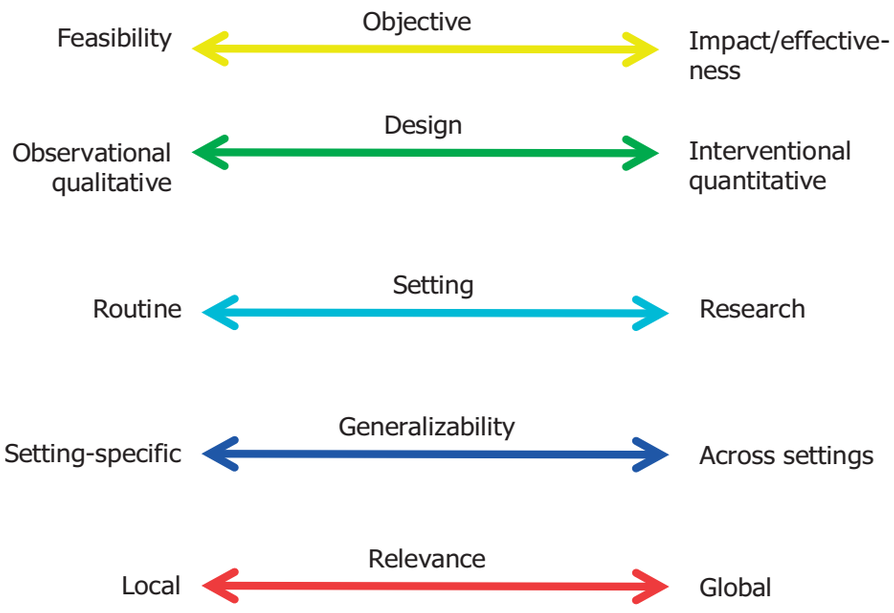
Promoting research is a key component of the *Stop TB Strategy*, which includes conducting "programme-based operational research" and "research on introducing new tools into practice" *(2)* (see **Box 1**). The importance of programme-based operational research is increasingly recognized and was recently identified as a major area on which global action is urgently needed *(3, 4)*. In its broad sense, operational research covers a large spectrum of activities, from local setting-oriented research to improve TB control programme performance, to international policy-guiding research, including the assessment of new interventions to improve TB control *(4)*. The most appropriate type and scale of operational research is largely dependent on the type of questions being addressed, the level and type of care services and the users concerned, as well as the anticipated general relevance of the results *(4)* (see **Figure 1**). At the national level, TB control programmes should carry out setting-oriented operational research projects involving partners, in order to address local or regional problems, and identify appropriate solutions *(5)*. At the international level, a robust evidence base is increasingly required for guiding policy-making (including the use of systematic reviews and GRADE[c] evaluation). Multicentre operational research projects are required to meet needs, address gaps in TB control, and to give rise to changes in international policy *(4)*.

A broad-based, concerted effort is needed to develop operational research capacity, allocate appropriate resources, and encourage all actors/

[c] See: Grading of Recommendations Assessment, Development and Evaluation (GRADE) Working Group: http://www.gradeworkinggroup.org

protagonists to work together *(6)*. The capacity to conduct operational research within NTCPs is often limited, due to lack of expertise, infrastructure, staff or funds. Research may also be constrained by the nature of the policy cycle, or an absence of the critical performance questioning what operational research entails. Investment in capacity building, which includes continuous training to undertake operational research, and interpret and act upon results, is essential to improving service delivery, and also to understanding where and why programmes do not work, as well as to guiding optimal implementation of new methods and strategies *(7)*. An enabling environment for performing operational research is key to achieving the full potential for TB control programme improvement at all levels, and managers should be aware of the benefits of building research capacities through collaboration with government research institutes, local universities and nongovernmental organizations (NGOs). Lastly, and importantly, the conduct of effective operational research relies on the availability of efficient monitoring and evaluation systems that can collect routine data reliably in order to analyse how the systems work and what can be done to improve on potential problems that are identified.

FIGURE 1. THE SCOPE OF TB OPERATIONAL RESEARCH



Adequate funding is crucial for quality operational research to provide results that are relevant to policy-makers. According to the Treatment Action Group (TAG) *(8)*, only US$ 34 million (6.8%) of the US$ 510 million invested in TB research and development in 2008 (by more than 70 reporting organizations) were for operational research, showing that only a limited proportion of funding is available for operational research compared to that for tool research and development research (i.e. diagnostics, treatment and vaccines). As operational research receives increasing attention, however, new funding opportunities are emerging to improve the use of existing technologies (e.g. increased case-finding, simplified treatment monitoring etc.) and to evaluate the impact of introducing new tools in various health systems and epidemiological settings. Among these opportunities, the United States President's Emergency Plan for AIDS Relief (PEPFAR) increasingly contributes to integrated TB and HIV activities in areas with high dual incidences, and the recently launched 'Phase II' states that "study proposals are encouraged to focus on bringing evidence into practice to improve service delivery and outcomes" *(9)*. Of equal note is the Wellcome Trust initiative to strengthen research capacity in Africa *(10)*.

The Global Fund to Fight AIDS, Tuberculosis and Malaria (Global Fund) is the largest international

funder of activities for TB control, accounting for 60% of all external financing of TB programmes. The Global Fund explicitly states that up to 10% of country proposal budgets should include monitoring, evaluation and operational research. The proportion of grants including an operational research component has increased from 19% in rounds 1–5 to 58% in round 7, showing the importance of operational research in TB control programmes. This provision of funds is sometimes not optimally used, however, due to limited local capacity for conducting operational research, or the absence of coordination mechanisms at country level to conduct appropriate research with suitable collaborators *(11)*. In this context, there is an urgent need to identify operational research priorities in order to help donors target the most pressing needs.

On this basis, the Stop TB Partnership Research Movement and the Global Fund initiated a process in February 2010 to identify ways to improve and rationalize operational research for TB control. This stepwise process included a systematic review, an expert meeting and a workshop, accompanied by wide stakeholder consultation. The objective was to critically address the increasing need for improved and rational operational research in TB control, to identify key areas where evidence was lacking for proper implementation of new and existing technologies, as well as novel service delivery models. As a result, five key areas were identified in which knowledge gaps hamper proper implementation of TB control activities:

(i) Access, screening and diagnosis of drug-susceptible and multi/extensively drug-resistant TB (M/XDR-TB);

(ii) Development of sustainable collaboration with all practitioners for TB care and control;

(iii) Prevention and treatment of TB in persons living with HIV;

(iv) Optimal access to and delivery of treatment and retreatment of drug-susceptible and M/XDR-TB; and

(v) Operational research capacity building.

In each of these five areas, critical questions were identified to improve TB control at the national and international levels, and to reach the *Global Plan to Stop TB* targets by 2015.

This document is the culmination of this undertaking and is intended to help countries to carry out operational research and prepare for grant applications to donors for operational research support. For each of the five areas above, a list of priority questions and corresponding rationales is provided, along with a synopsis of suitable methods to conduct operational research projects in respective areas. The document is intended to help programme managers, consultants and researchers who wish to conduct operational research to improve TB control at national, regional and international levels. We recognise that the priority research questions may differ depending on the country, the epidemiological setting, the health system, etc. We do not propose here a “one-size-fit-all” method for each of the potential research questions that may arise, but rather propose a “tool-kit” of potential methods to be used to address specific operational questions that may be considered critical to improve programme performance, optimize patients’ care, or adopt new TB control techniques.

BOX 1: THE SIX COMPONENTS OF THE WHO *STOP TB STRATEGY*

1. Pursue high-quality DOTS expansion and enhancement

- Secure political commitment, with adequate and sustained financing
- Ensure early case detection, and diagnosis through quality-assured bacteriology
- Provide standardized treatment with supervision, and patient support
- Ensure effective drug supply and management
- Monitor and evaluate performance and impact.

2. Address TB-HIV, MDR-TB, and the needs of poor and vulnerable populations

- Scale-up collaborative TB/HIV activities
- Scale-up prevention and management of MDR-TB
- Address the needs of TB contacts, and of poor and vulnerable populations.

3. Contribute to health system strengthening based on primary health care

- Help improve health policies, human resource development, financing, supplies, service delivery and information
- Strengthen infection control in health services, other congregate settings and households
- Upgrade laboratory networks, and implement the Practical Approach to Lung Health
- Adapt successful approaches from other fields and sectors, and foster action on the social determinants of health.

4. Engage all care providers

- Involve all public, voluntary, corporate and private providers through Public–Private Mix approaches
- Promote use of the *International Standards for Tuberculosis Care.*

5. Empower people with TB, and communities through partnership

- Pursue advocacy, communication and social mobilization
- Foster community participation in TB care, prevention and health promotion
- Promote use of the *Patients' Charter for Tuberculosis Care.*

6. Enable and promote research

- Conduct programme-based operational research
- Advocate for and participate in research to develop new diagnostics, drugs and vaccines.

REFERENCES

1. Global tuberculosis control: WHO report 2010. Geneva, World Health Organization, 2010 (WHO/HTM/TB/2010.7; also available at: http://www.who.int/tb/publications/2010/en/index.html).
2. Raviglione MC and Uplekar MW. WHO's new Stop TB Strategy. *The Lancet*, 2006;367:952–5.
3. Marais BJ et al. Scale-up of services and research priorities for diagnosis, management and control of tuberculosis: a call to action. *The Lancet,* 2010, (Published Online) May 19, 2010, DOI:10.1016/S0140-6736(10)60554–5.
4. Lienhardt C and Cobelens F. Operational research for improved TB control: the scope, the needs and the way forward. *International Journal of Tuberculosis and Lung Disease*, 2011; 15(1):6–13.
5. Harries A.D. Integration of operational research into National Tuberculosis Control Programmes. *Tuberculosis,* 2003; 83:143–147.
6. Zachariah R et al. Operational research in low-income countries: what, why, and how? *Lancet Infectious Diseases*, 2009; 9: 711–17
7. Laserson K et al. Capacity building for international tuberculosis control through operations research training. *International Journal of Tuberculosis and Lung Disease*, 2005. 9(2):145–150.
8. *Tuberculosis research and development: a critical analysis*. New York, Treatment Action Group, 2009.
9. *Public health evaluation concept submission guidance*. Washington DC, Office of the Global AIDS Coordinator/US State Department, 2010. (http://www.pepfar.gov/documents/organization/142846.pdf, accessed 2 January 2011).
10. Whitworth J et al. Strengthening capacity for health research in Africa. *The Lancet*, 2008; 372: 1590–93.
11. Xueref S. Operations research in the Global Fund: status and way ahead, in collaboration with partners. Presentation at the 5th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention. Cape Town, South Africa, 2009.

II. METHODS FOR DEVELOPING THIS DOCUMENT

This document was developed in five successive steps:

1. Expert group meeting

An expert group meeting was convened in Geneva, Switzerland on 22nd February 2010, to reflect on the need to rationalize operational research in TB control and address the translation gaps for implementation of new technologies and service designs for TB programmes, including TB/HIV and MDR-TB control. During the meeting, participants identified key areas where evidence was lacking for optimal implementation of novel technologies and service delivery models, and explored how operational research could best inform policy-making for better TB control at all levels. One outcome was a proposed workshop to outline the main directions for improved and rationalized implementation research in TB, and define the research priorities for better implementation of current and novel technologies. The following specific areas of focus were identified by participants for priority attention:

1. Access, screening and diagnosis of drug-susceptible and M/XDR-TB;
2. Development of sustainable collaboration with all practitioners for TB care and control;
3. Prevention and treatment of TB in HIV-infected patients;
4. Optimal access and delivery of treatment and retreatment of drug-susceptible and M/XDR-TB; and
5. Operational research capacity building.

To prepare for the subsequent workshop (see below), an in-depth overview of the situation was prepared by one or two expert(s) in the area, which was complemented by background papers from two discussants in order to encourage discussions.

2. Systematic review

The second step was a systematic review of operational research projects conducted over the past 10 years on the following themes: (i) clinical algorithms for diagnosis of smear-negative TB; (ii) implementation of isoniazid preventive therapy (IPT) in contacts of TB cases, and in HIV-infected people; and (iii) provision of second-line treatment for MDR-TB. The systematic review focused on studies of *effectiveness* (in contrast to efficacy), delivery and cost-effectiveness of these interventions, addressing objectives, design, study setting and scope. It showed that few studies had been completed and published on these interventions and that they were highly concentrated on a limited number of geographical and epidemiological settings. Comparative designs had hardly been used. These results highlighted clear gaps in operational research with regard to the potential to inform policy-making, and the need for coordinated action at regional and global levels.

3. Workshop

A workshop was convened on the 10–11th May 2010 in Geneva, Switzerland and assembled a wide group of scientists, public health specialists, national TB programme managers, clinicians, NGOs and community representatives. The objective of the workshop was to address the translation gaps for implementation of current and innovative technologies for TB control, and identify the priorities in operational research for improved TB control, including the improvement of current TB control methods and the uptake of novel technologies and service delivery models.

The areas identified during the expert group meeting (see above) were explored. Participants examined what research was needed to address the identified gaps ("what?"), the methods ("how"), the study population ("whom?"), and whether it should be conducted at local, national, or multicountry/international levels ("where?"). Results of the discussions were summarized by a rapporteur and discussed among all workshop delegates, resulting in the identification of critical research gaps and subsequent priorities in each of the five defined areas.

4. Stakeholder consultations

To invite comments and suggestions on these research questions, the report of the workshop (see above) was widely circulated to the working groups of the Stop TB Partnership.[d] Comments and suggestions received were included in the workshop report that served for the preparation of this document.

[d] Including the DOTS Expansion Working Groups and its subgroups: (Childhood TB, TB and poverty, Introducing new approaches and tools, Public–Private Mix); the TB/HIV Working Group; the Global Laboratory Initiative; the MDR/TB Working Group; the Working Group on New Diagnostics; the Working Group on New Drugs; and the Working Group on New Vaccines.

5. Development of specific methods

Methods and research study designs to address each of the research priorities identified above were developed as templates for use by NTCPs and their research partners. For each priority, the current document provides a synopsis of applicable research methods, according to the following outline:

- objective(s)
- design outline
- setting/study population
- methods: recruitment of subjects (eligibility criteria); intervention (as appropriate)
- expected endpoints
- analysis
- guidance for sample size calculation (and estimated number of participants)
- expected duration/timeline
- suitable scale
- estimated budget range.

Annex I provides a broad description of the main study designs used in operational research. The detail of these suggested methods and research study designs is provided in **Annex II**. In addition, a summary description of these various designs, with respective budgetary requirements, and estimated timelines is included at the end of each section in **Annex II**.

6. Scope of operational research projects

In pragmatic terms, TB operational research projects should be conducted with the aim of addressing three main objectives *(1)*: (i) to improve programme performance and outcomes; (ii) to assess the feasibility, effectiveness or impact of new strategies or interventions on TB control; and (iii) to collect data to guide policy recommendations on specific interventions (see Figure 1).

The first objective aims to assess deficiencies in TB control programmes and identify causes that are amenable to improvement by technical or managerial intervention. In that situation, research questions are *setting-oriented* and results are *setting-specific (1–3)*. Actors and protagonists are primarily TB programmes and/or the researchers commissioned for the studies, and the users are the health-care providers and/or programme managers. The scale can be local or national. Despite limited generalizability, methodological issues may be relevant to other high-burden countries or settings, so publication of the results is suitable for sharing of examples. Research questions are generated through the identification of problems or challenges encountered in programme activities, based on the review of locally-collected surveillance or programme-based data *(4)*, or upon outcomes of qualitative studies with properly defined methods that articulate the social, gender, cultural and economic aspects of the problem, in relation to where the research is undertaken *(5–7)*. Questions may also aim at identifying specific risk factors or the causal effect of given variables. These can lead to testing targeted interventions aimed at improving TB control performance locally, for instance methods for increasing case detection *(8, 9)*, improving treatment adherence *(10)* or for encouraging collaboration with the private health sector (*11*). Other disciplines such as health economics may also be useful *(12)*.

The second objective includes studies of new interventions to improve TB control, such as the effective and efficient use of new tools (diagnostics, treatments and vaccines) *(13)*, but can also assess novel algorithms or combinations of tools, or new approaches to care delivery *(14)*. Once efficacy of new interventions has been established in relevant trials or validity studies (that typically lead to approval by regulatory authorities and/or endorsement by international organizations such as the WHO), operational research needs to be conducted to determine the conditions/requirements under which these interventions can be effectively implemented through the assessment of their effectiveness, acceptability, feasibility and affordability, as well as their impact on the health system when applied under routine conditions *(15–17)*. Actors and protagonists are primarily TB programmes and/or the researchers commissioned for the studies, as well as NGOs, that can play a role in the evaluation and incorporation of new tools into the health system, for example. The users include policy-makers, care-providers and/or programme managers of the country where the study is undertaken, but also potentially in other countries with similar epidemiological patterns and health-care systems. The scale ranges from national to regional/international. Studies are therefore best done in several countries/settings, and reported in such a way that TB programme managers and policy-makers can use the results to adapt the tested implementation strategy to their own setting *(1)*.

In relation to the third objective, operational research projects are being conducted to inform international policy recommendations. Initial evaluations of new interventions under controlled conditions (e.g. clinical, laboratory etc.) need to be complemented with evaluation in real-life, programmatic conditions *(18)* in order to examine the optimal possibilities for uptake and use of novel interventions in populations, and assess their potential impact on patient outcomes (e.g. reduced treatment delay, improved cure rate, quicker return to work, reduced risk of relapse and drug resistance), as well as on the organization and structure of health services *(12)*. Relevant actors and protagonists are primarily researchers working in collaboration with TB programmes. The users are national and international policy-makers. These studies are conducted on a regional or global scale.

Since generalizing conclusions across settings is of primary importance, the choice of study locations is dictated by how representative the setting is in terms of epidemiological patterns and health-system structure, as well as local research capacity to produce high-quality data. Although most research is carried out within NTCPs, there is also a need for additional research capacity (e.g. laboratory, clinical aspects, data management etc.). Rigorous methods must be used, driven by results from safety, efficacy and demonstration studies, and should include Phase IV trials, cluster-randomized trials, cost-effectiveness studies, and impact evaluation studies *(19–22)*. By implementing similar study protocols in various countries, data can be compiled into combined databases and analysed across countries.

REFERENCES

1. Lienhardt C and Cobelens F. Operational research for improved TB control: the scope, the needs and the way forward *International Journal of Tuberculosis and Lung Disease*, 2011; 15(1):6–13.
2. Harries AD. Integration of operational research into National Tuberculosis Control Programmes. *Tuberculosis*, 2003; 83:143–147.
3. Zachariah R et al. Operational research in low-income countries: what, why, and how? *Lancet Infectious Diseases*, 2009; 9:711–17.
4. Nunn P et al. The research agenda for improving health policy, systems performance, and service delivery for tuberculosis control: a WHO perspective. *Bulletin of the World Health Organization*, 2002; 80(6):471–476.
5. Ogden JA. Improving Tuberculosis Control: Social Science Inputs. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 2000; 94:135–140.
6. Harper M et al. Identifying the determinants of tuberculosis control in resource-poor countries: insights from a qualitative study in The Gambia. *Transactions of the Royal Society of Tropical Medicine and Hygiene,* 2003; 97:506–10.
7. Noyes J and Popay J. Directly-observed therapy and tuberculosis: how can a systematic review of qualitative research contribute to improving services? A qualitative meta-synthesis. *Journal of Advanced Nursing*, 2007; 57(3),227–243.
8. Ramsay A et al. Front-loading sputum microscopy services: an opportunity to optimise smear-based case detection of tuberculosis in high prevalence countries. *Journal of Tropical Medicine*, 2009; 398767. Epub 2009, Mar 15.
9. Miller AC et al. Controlled trial of active tuberculosis case finding in a Brazilian favela. *International Journal of Tuberculosis and Lung Disease*, 2010; 14(6):720–726.
10. Hane F et al. Identifying barriers to effective tuberculosis control in Senegal: an anthropological approach. *International Journal of Tuberculosis and Lung Disease,* 2007; 11(5):539–43.
11. Uplekar M. Involving private health-care providers in delivery of TB case: global strategy. *Tuberculosis*, 2003; 83(1-3):156–64.

12. Vassall A et al. Reforming tuberculosis control in Ukraine: results of pilot projects and implications for the national scale-up of DOTS. *Health Policy Planning,* 2009 Jan; 24(1):55–62.

13. Pai M, Minion J, Steingart K and Ramsay A. New and improved tuberculosis diagnostics: evidence, policy, practice, and impact. *Current Opinion in Pulmonary Medicine*, 2010; 16(3):271–84.

14. Ramsay A et al. New policies, New technologies: Modeling the potential for improved smear microscopy services in Malawi. *PLoS One*, 2009 Nov 10; 4(11):e7760.

15. Sanders D, Haines A. Implementation research is needed to achieve international health goals. *PLoS Med*, 2006; 3: e186.

16. Ramsay A, Steingart KR and Pai M. Assessing the impact of new diagnostics on TB control. *International Journal of Tuberculosis and Lung Disease,* 2010; 14(12):1506–7.

17. Mann G et al. Beyond accuracy: creating a comprehensive evidence base for tuberculosis tools. *International Journal of Tuberculosis and Lung Disease,* 2010; 14(12): 1518–1524.

18. Victora CG, Habicht JP and Bryce J. Evidence-based public health: moving beyond randomized trials. *American Journal of Public Health*, 2004; 94:400–405.

19. Godfrey-Faussett P. District-randomized phased implementation: strengthening the evidence base for cotrimoxazole for HIV-positive tuberculosis patients. *AIDS,* 2003; 17(7):1079–1081.

20. van Cleeff MRA et al. The role and performance of chest X-ray for the diagnosis of tuberculosis: a cost-effectiveness analysis in Nairobi, Kenya. *BioMed Central Infectious Diseases*. 2005; 5:111.

21. Grosskurth H et al. Impact of improved treatment of sexually transmitted diseases on HIV infection in rural Tanzania: randomised controlled trial. *The Lancet,* 1995; 346(8974):530–6.

22. Thiam S et al. Improving adherence to tuberculosis treatment in a resource-poor setting: A Cluster Randomised Controlled Trial. *Journal of the American Medical Association*, 2007; 297(4):380–6.

III. PRIORITY AREAS FOR OPERATIONAL RESEARCH IN TB CARE AND CONTROL

1. Improving access, screening and diagnosis of TB

TB diagnosis in most endemic countries relies heavily upon direct sputum smear microscopy, as this is most often the only simple test that can be used below the reference laboratory level. Currently, however, only about 60% of all infectious TB cases are being detected with this test, and a proportion of patients with active TB (i.e. those listed in the laboratory registers as having at least one positive smear) do not collect their test results, so do not receive appropriate treatment *(1)*. In addition, among the estimated half a million cases of MDR-TB that occur globally each year, only a tiny fraction are identified and treated appropriately. From a diagnostic perspective, this is largely due to services not being accessible to patients. Diagnostic services for drug-resistant TB (DR-TB) are based on complex technologies that require sophisticated, biosafe laboratories with highly-trained staff, and so are rarely available outside of national reference laboratories. As transport of specimens from the periphery to the reference laboratory(ies) can also be problematic, TB diagnostic services may be non- or dysfunctional. Similarly, since direct smear microscopy is less sensitive in patients with HIV-associated TB, further testing using complex technologies in sophisticated laboratories is needed to reliably diagnose HIV/TB co-infection *(2)*. New, simple and inexpensive tools are needed to identify the various forms of TB (including DR-TB and HIV-associated TB), particularly at the lower levels of health services.

New tools are becoming available *(3)*. Since 2007, WHO has endorsed the use of over 10 new TB diagnostic tools (technologies or approaches) that, if used wisely, could improve TB control considerably. In 2010, WHO endorsed a new automated real-time nucleic acid amplification technology (NAAT) for rapid and simultaneous detection of TB and rifampicin resistance (the Xpert MTB/RIF system) that offers the prospects of drastically improved diagnosis of active TB and MDR-TB. This test, however, requires significant initial capital investment and subsequent running costs for sustainability.

In a recent survey on the use of diagnostic tools in 16 high-burden TB countries from 2007 to 2009, 50% of these countries reported using TB diagnostic tools recommended by WHO; national TB control programme managers reported diverse challenges in the implementation of new diagnostics, but no impact assessment of their introduction on TB control was carried out *(4)*. In general, there is insufficient evidence available to determine which package of current and newly-developed diagnostic tests would work best in a given set of circumstances, and there is as yet little guidance to countries on what new diagnostic tools, or combinations of tools, should be implemented in particular epidemiological/health system settings or for different risk groups, and at what level of the health service it should be done. Operational research with collection of multiple country experiences could act as the starting point to better guide diagnostic scale-up, and regional or global policy in the future. This requires the creation of an environment that is conducive for operational research around TB diagnostics and diagnostics services, preferably using a multi-disciplinary approach, starting with careful situation analysis.

Outstanding questions for optimization of TB diagnosis

i. Improving access to TB diagnosis

Based on the currently used TB diagnostic tests (sputum smear microscopy, chest X-ray, culture/drug susceptibility testing (DST)), what are the various socioeconomic, health system-related and qualitative barriers that influence TB diagnosis at the patient and health-provider levels (in terms of timeliness of diagnosis, convenience and cost to the patient, as well as prevention of primary treatment default)? Which interventions (decentralization to primary level facilities, decentralization into communities, use of mobile clinics, etc.) would be most effective in overcoming these

barriers? How can diagnostic services be brought closer to the community (decentralization, active case-finding, mobile systems, etc.), and how can they be integrated into the general health system, bearing in mind that the nature of the technology employed and how it is delivered will determine patients' access to the service?

ii. Improving screening of TB suspects and high-risk groups

Who should be screened, what should they be screened for, and how should they be screened? For this, higher-risk groups for different forms of TB, to which intensive case-finding can be targeted, must be identified (e.g. people living with HIV (PLHIV), prisoners, vulnerable groups, MDR-TB suspects, patient contacts, people with risk factors such as diabetes, smoking, alcohol or drug use). Appropriate screening algorithms and test methods (e.g. symptom questionnaire, conventional or digital chest X-rays, light-emitting diode (LED) microscopy and/or front-loaded microscopy and/or other laboratory tools) must be defined. Similarly, reliable methods need to be established for ruling out active TB in various screening settings, including prior to IPT. For instance, recently revised WHO guidelines on IPT recommend the use of a simplified screening algorithm based on the absence of all four clinical TB symptoms (current cough, night sweats, fever and weight loss) to help identify PLHIV who have less likelihood of active TB disease and hence are eligible for IPT *(5)*. This simplified symptom-based algorithm should be used for all adults living with HIV, including pregnant women, people who are receiving ART, and those who have successfully completed TB treatment. However, the way in which these symptoms are confirmed or excluded is likely to be context-specific, and screening approaches may vary among risk groups and so will the yield of screening approaches. Subsequently, data are needed to evaluate the performance of the new IPT guidelines locally.

iii. Assessing the role of new diagnostics tools to improve practices

As mentioned above, WHO has recently endorsed optimized smear microscopy approaches, a variety of commercial and non-commercial options for culture and DST (e.g. microscopic-observation drug susceptibility (MODS) assay, colorimetric redox indicator and nitrate reductase assay), and most recently the Xpert MTB/RIF system. How should these new tools be introduced into health systems? What would be the appropriate screening and diagnostic algorithms in different settings and for various risk groups? What would be their contribution to improved case detection and treatment of drug-sensitive (DS) and DR-TB? But, also, what will be the human resource implications of introducing new tools (training, number and cadre of staff)? What are the infrastructure implications (such as equipment, laboratory layout and safety installations)? What will be the projected impact of going to scale with a new tool (e.g. cost savings to patients in relation to income, cost savings to health providers or the health system, effects on transmission of improved infection control as a result of the new tool)? At a broader scale, how can TB diagnostic services be integrated with other diagnostic services for infectious as well as noncommunicable diseases? The delivery of these services will clearly depend upon the existence of a functional and interconnected, tiered health system. Operational research must address the issues associated with dysfunctions in specimen and patient referral systems, as well as the communication of and response to laboratory results.

iv. Active case-finding

Introducing optimal combinations of diagnostic tools at all levels of health care and improving access to early diagnosis for all TB cases are key to decreasing the period of infectiousness before diagnosis as well as improving patients' outcomes. Active case-finding (ACF) has high potential to rapidly improve TB control. Recently, ACF in communities has shown high potential to substantially improve case detection, with rapid declines in undiagnosed TB *(6)* and increased numbers of TB cases diagnosed when ACF was added to routine facility-based DOTS *(6–9)*. However, with the exception of TB screening in PLHIV, there has been very little research into ACF in high TB burden countries, as this was not part of international policy until very recently. Further operational research is needed to identify how best to introduce ACF in different epidemiological settings.

The choice of ACF strategy should be guided by the local epidemiology of undiagnosed TB disease. In high-prevalence settings (e.g. southern Africa, crowded urban communities, prisons) an untargeted approach that regards all adults as being at high risk may be more appropriate than one that relies on identification of individual risk factors for disease such as HIV infection, diabetes or recent close

contact with infectious TB cases (especially if these risk factors are themselves mostly undiagnosed). However, in low and medium TB burden settings, an untargeted approach is likely to have a very low yield in most communities.

Even when definitions of TB suspects and diagnostic algorithms are identical to those used in facility-based DOTS (sputum microscopy in patients self-reporting chronic cough), community-based interventions (e.g. using mobile outreach vans or temporary microscopy clinics) can still have high yield and impact *(6)*. More intensive approaches to seeking out people with unreported cough (such as door-to-door enquiry) may be, paradoxically, less effective – perhaps through imposing unwelcome, unannounced visits to people's homes. Not all patients with undiagnosed TB have chronic cough, however, and so strategies that successfully investigate everyone, regardless of symptoms, or use a broader definition of TB symptoms, will have a higher yield if well implemented, but will be more challenging and resource intensive. In this respect, cost-effectiveness studies that compare various approaches in specified settings based on observed data (e.g. from pilot projects) can be extremely informative.

v. Building accessible, effective and efficient diagnostic services with new diagnostic tools

The major question across all the above areas is: What combination of diagnostic tools should be introduced and what determines the appropriateness of particular combinations to given national programmes/health services? Until this question is addressed, most of the work in areas *i–iii* above will remain unfocused. This requires a broad approach to revising existing clinical diagnostic algorithms and developing new algorithms incorporating existing and new diagnostic tests, taking into account the prevalence of HIV and HIV-associated TB, the prevalence of M/XDR-TB, as well as infrastructural issues such as transport and health systems assessment. Such changes would have major implications for the administration of NTCPs and ministries of health. Starting work in this area will require the identification of 'best-fit' packages of particular diagnostics that might be used effectively and efficiently in a given epidemiological and infrastructural situation, following clearly laid-out algorithms that give the highest likelihood of correctly diagnosing patients with TB quickly and at reasonable cost.

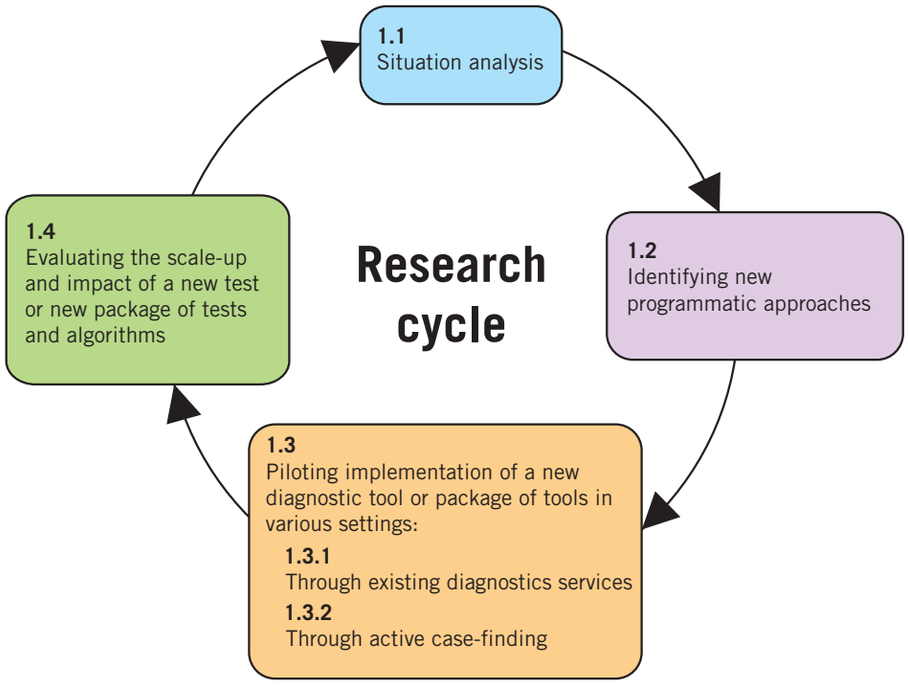
vi. Evaluation of the impact of new tests or new approaches

Determining the impact of any new diagnostic test (or package of tests) not only requires the assessment of the test's sensitivity and specificity with all forms of TB and in all populations affected by TB (including those co-infected with HIV), but also its effect on important patient outcomes (such as cure rate and mortality), accessibility (particularly to poor and vulnerable TB suspects), affordability (to the health system and to the patient) as well as staffing and infrastructure requirements. These data can be collected through national level operational research, and should be collated to help inform decisions on whether and how a test can be incorporated into current policy and practice. An Impact Assessment Framework has been developed to help guide this process and is shown in **Annex III** *(10)*.

Priority areas for operational research

The outstanding research questions have been prioritized to facilitate operational research that leads to improvements in the accessibility, quality and scope of diagnostic services for TB. Knowledge is needed at all levels from local to international and the research questions need to be asked at all levels. There may be different answers at different levels, but overall, there is a logical series of successive research projects to be conducted, as shown in **Figure 2**.

FIGURE 2. CYCLE OF RESEARCH ACTIVITIES FOR IMPROVED ACCESS, SCREENING AND DIAGNOSIS OF TB



1.1 Situation analysis

This comprises the baseline assessment. Studies need to be conducted to identify the local barriers to access to diagnosis of TB in order to allow better use of *existing* systems for diagnosing all forms of TB in various populations (including difficult-to-reach populations) and specific risk groups (TB suspects, DS-TB, retreatment, DR-TB, PLHIV, children, people with diabetes or other risk factors).

1.2 Identifying new programmatic approaches

This step is aimed at determining the best approaches that include revised clinical algorithms for TB diagnosis, and help in defining the most effective use of new diagnostic tool(s) in specific settings and populations (i.e. screening or confirmatory, rule-in or rule-out, etc.) so as to maximize their impact.

The example of Xpert MTB/RIF system is informative here. In the third quarter of 2010, WHO convened an expert group meeting to evaluate existing accuracy and operational data about Xpert MTB/RIF and reported the findings to the WHO Strategic and Technical Advisory Group for TB (STAG-TB). Following STAG-TB's recommendation to endorse the new diagnostic tool, WHO organized a wide-scale consultation to inform roll-out of Xpert MTB/RIF. A roadmap for implementation was prepared, that included operational research to validate interim screening and diagnostic algorithms, inform anticipated changes in TB case and outcomes definitions, and provide early data on programmatic aspects and impact information.[e] Other examples of such approaches include: the development of improved clinical algorithms for smear-negative TB in high, medium and low HIV prevalence settings with and without access to (digital or conventional) chest X-ray; or the identification of specific risk factor profiles for MDR-TB in different settings, that could be used for the presumptive identification of suspect MDR-TB cases or identification of MDR-TB risk groups.

[e] Recommendations are as follows:

i. In high MDR-TB settings: persons at risk of MDR-TB (e.g. treatment failures, other retreatment cases, close contacts of MDR-TB cases) should be tested using Xpert MTB/RIF as the primary diagnostic test;

ii. In high HIV prevalence settings: PLHIV with signs and symptoms of TB, those seriously ill and suspected of having TB regardless of HIV status, and those with unknown HIV status presenting with strong clinical evidence of HIV infection, should be tested using Xpert MTB/RIF as the primary diagnostic test;

iii. In other settings: Xpert MTB/RIF is recommended as the primary diagnostic test where available, including in PLHIV in these settings, or as a follow-on test (at a higher level of the health service) after screening by sputum smear microscopy (at lower level of the health service) or after screening by chest radiography.

1.3 Piloting implementation of a new diagnostic tool or package of tools in various settings

1.3.1 Through existing diagnostic services (routine health service provision)

Identification of new programmatic approaches leads to questions regarding which infrastructure and delivery systems are appropriate for implementation of a new diagnostic test in a given country: Does the new test perform as well as expected for diagnosing patients who would otherwise remain undiagnosed or be diagnosed much later (effectiveness)? Are health-care providers willing and able to utilize the new test? Does the new test facilitate equitable access for all patients? Will it reduce diagnostic delay? What is the impact on the main patient treatment outcomes? What are the feasibility and conditions for scale-up? What is the cost-effectiveness of the new tool using locally-relevant costing and outcome data and what is the most cost-effective positioning of the test in a diagnostic algorithm given epidemiological and economic conditions?

In the epidemiological context of predominant DS-TB, the example of the recently endorsed Xpert MTB/RIF is, here also, informative: What will be the impact of the introduction of the cartridge-based NAAT or the rapid culture, or a combination of those? What is the role of chest X-ray in the diagnostic pathway in situations where MDR-TB and HIV associated TB is of lesser concern? What proportion of newly-detected cases will have access to treatment services? How cost-effective is the new combined approach in routine settings? What would be the impact on treatment and patient management? What would be the impact on access to care by different socioeconomic groups? As the answers to these questions largely depend on specific settings, global/regional expert meetings can provide guidance to countries in identifying the likely provisional algorithms for particular settings (see **section 1.2** above).

As scale-up of Xpert MTB/RIF is likely to take time, and the tool may not be used everywhere, other approaches for detecting new TB cases that have been endorsed by WHO and which have not yet been widely implemented through research can be investigated - for example through the following sequence:

(i) Improving technical performance of sputum smear microscopy services for TB: WHO recently endorsed a more sensitive definition of a smear-positive TB case and endorsed tools to reduce the workload in smear microscopy labs (two-specimen approach) and increase sensitivity of sputum smear microscopy (using LED-based fluorescence microscopy). Can a combination of these reduce the workload in smear microscopy labs and increase the number of sputum smear-positive cases detected in routine settings? Does it increase the TB notifications and number of patients cured? How cost-effective is the dual approach?

(ii) Improving treatment access for sputum smear-positive TB cases detected: Can same-day smear examination (with or without same-day reporting and treatment initiation/ referral) reduce initial patient default? Does it lead to improved access to treatment? How cost-effective is the approach? Are there opportunities for improvement of infection control practices through reorganization of patient flow in waiting rooms?

(iii) Improving the presumptive (non-bacteriological) detection of smear-negative TB (including HIV-associated TB) through smear-negative clinical algorithms: With improved sputum smear microscopy services and access to treatment, can improved clinical algorithms be applied in routine settings to increase the number of TB cases detected and their access to treatment?

In the context of high MDR-TB rates, the question may be different. For example, how to improve the detection of MDR-TB cases using rapid culture-based techniques and/or NAAT (including line probe assays (LPA) and cartridge-based polymerase chain reaction (PCR))? What increase in confirmed MDR-TB case detection will result from the introduction of rapid identification of *M. tuberculosis* isolates and rapid DST? What proportion of detected cases will have access to and use second-line treatment? How cost-effective is the approach in routine settings? Here again, the situation will be highly dependent on the prevalence of HIV-associated or DR-TB, together with the level of laboratory and other infrastructure (including transport and communication infrastructures) available in different settings. Global and regional expert group meetings should provide guidance to countries in identifying the likely provisional 'best-fit' packages of interventions for specific settings.

1.3.2 Through active TB case-finding

In addition to the introduction of combinations of new and current diagnostic tools at all health-care levels, enhanced and active TB case-finding activities must be developed and strategized in order to improve case detection. The following related operational research questions should be addressed.

(i) *Defining TB suspects and evaluating different screening algorithms:* The cost and yield of using different ACF strategies to define who and how to screen needs systematic evaluation. This includes evaluation of standard community-based interventions using traditional approaches (e.g. smear microscopy to investigate self-reported chronic cough) as well as comparison with promising alternatives, such as mobile digital radiography. Cost-effectiveness/cost utility analyses should be an important part of these evaluations.

(ii) *Service delivery to the community*: Ideally ACF should be integrated into existing community-level services rather than delivered as stand-alone 'vertical' interventions. However, some notable examples of disappointingly low yield have been reported when an integrated approach has been taken *(11, 12)* (see **Annex V**). Therefore, different models of service delivery need to be developed and compared in urban and rural settings with high and low TB burdens.

(iii) *Sustained effect on TB incidence of ACF in high and medium/low TB burden countries and communities*: ACF in the general community has potential to provide the rapid gains in TB control that are needed to accelerate global TB control, but many uncertainties remain. Careful operational research to measure the effect of ACF on TB incidence needs to go hand-in-hand with sustained implementation of ACF strategies that are found to be effective (increased numbers of cases diagnosed initially) in different settings.

1.4 Evaluating the impact of scale-up of a new test or new package of tests

After a new test (or package of tests) has been implemented through research, as outlined in **section 1.3** above, policy-makers have to make decisions about whether to go to full scale with the new algorithms or modifications of these. At this stage the systematic presentation of data obtained through application of the Impact Assessment Framework (see **Annex III**) during the implementation studies can promote rational decision-making about whether and how to go to scale. It is important that data are generalizable at this stage, so that policy-makers can assess the potential contribution to overall improved case detection and treatment of DS- and DR-TB. Operational modelling is helpful at this stage to evaluate the likely inputs and costs of going to scale, and transmission modelling can indicate the likely impacts on TB transmission, and, over time, the impacts on TB epidemiology. Combining these two approaches (as proposed in Layers 4 and 5 of the Impact Assessment Framework) can make it possible to present different costed options for the process of final scale-up to policy-makers to facilitate the decision-making process.

Once a decision has been made to go to full, national scale with a new test or package of tests, the overall public health and societal consequences need to be documented. Projections can then be checked and final benefits and challenges documented. For this stage, implementation research that focuses on audit, before-and-after assessments, and monitoring and evaluation are appropriate (see **Annex II**).

*Note: it should be noted, however, that it is not always necessary to conduct detailed research in all areas of the research cycle, if there is sufficient knowledge in existing situation analyses (see 1.1 above) and identification of new programmatic approaches (see **section 1.2**) for a country to proceed to implementation through research (see **section 1.3**). The Xpert MTB/RIF system provides a potential example to illustrate this. It is widely documented in resource-poor settings that the repeat visits to health facilities required for sputum submission for smear microscopy is a significant barrier to diagnosis (13–17). The Xpert MTB/RIF offers the potential to deliver an accurate TB diagnosis within two hours of submission of a single sputum specimen. WHO has developed potential algorithms for different epidemiological situations in which Xpert MTB/RIF can be used (see **section 1.2**). In this example, therefore, a given country might choose to proceed straight to implementing Xpert MTB/RIF through phased research that permits documentation of operational requirements/inputs as well as the clinical and epidemiological effects/ outputs (see **section 1.3**).*

2 Developing sustainable collaboration with all care-providers for TB care and control

Global TB case detection is stagnating at around 60% despite strengthened TB programmes. In many countries, a significant proportion of TB suspects and cases, including poor and vulnerable populations, present themselves to a range of public and private care-providers that are not linked to NTCPs *(18)*. These include informal and formal, commercial and non-profit, individual and institutional private sector care-providers such as traditional healers, pharmacists, general practitioners, private clinics and hospitals, NGOs and faith-based organizations (FBOs), and employee health services by the business sector, as well as public sector care-providers such as general and speciality public hospitals, academic institutions, prison and military health services. Evidence shows that TB diagnosis and treatment practices of many non-programme care-providers are inappropriate and that care-seeking from diverse care-providers hampers access to quality TB care, causes delays in TB diagnosis and imposes financial burden on patients *(19)*. Furthermore, it is estimated that only about 5% of people with MDR-TB are managed within TB programmes. Several 'public–private' and 'public–public' mix' (PPM) projects have demonstrated the feasibility, effectiveness, cost-effectiveness and scalability of engaging non-programme care-providers in TB care and control in diverse country settings. In some settings, PPM has also been shown to improve access, enhance equity, reduce diagnostic delays and reduce costs of care for TB patients. Consequently, WHO advises countries to undertake baseline and periodic national situation assessment (NSA) to determine the need and scope of implementing and scaling up of PPM. A tool to help conduct NSAs and guidance for PPM implementation is available and has been used effectively by many countries *(20)*. By 2008, 58 out of 93 active, TB-related country grants from the Global Fund had a PPM component, amounting to about 5% of the total allocation.

Outstanding questions for optimization of collaboration with all care-providers

i. Improving and scaling up existing approaches to engaging all care-providers

While there are examples of PPM projects that are being taken to scale, knowledge gaps persist in relation to models or approaches for nationwide PPM scale-up. It remains unclear how to prioritize providers for engagement. More needs to be learnt also about specific models and approaches for scale-up, such as the use of incentives and enablers, the use of novel regulatory approaches, and the use of social marketing and franchising. Provider segmentation and adaptation of approaches to fit specific groups of providers who are targeted to provide specific services – such as TB suspect identification and referral, TB treatment support, etc. – need to be better understood, and the role of PPM in the broader aspects of ACF among symptomatic patients requires clarification. Effectiveness and feasibility of PPM implemented as part of the Practical Approach to Lung Health (PAL) also need to be investigated. Better knowledge of issues related to quality in TB

care and control is needed as PPM initiatives are taken to scale, both in terms of adherence to the *International Standards of TB Care,*[f] and from the perspective of patients.

ii. Measuring the contribution of different provider groups to TB care and control

Measurement of the contribution of diverse care-providers to a variety of TB control tasks is difficult and may put an undue burden on recording and reporting systems. The components of PPM contribution that need to be measured at different levels (e.g. referral, microscopy, treatment, directly-observed treatment, default retrieval) all need to be tailored to country contexts. Tied to this is the need to understand the resource requirements from the programme perspective for scale-up and, thus, the ability to weigh PPM outputs and outcomes against resource inputs, which can help to monitor PPM cost-effectiveness as initiatives are taken to scale.

iii. Encouraging involvement of as yet unengaged providers

Gaps exist in the provision of TB care at the national and local levels. To identify ways of filling these gaps, it will be helpful to determine what models are appropriate for the specific national/local context; whether there are additional providers (as yet uninvolved in PPM) that could provide services; and to assess the effectiveness of models that involve these providers.

iv. Encouraging involvement of the non-public sector in MDR-TB management and TB/HIV collaborative activities

In some countries, much of the MDR-TB management is currently carried out by NGOs and private medical practitioners. Similarly, TB/HIV collaborative activities are being performed by multiple providers, with variable levels of coordination. It will be important to compile information on such approaches, and their quality and effectiveness.

v. Developing and assessing responses to changing involvement of diverse providers in TB care and control

As new TB diagnostic tests and drugs become available, it will be important to ensure they are used rationally by all providers. New diagnostics could mean the end of the distinction between pulmonary smear-positive and pulmonary smear-negative TB, with consequent changes in treatment protocols and recording. New drugs may lead to different treatment schedules. While it is expected that within government provision use of the new diagnostics and drugs will be well-regulated, there is a need to ensure that private practitioners use them rationally, to avoid misdiagnosis and development of resistance to new drugs.

vi. Encouraging introduction of regulatory approaches to collaborating care-providers

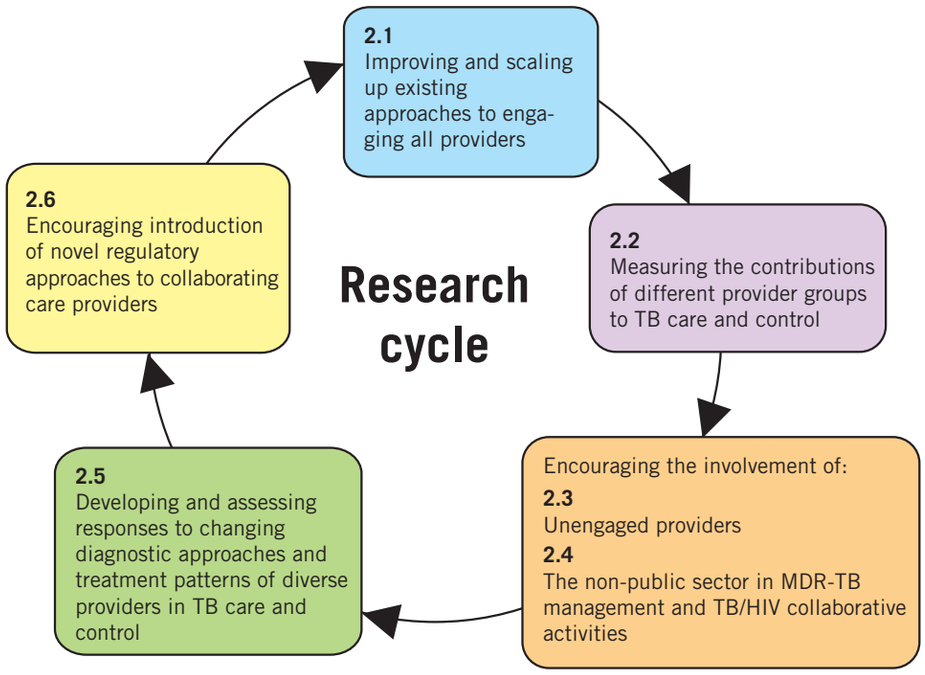
Countries are likely to have very different approaches to regulation of different aspects of TB care, including mandatory TB case notification, certification and accreditation, and restricting access to anti-TB drugs to collaborating care-providers. Some are already using regulation successfully; others have regulation but have problems with enforcement; while others have not yet attempted regulation. Clearer understanding of what works and what does not in different contexts will help countries to determine what approaches might be appropriate in their own settings; what enabling factors can be used; and whether potential disabling factors exist in their context.

Priority areas for operational research

The outstanding research questions have been prioritized to facilitate operational research that leads to optimization of collaboration with all practitioners (see **Figure 3**).

[f] Available at: http://www.who.int/tb/publications/2006/istc/en/index.html

FIGURE 3: CYCLE OF RESEARCH ACTIVITIES FOR OPTIMIZING COLLABORATION WITH ALL CARE PROVIDERS



2.1 Improving and scaling-up existing approaches to engaging all providers

The objective is to develop an evidence base of different PPM models and approaches to scale-up that include contextualized analyses of reasons for success/failure as well as mechanisms to create demand for quality services. This includes assessing enablers and incentives for various care-providers and the various mechanisms to fund scale-up.

2.2 Measuring the contributions of different provider groups to TB care and control

As various care-providers (e.g. private, NGOs, FBOs, public health units outside the Ministry of Health, etc.) intervene in TB control, it is important to assess their contributions, as well as assess their respective abilities to improve user access, case detection and outcomes for underserved groups, and reduce diagnostic delays and costs of care. This will allow estimation of the needs and resource requirements for scale-up.

2.3 Encouraging involvement of unengaged providers

It is important to assess potential models and approaches to involving various care-providers who are not yet engaged in a PPM framework. This includes identifying the potential new providers that could provide accessible and effective services, and assessing the effectiveness of models involving these newly-identified care-providers.

2.4 Encouraging involvement of the non-public sector in MDR-TB management and TB/HIV collaborative activities

The objective is to develop an evidence base of different models and approaches, through the identification of potential providers that could provide accessible and effective services for MDR-TB and TB/HIV management and the assessment of effectiveness of various models for PPM in MDR-TB and TB/HIV care and prevention.

2.5 Developing and assessing responses to changing diagnostic approaches and treatment patterns of diverse providers in TB care and control

The objective is to identify and assess ways to ensure rational use of new diagnostics and drugs in the private sector.

2.6 Encouraging introduction of novel regulatory approaches to collaborating care-providers

This research process will require the development of an evidence base of regulatory approaches (such as mandatory TB case notification, certification and accreditation) that includes contextualized analyses of reasons for success/failure, and then assessment of the possibility to develop locally appropriate regulatory approaches.

3 Prevention and treatment of TB in persons living with HIV

Optimal control of TB in high-HIV burden areas requires implementation of collaborative TB/HIV interventions through a sound policy and programme environment that gives due consideration to the local context, the respective epidemiology of TB and HIV, as well as the health system infrastructure that determines service delivery models. System-wide differences between HIV and TB care-providers and stakeholders – and operational difficulties for providing effective and appropriate interventions – have contributed to the lesser implementation and scaling up of collaborative activities *(21)*. It is therefore important to identify measures that would facilitate wider implementation and scaling up of collaborative TB/HIV interventions through effective service delivery models, including community-based interventions.

Collaborative TB/HIV interventions include the prevention of TB in PLHIV, the joint treatment of TB and HIV in people dually infected and improved infection control and prevention in various health settings *(22)*. In HIV-positive individuals, IPT has been shown to reduce the risk of TB (Relative Risk (RR): 0.64, 95% CI 0.51 to 0.81), with almost all of the protective effect in individuals with a positive tuberculin skin test (RR 0.38, 95% CI 0.25 to 0.57) compared to those who had a negative test (RR: 0.83, 95% CI 0.58 to 1.18) *(23)*. In this population, however, diagnosis of latent TB infection (LTBI) and exclusion of active TB are difficult, especially in resource-poor, high-prevalence settings *(24)*. The WHO 'Three I's' strategy *(22)* recommends that PLHIV are screened for tuberculosis, and, if tuberculosis is ruled out, IPT should be provided. Despite this recommendation, less than 1% of PLHIV were started on IPT in 2008 *(25)*. Among PLHIV who develop TB disease, mortality remains significantly higher than among HIV-negative patients. Early initiation of co-trimoxazole and ART can reduce mortality, but linking HIV-infected TB patients to HIV care and treatment has proven challenging. Recent data from South Africa have shown that a synergistic approach to delivery of IPT and ART may result in a decline in risk of active TB *(26)*. Therefore, operational research is needed in high-burden countries to optimize prevention and treatment of TB in PLHIV and address the several barriers that may occur at the level of screening, diagnosis, treatment and prevention *(27, 2)*.

Outstanding operational questions for improved TB/HIV core group activities

i. Barriers to TB diagnosis

Among PLHIV who have active TB, in addition to the lack of sensitive and rapid diagnostic tools, barriers to TB diagnosis include the lack of awareness by HIV care-providers about why and how they should screen for TB disease, inadequate recording, reporting and monitoring of TB screening in HIV care and treatment settings, and limited access to TB diagnostics in HIV care and treatment settings. A simplified screening algorithm was recently recommended by WHO to identify those PLHIV who have less likelihood of active TB disease and hence are eligible for IPT *(5)*. This algorithm relies on the absence of the four main clinical symptoms (current cough, night sweats, fever, and weight loss), to identify those PLHIV who are eligible to receive IPT. This simplified symptom-based algorithm should be used for all adults living with HIV, including pregnant women and people receiving ART, and should therefore be operationalized in various settings. Data are therefore needed on the local performance of this new guideline.

ii. Barriers to IPT initiation

Among PLHIV who have been screened and are deemed eligible for IPT, barriers to IPT initiation

remain formidable, despite the presence of substantial clinical research to address these. For example, many programmes remain unclear about whom they should prioritize for IPT (i.e. whether IPT should be offered to all HIV patients who do not have active TB disease, or whether it should be restricted to those with specific tuberculin skin test results or to patients above or below a specific CD4 count threshold). Similarly, there are questions about the optimum mode of delivery (e.g. in pre-ART and in ART clinics or through home care services; in HIV facilities or TB facilities), timing and duration of therapy, frequency and type of monitoring for adherence, toxicity, and breakthrough TB disease, and methods to train and motivate health-care workers for use of IPT in addition to ART and co-trimoxazole therapy among PLHIV.

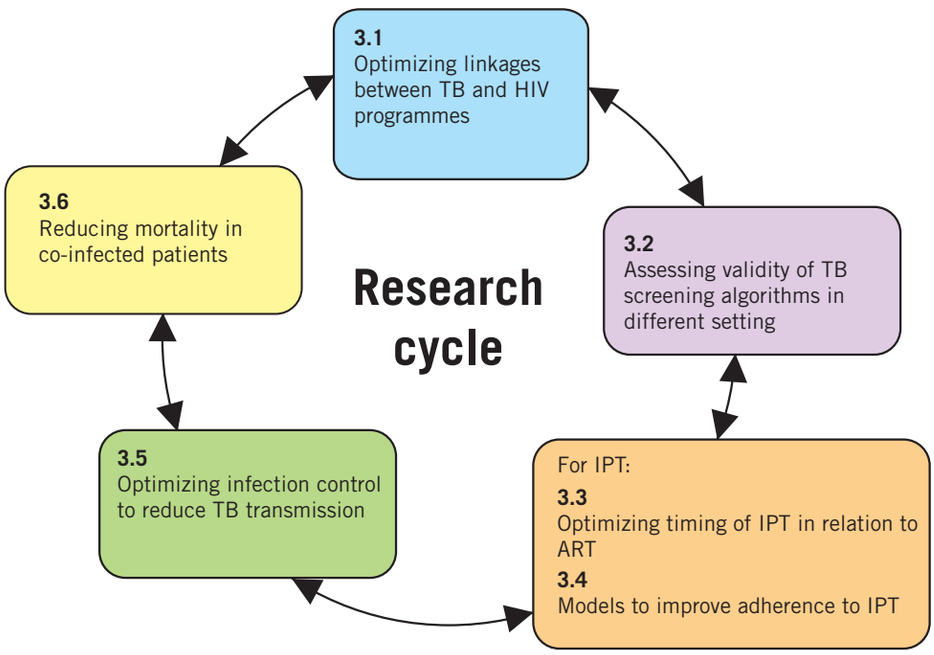
iii. Barriers to optimal combined TB/HIV diagnosis and treatment

HIV testing has been scaled up rapidly among patients who seek care in TB clinics. Nevertheless, in many settings, particularly countries with concentrated or low-level HIV epidemics, major questions remain about whether or how to implement targeted, as opposed to universal, HIV testing of TB patients. When TB patients are diagnosed with HIV infection, they should immediately be evaluated for HIV care and treatment, particularly co-trimoxazole and ART, to reduce short-term mortality. Nevertheless, the optimum models for joint TB and HIV diagnosis and treatment remain ill-defined at the patient and public health level.

Priority areas for operational research

The outstanding research questions have been prioritized to facilitate operational research that leads to optimal control of TB in high-HIV burden areas (see **Figure 4**).

FIGURE 4: CYCLE OF RESEARCH ACTIVITIES FOR COMBINED TB/HIV MANAGEMENT



3.1 Optimizing linkages between TB and HIV programmes

Key questions to address here are: What forms of cross-referral, co-location of services, and community participation would increase: (i) the proportion of PLHIV who are screened for TB; (ii) the proportion of PLHIVs in whom IPT is initiated; and (iii) the proportion of PLHIVs who survive during treatment of TB disease? Of note, specific attention should be paid to studying these questions in high priority populations, such as children, injection drug users and prisoners. Specific studies are required to determine the best linkages between TB and HIV programmes. In particular, the following questions need to be addressed:

- What are the best strategies to integrate and deliver joint TB/HIV interventions, including ART, at the community- and health sector level to TB/HIV co-infected adults, children and their families?

- What are the optimal models of community participation (i.e. effective, feasible, acceptable and sustainable) for enhanced TB case-finding and early HIV detection, in order to reduce any delay in initiation of TB and HIV care, and reduce TB and HIV transmission?

- How can the cost-effectiveness of joint TB/HIV interventions delivered through community approach and through health facilities be estimated?

- What are the best delivery models of collaborative TB/HIV interventions for most at risk and special populations in all settings, including those with different TB and HIV epidemiology and epidemic states?

3.2 Assessing the validity of TB screening algorithms in different settings

In PLHIVs who attend health facilities, does implementation of the WHO-recommended algorithm for TB screening increase the proportion of subjects screened for TB, reduce the proportion of PLHIV who develop TB disease during IPT, and reduce mortality during TB treatment, compared with current policy or more intensive TB screening strategies (e.g. microbiologic evaluation of all PLHIV)? Specific attention should be paid to studying this question in different contexts in which TB screening might occur, such as HIV counselling and testing centres, HIV clinics, community-based case-finding, and household contact investigations. The following questions need to be addressed:

- What is the best model to eliminate diagnostic delay, hasten treatment initiation for TB and reduce mortality, using existing tools including the effectiveness of the revised WHO algorithm for smear-negative TB among HIV-infected TB suspects?

- What are the best strategies to promote and scale-up integrated screening of HIV infection and TB infection and disease among household contacts of HIV-infected TB patients?

- What are the best operational models for active and enhanced case-finding of TB among PLHIV in HIV service facilities and at community level, in both high and low HIV prevalence settings?

3.3 Optimizing timing of IPT in relation to ART

In PLHIV eligible for both IPT and ART, what is the appropriate time to initiate IPT, and what is the optimal duration, safety, efficacy and cost-effectiveness of IPT – alone or with ART – in reducing the risk of active TB and mortality compared to ART alone among PLHIV, particularly under programme conditions?

3.4 Models to improve adherence to IPT

In PLHIV initiating IPT, what models of medication delivery, clinical monitoring and community support reduce rates of default during IPT, reduce the incidence of breakthrough TB, and reduce the occurrence of severe adverse events? What are the best operational models to scale-up IPT in HIV care settings, including frequency of symptom screening, monitoring tools and measures to maintain high adherence? The following questions need to be addressed:

- What are the best strategies to obtain optimal medication delivery, community support and clinical monitoring of IPT in PLHIV in order to maximise adherence?

- How to best model and forecast the operational requirements and full economic costs of going to scale in administering IPT in HIV care settings?

3.5 Optimizing infection control to reduce TB transmission

In HIV care and treatment settings, does a standardized package of infection control (IC) interventions reduce nosocomial TB transmission compared with current policy and practices? Do selected IC interventions reduce nosocomial TB transmission? For both of these questions, the best recognized indicator to measure (in operational research studies) is TB infection rates in health-care workers. With this in mind, the following operational research challenges need to be addressed:

- What are the best infection control interventions that effectively reduce *M. tuberculosis* transmission (both drug-susceptible and -resistant) in health care-settings, at home and in the community?

- What are the best operational models, i.e. practical, feasible, easily reproducible and effective, to implement and monitor infection control measures in health facilities?

- What are the best operational models to assess the impact of infection control measures in reducing the spread of *M. tuberculosis* to HIV-infected adults and children?

3.6 Reducing mortality in co-infected patients

In PLHIV treated for TB, what factors are associated with death during TB treatment and, among those who die, what are the most common causes of death? Although use of ART and co-trimoxazole have been clearly documented to reduce mortality during TB treatment, specific attention should be paid to identifying whether the absence of such treatment is responsible for ongoing, high mortality rates or whether additional, modifiable risk factors can be identified.

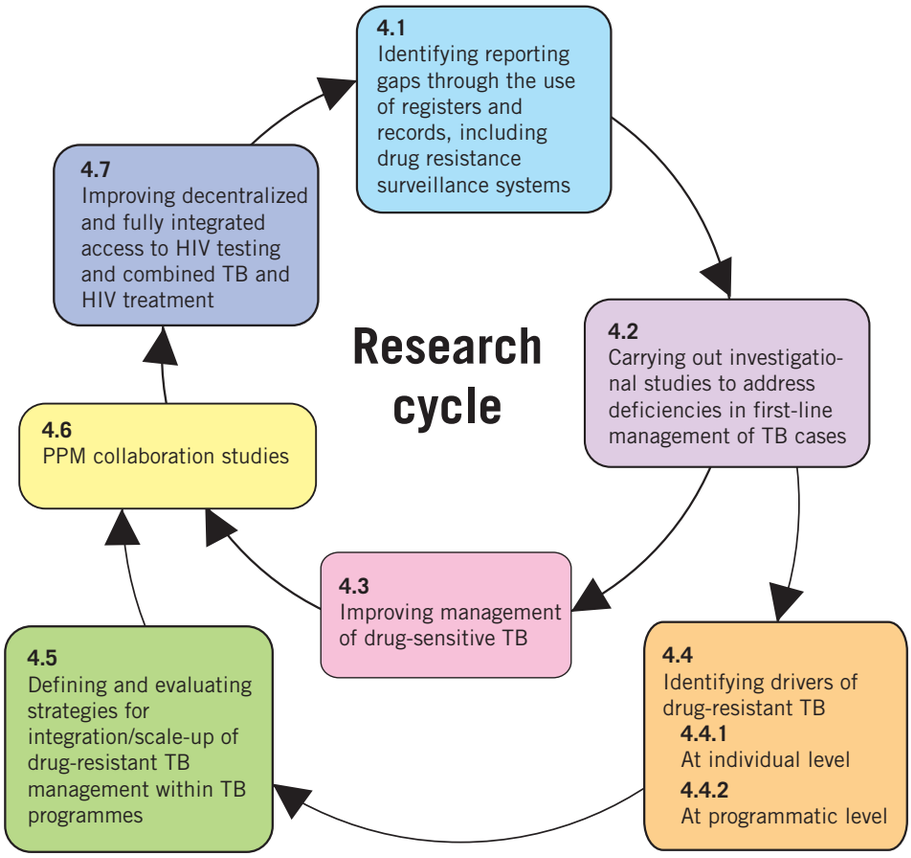
4 Treatment of drug-susceptible and M/XDR-TB: optimal access, delivery and community participation

Access to health care is the cornerstone of TB control programmes, which must ensure that all detected patients receive a full course of treatment *(29)*. This includes establishing effective treatment as well as effective strategies to support the process of care from detection of disease through the completion of appropriate treatment *(30)*. Limited access and poor adherence to treatment remain, however, major obstacles in the global fight against TB. In 2008, 39% of TB cases were not reported or detected, 93% of MDR-TB cases were not diagnosed and even more were not treated with an approved Green Light Committee (GLC) regimen, and 93% of HIV-infected TB patients were not started on ART *(25)*. Patient and health-system factors contribute to these problems. Major diagnostic delays compromise treatment outcomes and increase TB transmission. In 2008, the treatment success rate for new patients with smear-positive pulmonary TB was 87%, but this apparent success masked a number of operational challenges: for MDR-TB, only 30 000 (7%) out of the 440 000 (95%CI 390,000–510,000) MDR-TB cases estimated to have emerged in 2008 globally, were notified, and nearly 6000 of them (1.4%) were put on treatment *(31)*. Treatment success of a cohort of 4500 MDR-TB patients treated under programmatic conditions in 2004–6 was 60%. Access to treatment for MDR-TB remains one of the major problems facing TB control today. Only a small fraction of the tens of thousands of diagnosed patients are receiving appropriate care. Even among the 72 000 patients approved for GLC treatment, only 19 000 have actually been enrolled. Operational research is thus critically needed to improve access to care for drug-susceptible (DS) and drug-resistant (DR) TB patients. We propose here a cycle of research that addresses both DS- and DR-TB aspects, also taking into account HIV co-infection, with actions that can be conducted in parallel, arising from a common basis of core activities i.e. identification of the reporting gaps and identification of deficiencies in first-line management of TB cases.

Priority areas for operational research

The outstanding research questions have been prioritized to facilitate operational research that leads to improved access and delivery of treatment of drug-susceptible and M/XDR tuberculosis (see **Figure 5**).

FIGURE 5: CYCLE OF RESEARCH ACTIVITIES FOR IMPROVED ACCESS AND DELIVERY OF TREATMENT OF DRUG-SUSCEPTIBLE AND M/XDR TUBERCULOSIS



4.1 Identifying reporting gaps

Reporting gaps exist at several levels and give false estimates of TB cases being detected and under treatment. For instance, it has been reported that 5 to 15% of patients with smear-positive pulmonary TB are not entered into TB registers and fail to be counted *(32)*. It is therefore important to determine treatment outcomes of new smear-positive pulmonary TB at different sources: comparing sputum laboratory registers with TB patient registers, for example, or comparing TB treatment outcomes by treatment cards, registers and quarterly reports. This should also be done for patients enrolled in TB re-treatment regimens. In addition, reporting gaps also arise from the lack of registration of TB patients detected and treated in the private sector or in public sectors outside the NTCPs, and efforts should be made to count them in (see **section 2**, above).

4.2 Carrying out investigational studies to address deficiencies in first-line management of TB

Avoiding missed treatment doses is beneficial, since not all who receive irregular treatment will default but may be at high risk for acquired drug resistance. Similarly, defaulting from treatment may contribute to drug resistance. Efforts are needed, under programmatic conditions, to reduce irregular adherence or default from treatment; and to reduce selective pressure for resistant organisms among patients who take treatment irregularly, but do not default. Studies to understand reasons behind default (both primary and secondary), poor adherence, missed doses, and drug stock-outs will help inform strategies to reduce each of these challenges, and reduce the risk of drug resistance.

4.3 Improving management of drug-sensitive TB

Strategies should be developed, piloted, and rolled out to improve first-line management of drug-sensitive TB, based on the findings generated in 4.2 above.

4.4 Identifying drivers for drug-resistant TB

Studies conducted under 4.2 (above) will also provide information that will help identifying the various risk factors that contribute to the development of drug-resistant TB in the local or national context. One way of identifying these risk factors would be to investigate unexpected increases in numbers of MDR-TB patients, as observed through routine surveillance of multidrug resistance among risk groups, especially previously treated patients. Such investigations can provide important information as to which factors, operating at either individual patient level or at the health service (programmatic) level, are important drivers of the local/national increase in MDR-TB cases. Of particular importance are risk factors that are amenable to intervention. These factors will be investigated both at *individual* and *programmatic* level.

4.5 Defining and evaluating strategies for integration/scale-up of drug-resistant TB management within TB control programmes

From the studies outlined above, optimal strategies can be developed to identify and target those patients most at risk of developing, or having, MDR-TB. This includes evaluating the operational steps and time required to introduce new diagnostics for identification of drug resistance, as well as optimal delivery of treatment in the context of programmatic management of DR-TB (for instance, models for maximizing adherence to treatment): What are the best strategies to scale-up DR-TB management into TB control programmes in relation to provision of 2nd-line treatment (e.g. inpatient vs. ambulatory treatment vs. community-based care, use of incentives and enablers to enhance adherence to treatment, social support, community involvement etc.)? It should be noted that different methods will require different technologies/infrastructure, but many processes will likely be identical.

4.6 PPM collaboration studies

Previous questions will address the role of private practitioners in the detection, diagnosis and treatment of both DS- and DR-TB. To achieve this, it is necessary to get realistic data on numbers of DS- and DR-TB cases treated in the private sector and about treatment outcomes and DR rates among those who failed or relapsed. Why do people use the private sector rather than the public sector? How to engage with the private sector for the treatment of DS- and DR-TB?

4.7 Improving decentralized and fully integrated access to TB and ART treatment

Previous sections addressed the challenges of detecting and treating both DS- and DR-TB most efficiently. In addition, this has to be contextualized in relation to the HIV epidemic, as all TB patients, whether drug-sensitive or drug-resistant, should start ART early if they are co-infected with HIV, and the provision of both TB and ART drugs should be in the same facility. There are similar questions for both conditions: How can joint treatment be provided at health centres? How can communities be better engaged (structures, support, links with traditional systems)? Should TB programmes have their own stock of tests, co-trimoxazole and/or ART? Should they place orders to HIV programmes on behalf of patients receiving TB treatment, or should they have to refer patients to HIV programmes for all the above?

In addition, there are specific operational and epidemiological questions related to HIV-infected DR-TB, such as the precise causes of HIV-driven MDR-TB outbreaks if these occur, and the potential of IPT for inducing isoniazid-resistance and MDR-TB in the long run.

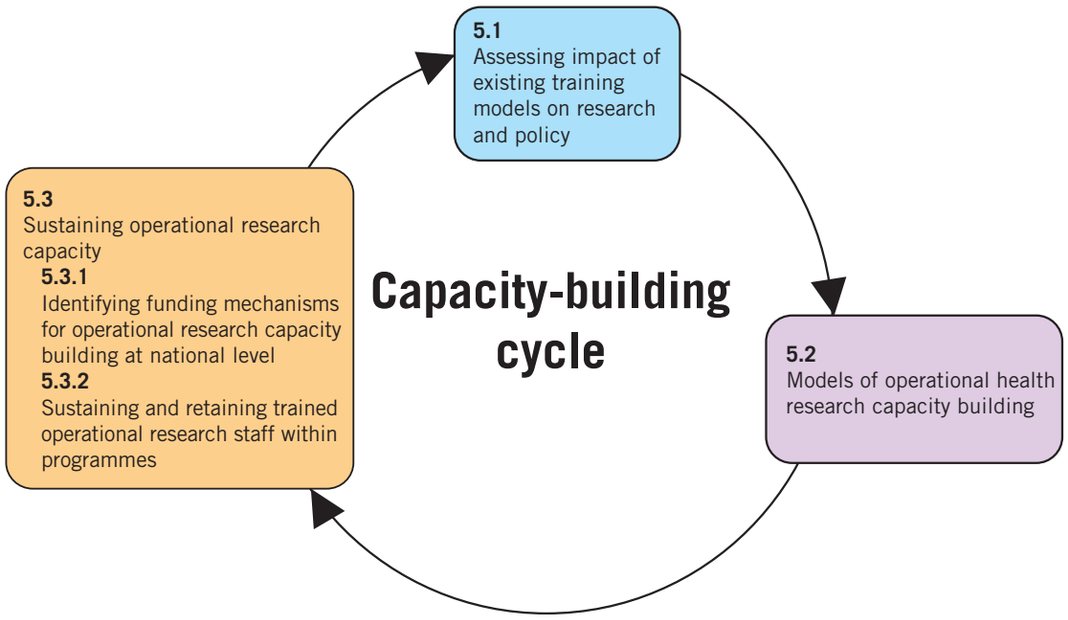
5 Capacity building for operational research

Despite international interest in the subject of operational research, very little research is conducted or published from resource-limited settings where the greatest burden of tuberculosis resides. Yet major funders such as the Global Fund explicitly state that up to 10% of country proposal budgets should include monitoring, evaluation and operational research *(33)*. There is a widespread consensus that operational research is important at local/national level to improve programme performance and at international level to guide policy recommendations *(34)*. Questions remain, however, on the issue of where operational research should be located (e.g. NTCPs, governmental institutions, universities, NGO, or partnerships?), on the nature of research capacity to develop and approaches to capacity building, as well as on the evaluation of outcomes of training *(34)*. NTCPs may lack essential expertise, infrastructure, staff, funds, policy cycle, and/or professional culture, and there may be only weak linkages between programmers and researchers. The key aspects of capacity building/strengthening in operational research at programme levels are the following: (i) operational research should be embedded in the NTCPs strategic plan; (ii) there should be an operational research focal point in programmes; (iii) operational research projects should conclude with clear results to alter/improve programme performance (see **Box 2**). It is particularly important to be aware of the distinction between *capacity building* (providing the ability to individuals, organizations or systems to perform and utilize health research effectively, efficiently, and sustainably) *(35)* and *training* (which is an organized activity aimed at imparting information and/or instructions to improve the recipient's performance or to help him/her attain a required level of knowledge or skill).[g] Training may be one component of capacity building that should encompass the generation of research agendas/topics as well as the uptake and utilization of research outputs. It can be helpful to envisage these as extremes of a spectrum with training in its simplest form at one end of the spectrum, most of the time unlinked to outputs such as peer-reviewed publications that enable an individual or organization to utilize health research. At the other end of the spectrum a country-based capacity-building programme could include direct participation of policy-makers who should directly utilize health research outputs. Between these two extremes, there may be training programmes delivering some elements of capacity building (such as peer-reviewed publications) but without being sufficiently embedded within a national system to deliver full changes to the functioning of an organization or system as a result. As a starting point, it might be helpful to use the needs assessment conducted with health professionals working with TB, TB/HIV, MDR-TB, who voluntarily complete locally-adapted questionnaires on preferred learning modes, research training needs, such as laboratory, clinical, epidemiology, biostatistics/data analysis, social science research, bioethics, and skills desired, such as grant writing and data management. This approach could be helpful to identify the further steps for the capacity building for operational research locally or nationally.

[g] See: http://www.businessdictionary.com/definition/training.html

Priority areas for capacity building for operational research

FIGURE 6: CYCLE OF RESEARCH ACTIVITIES FOR IMPROVED CAPACITY BUILDING FOR OPERATIONAL RESEARCH IN TB



5.1 Assessing the impact of existing training models in terms of products and measurable outcomes

Although there is a substantial literature available on different approaches to educational evaluations, few innovative interventions for work-based education of health professionals in developing countries have been adequately evaluated *(37, 38)*. It is therefore important to evaluate the impact of existing models in terms of products and outcomes, such as the number and type of publications, the impact indicators for policy and practice, or the users' satisfaction (e.g. patient, health-care workers, managers or community leaders). It is also important to track what happens to persons who attended training courses and whether they continue with research after the training and continue to undertake operational research and use this research to inform policy and practice. Accurate and reliable indicators of impact need to be defined to assess the quality of operational research projects and the capacity for translation into evidence-based programme practice. Examples of potential indicators are given in **Annex VI**.

5.2 Models of operational health research capacity building

Several models of operational research capacity building have been implemented, including country-specific experiences, and are detailed in **Annex VI**. Lessons learnt to date are summarized in **Box 2**. Prospective case study methodology can be used to learn further lessons from other models as they are introduced (*36*).

5.3 Sustaining operational research capacity at the national level

5.3.1 Identifying funding mechanisms for operational research capacity building at national level

What sort of funding mechanism (beyond Global Fund budgets) can allow operational research capacity to be built and sustained at the national level? Examples in the literature include money from central government training budgets, contributions from project participants towards the cost of

courses, income from selling consultancy services, and externally funded research grants (*36*).

5.3.2 *What are possible ways of sustaining and retaining trained operational research staff within programmes?*

Along with ongoing funding, sustained staff retention is another indicator of successful operational health research capacity building, and lessons for retaining and sustaining trained staff should be extracted from the evaluation described in **section 5.1** (above). In the literature *(36)*, indicators of sustainable capacity building increase in complexity as activities mature and include

- early engagement of stakeholders; explicit plans for scale up; strategies for influencing policies; quality assessments (*awareness and experiential stages)*
- improved resources; institutionalisation of activities; innovation *(expansion stage)*
- funding for core activities secured; management and decision-making led by southern partners *(consolidation stage)*.

BOX 2: TEN KEY ENABLING FACTORS FOR CAPACITY BUILDING/STRENGTHENING IN OPERATIONAL RESEARCH AT PROGRAMME LEVEL (FROM REFERENCE *34*)

(i) Operational research should be embedded in a national TB programme strategic plan;

(ii) In each national TB programme, there should be an operational research focal point (supported by other field staff) that supports the programme manager and who coordinates and sets the national research priorities;

(iii) Programme staff engaged in operational research should be encouraged and motivated through 'on-the-job' training and supervision, with dedicated time and opportunity for research activities, provided with opportunities to make presentations at national and international conferences, research bonuses and small grants;

(iv) There should be adequate operational research infrastructure (e.g. room space, computers, internet, stationary) and implementation support (e.g. motorcycles);

(v) At the country level, there should be a shift towards a 'partnership model' in operational research that is inclusive of academic institutions, NGOs and community-based associations, so that the comparative advantages of each group are harnessed.

(vi) Operational research training should be based on strict selection criteria, must be output oriented, and should involve a strong element of mentorship. The operational research training model should be practical and geared towards providing practical skills for both conducting and publishing research (e.g. the UNION-MSF operational research approach to training);

(vii) There should be long-term career opportunities at programme level through operational research fellowships (junior and senior);

(viii) Operational research projects should conclude with clear results that can influence or alter/improve programme performance;

(ix) Funding and resources for operational research need to be built into the programme so as to avoid foreign or academic institutions developing a monopoly on funding, time and mandate for research and consequently, the associated power of decisions;

(x) Programme researchers should be represented on the Global Fund Country Coordinating Mechanism so that the rationale for and decisions on research funding are made at the highest level.

REFERENCES

1. Guillerm M. *Tuberculosis Diagnosis and Drug Sensitivity Testing: an Overview of the Current Diagnostics Pipeline.* Campaign for Access to Essential Medicines. Paris, Médecins Sans Frontières, 2006.
2. Perkins M and Cunningham J. Facing the Crisis: Improving the Diagnosis of Tuberculosis in the HIV Era*. Journal of Infectious Diseases*, 2007;196: S15–27.
3. Pai M, Minion J, Steingart K and Ramsay A. New and improved tuberculosis diagnostics: evidence, policy, practice, and impact. *Current Opinion in Pulmonary Medicine*, 2010; 16(3):271–84.
4. Van Kampen SC, Ramsay AR, Anthony RM and Klatser PR. Retooling National TB Control Programmes (NTPs) with New Diagnostics: The NTP Perspective. *PLoS ONE*, 2010; 5(7):e11649.
5. *Guidelines for intensified tuberculosis case finding and isoniazid preventive therapy for people living with HIV in resource constrained settings*. Geneva, World Health Organization, 2010.
6. Corbett EL et al. Comparison of two active case-finding strategies for community-based diagnosis of symptomatic smear-positive tuberculosis and control of infectious tuberculosis in Harare, Zimbabwe (DETECTB): a cluster-randomised trial. *The Lancet*, 2010; 376(9748):1244–53.
7. Datiko DG and Lindtjorn B. Health extension workers improve tuberculosis case detection and treatment success in southern Ethiopia: a community randomized trial. *PLoS ONE*, 2009; 4(5):e5443.
8. Getahun H and Maher D. Contribution of 'TB clubs' to tuberculosis control in a rural district in Ethiopia. *International Journal of Tuberculosis and Lung Disease*, 2000; 4(2):174–8.
9. Sekandi JN, Neuhauser D, Smyth K and Whalen CC. Active case finding of undetected tuberculosis among chronic coughers in a slum setting in Kampala, Uganda. *International Journal of Tuberculosis and Lung Disease*, 2009; 13(4):508–13.
10. Mann G et al. Beyond accuracy: creating a comprehensive evidence base for TB diagnostic tools. *International Journal of Tuberculosis and Lung Disease, 2010;* 14 (12):1518–1524
11. Pronyk PM et al. Active case finding: understanding the burden of tuberculosis in rural South Africa. *International Journal of Tuberculosis and Lung Disease*, 2001; 5(7):611–8.
12. Shargie EB, Morkve O and Lindtjorn B. Tuberculosis case-finding through a village outreach programme in a rural setting in southern Ethiopia: community randomized trial. *Bulletin of the World Health Organization*, 2006; 84:11–119.
13. Kemp J et al. Is tuberculosis diagnosis a barrier to care? *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 1996; 90:472.
14. Chandrasekaran V, et al. Factors leading to tuberculosis diagnostic drop-out and delayed treatment initiation in Chennai, India. *International Journal of Tuberculosis and Lung Disease,* 2005; 9(S10):172.
15. Squire SB et al. Lost smear-positive pulmonary tuberculosis cases: where are they and why did we lose them? *International Journal of Tuberculosis and Lung Disease,* 2005; 9(1):25–31.
16. Botha E, et al. Initial default from tuberculosis treatment: how often does it happen and what are the reasons? *International Journal of Tuberculosis and Lung Disease,* 2008; 12(7):820–3.
17. Khan MS, et al. Default during TB diagnosis: quantifying the problem. *Tropical Medicine and International Health*, 2009; 14:1437–41.
18. Uplekar M. Involving private health-care providers in delivery of TB case: global strategy. *Tuberculosis* (Edinb), 2003; 83(1-3): 156–64.
19. Kemp JR et al. Can Malawi's poor afford free tuberculosis services? Patient and household costs associated with a tuberculosis diagnosis in Lilongwe. *Bulletin of the World Health Organization*, 2007; 85(8):580–5.
20. *A tool for National Situation Assessment - Public–private mix for TB care and control* (WHO/HTM/TB/2007.391). Geneva, World Health Organization, 2007.

21. *Priority research questions for TB/HIV in HIV-prevalent and resource-limited settings* (WHO/HTM/TB/2010.8). Geneva, World Health Organization, 2010.

22. *WHO Three I's meeting. Intensified case finding, isoniazid preventive therapy and TB infection control for people living with HIV. Report of a joint World Health Organization HIV/AIDS and TB department meeting (2-4 April 2008).* Geneva, World Health Organization, 2009.(http://www.who.int/tb/publications/2009/who_3Is_meeting_report.pdf, accessed 7 January 2010).

23. Woldehanna S and Volmink J. Treatment of latent tuberculosis infection in HIV infected persons. *Cochrane Database Systematic Review,* 2006; (3): CD000171.

24. Aït-Khaled N et al. Isoniazid preventive therapy for people living with HIV: public health challenges and implementation issues. *International Journal of Tuberculosis and Lung Disease,* 2009; 13(8): 927–35.

25. *Global tuberculosis control: a short update to the 2009 report* (WHO/HTM/TB/2009.426). Geneva, World Health Organization, 2009.

26. Golub JE et al. Isoniazid preventive therapy, HAART and tuberculosis risk in HIV-infected adults in South Africa: a prospective cohort. *AIDS*, 2009; 23: 631–636.

27. Getahun H et al. Implementation of Isoniazid preventive therapy for people living with HIV worldwide: barriers and solutions. *AIDS*, 2010; 24 (suppl 5): S57–S65.

28. Sester M et al. Challenges and perspectives for improved management of HIV/TB coinfection. *European Respiratory Journal,* 2010; 36:1242–1247.

29. Lönnroth K et al. Tuberculosis control and elimination 2010-50: cure, care, and social development. *The Lancet*, 2010; 375(9728): 1814–29.

30. Lienhardt C and Rustomjee R. Improving tuberculosis control: an interdisciplinary approach. *The Lancet,* 2006;367: 949–50.

31. Multidrug and extensively drug-resistant TB (M/XDR-TB): 2010 Global Report on Surveillance and Response (WHO/HTM/TB/2010.3). Geneva, World Health Organization, 2010.

32. Botha E et al. Initial default from tuberculosis treatment: how often does it happen and what are the reasons? *International Journal of Tuberculosis and Lung Disease,* 2008; 12(7):820–3.

33. Xueref S. Operations research in the Global Fund: status and way ahead, in collaboration with partners. Presentation at the 5th International AIDS Society Conference on HIV Pathogenesis, Treatment and Prevention. Cape Town, South Africa, 2009.

34. Zachariah R et al. Operational research in low-income countries: what, why, and how? *Lancet Infectious Diseases,* 2009; 9:711–17.

35. Bates I et al. Evaluating health research capacity building: an evidence-based tool. *PLoS Medicine*, 2006; 3(8):e299.

36. Bates I et al. Indicators of sustainable capacity building for health research: analysis of four African case studies. *Health Research Policy and Systems,* 2011, 9:14doi:10.1186/1478-4505-9-14.

37. Cantillon P. Evaluation: beyond the rhetoric. *Journal of Evaluation in Clinical Practice*, 1999; 5(3):265–8.

38. Prideaux D. Researching the outcomes of educational interventions: a matter of design. RTCs have important limitations in evaluating educational interventions. *BMJ*, 2002; 324(7330):126–7.

IV. COMMUNITY INVOLVEMENT IN TUBERCULOSIS RESEARCH

Introduction

While communities at risk have been both drivers and partners in HIV research (1), their important role in TB research is yet to be fully realized. Involvement of communities in tuberculosis care and prevention is currently on the international agenda (2, 3). This creates opportunities and indicates the urgency to also engage communities in TB research.

In this section, the advantages of involving communities in operational research for TB care and control are presented, and strategies for engaging risk groups, (ex-)patients and high burden communities are proposed. We intend to provide researchers with some tools and ideas to help them collaborate effectively with communities in the identification of research priorities, the conduct of research projects and dissemination of the results.

'Community involvement' is more than the token inclusion of a former TB patient on staff or the setting up of a Community Advisory Board (CAB) for a clinical trial. It implies a fundamental shift in the way we think about and implement studies. Community involvement[h] in research is founded on the core principle that people who are affected by research have a right to have a say in what is being studied and how, and ensure that they can take advantage of the benefits of research findings.

In addition to being a legal requirement in some countries (e.g. Brazil, Namibia, Kenya, Zimbabwe), capacity building and collaboration with members of risk groups enhances the quality of the research, because they are the experts on the challenges they experience and the solutions they require (*4*). The aim of community involvement is to recognize the public's stake in determining the way research is formulated, commissioned, undertaken, disseminated and translated into policies and service delivery. Many TB researchers are embracing a TB version of the GIPA principle (Greater Involvement of People Living with HIV/AIDs) called GIPT (Greater Involvement of People with TB) to reflect their belief that TB patients have been neglected stakeholders for too long.[i]

'Communities at risk' refers to both vulnerable groups such as the malnourished or people with diabetes as well as residents in geographical areas with a high TB prevalence. Some communities share a set of social, cultural, racial or linguistic characteristics that unite them. However other 'communities' are simply identified by a common risk behaviour or shared geography such as injection drug users, prisoners, smokers, miners, or laboratory personnel. Some refer to 'Most At Risk Populations (MARPs)' to describe communities whose risks for TB are much higher than the general population. Often these communities are hidden due to stigmatization, and extra efforts and expertise are required to reach and to serve them well.

The following objectives of community involvement have been proposed (*4*):

- To ensure the relevance of research,
- To assess and assure that research is culturally and practically acceptable in the context it is intended,
- To ensure that community disruption by the research project is minimized,
- To avoid injustice, by ensuring a fair distribution of the benefits of research,
- To take into account the ethical hazards that may be part of the social, economic, and political landscape of the community (e.g. corruption),
- To ensure that research practices are in tune with local norms and values.

The requirements for successful community involvement include: a long-term commitment, an appreciation of the potential synergies, and a willingness to cross-train and support all the people involved (6).

One comprehensive model of collaboration is termed 'community-based participatory research' (CBPR). While, there are many definitions for CBPR, one definition is "a collaborative approach to research that combines methods of inquiry with community capacity-building strategies to bridge the gap between knowledge produced through

[h] The terms 'community engagement' or 'participation' are considered synonyms for 'community involvement'.

[i] http://www.worldcarecouncil.org/content/gipt-principles-new-driver-road-stop-tb

research and what is practiced in communities to improve health" (*7*). In CBPR projects, the community participates fully in all aspects of the research process.

Types of community involvement in operational research

Communities can play key roles in many different activities in the course of a research project:

(i) Priority-setting

(ii) Research design

(iii) Ethical evaluation

(iv) Protection of study participants

(v) Data collection

(vi) Interpretation of findings

(vii) Dissemination of findings

(viii) Translation into action

Community involvement in each aspect of research is considered in more detail below.

(i) Priority-setting

The public and patients often have a more holistic view of health and wellbeing than disease-specific researchers. This may result in priority-setting that differs from that of health workers or researchers, e.g. more geared towards cross-cutting issues, quality of life, and social determinants. For example, while trialists may narrowly focus on improved treatment regimens for TB, communities may see the need to answer broader primary prevention questions such as how to address the structural conditions (e.g. poverty, malnutrition, marginalization) that put them at risk for a range of diseases.

The research questions that have been prioritised by experts in this document will be addressed at country level. During this country-level research process, it is recommended that at-risk communities are involved to confirm their local relevance and priority. Various methods have been tested to involve the community (often ex-patients) in priority-setting. An example is the 'Dialogue Model' developed by Abma et al. for patient participation in health research agenda-setting (*8*). This model has six phases: *exploration, consultation, prioritization, integration, programming and implementation.*

A key question is: Who should be responsible for involving the community in identifying priority research questions? Researchers may not have the knowledge or funding to organize community involvement and/or consultation. Furthermore, they often work on one specific dimension of TB and may not be sufficiently knowledgeable of the whole field to respond to community queries. Donor agencies increasingly seek guidance from most-at-risk-populations when they prepare a new research programme or call for proposals. Also organizations of communities at-risk may take the lead, and may use the resulting, prioritized research agenda for advocacy purposes. Researchers should be aware that ex-TB patients and leaders of at-risk groups may not always be representative of all TB patients or people at-risk. Governmental organizations may also take a role especially in priority-setting that transcends one vertical disease program.

(ii) Research design

Reasons for involving the community in study design can be to increase participation and collaboration of the community; to improve the quality of research design; to ensure that study questionnaires pose questions in ways that are understandable and acceptable; to ensure that issues important to the community are included in the study design, etc. At-risk communities can play a pivotal role in developing recruitment and retention strategies or questionnaire design. When there are different study design options, community stakeholders can weigh in on which makes most sense locally.

Communities that may appear homogeneous from the outside may in fact be very diverse. Understanding the heterogeneity, social codes, and hierarchies often requires an insider's perspective. For example research with prisoners requires an insider's knowledge of the parallel power structures that govern inmates' participation in research endeavours (*9*). Acknowledging the heterogeneity of communities also means assuring that vulnerable sub-groups are heard, not just the main power brokers and gatekeepers.

There are several possibilities for involving at-risk communities in research design: community members might be members of the steering committee or technical advisory committee of research studies, or community members might work in equal partnership with the researchers and other stakeholders on the research project. When involving (former) TB patients and members of stigmatized groups (men who have sex with men, ex-prisoners, sex workers, undocumented migrants, etc.) as researchers, it is important to respect their right to control to whom and how they disclose their member status.[j]

A suitable way of involving communities is through the establishment of community advisory boards (CABs) - see the example of the Global Alliance for TB Drug Development.[k] These boards are often comprised of community members involved in local leadership, social services or health care delivery. They provide input on trial implementation and work towards greater community understanding of TB and the research process. Thus, a CAB established in KwaZulu-Natal, South Africa, contributed to promoting and facilitating relevant research and TB control activities (*10*). TB drug and vaccine trialists have engaged community leaders, traditional birth attendants, teachers, religious leaders, and parents in high-burden regions of Uganda, Kenya, and Mozambique in conversations about the potential need for invasive TB diagnostic procedures in children to anticipate and strategize ways to assuage their concerns in the pre-trial design phase (*11*).

Participatory research and evaluation implies early and sustained involvement by target groups in assessing the outcomes of activities designed to help them. It is based on the premise that communities are ideally positioned to help national TB programmes and NGOs to do the work as well as to measure the outcomes. Participatory evaluation often involves the use of graphics, maps and interactive techniques, particularly where low literacy and language diversity may limit the utility of written measurement tools.

(iii) Ethical evaluation

Research protocols must be submitted for consideration, comment, guidance and approval to a research ethics committee before the study begins (according to Article 15 of the World Medical Association Declaration of Helsinki).[l] The *Operational Guidelines for Ethics Committees that Review Biomedical Research*,[m] published by the WHO/Special Programme for Research and Training in Tropical Diseases (TDR), suggests that an ethical committee should include a lay person. An ethical committee should include at least one member whose primary area of expertise is in a non-scientific area. The Council for International Organizations of Medical Sciences presumes that lay people qualified to represent the cultural and moral values of the community, are included in the ethical committee.[n] They ensure that the rights of the research participants are respected. The Family Health International Office of International Research Ethics has developed a curriculum to empower community representatives through training and education to act as a competent voice for research participants worldwide: Research Ethics Training Curriculum for Community Representatives.[o]

(iv) Protection of study participants

The Declaration of Helsinki states that in medical research involving competent human beings, each potential participant must be adequately informed of the aims, methods, sources of funding, any possible conflicts of interest, institutional affiliations of the researcher, the anticipated benefits and potential risks of the study and the discomfort it may entail, and any other relevant aspects of the study. This is generally done through obtaining informed consent from the individual to contribute to the trial/study.

Involving community members in the design of the informed consent process can ensure that potential participants have a better grasp of the real risks and benefits of a study. While scientists may know the physical risks of study participation,

[j] See: http://www.aidsalliance.org/TechnicalThemeDetails.aspx?Id=34

[k] See: http://www.tballiance.org/aaa/endemic.php

[l] See: http://www.wma.net/en/30publications/10policies/b3/index.html

[m] See: http://apps.who.int/tdr/publications/training-guideline-publications/operational-guidelines-ethics-biomedical-research/pdf/ethics.pdf

[n] See: http://www.cioms.ch/publications/layout_guide2002.pdf

[o] See: http://www.fhi.org/en/RH/Training/trainmat/ethicscurr/retccr.htm

community members are better positioned to understand and mitigate the potential social harms (e.g. stigmatization, loss of employment, etc.) that can come from joining certain types of research. The community can also have a role in monitoring the procedures for obtaining informed consent. For example, is the informed consent process implemented as rigorously as was planned? History shows that community activist groups may oppose research projects that bypass them in the ethical review process. For example, two tenofovir trials were suspended due to the failure to engage activists in compensation issues (*12*).

The International AIDS Vaccine Initiative[p] has developed an AIDS vaccine literacy toolkit (Vaxlit). One of the positive outcomes of the vaccine literacy initiative is that potential trial volunteers make informed, independent decisions regarding participation.

(v) Data collection

Involving community members in the data collection phase may have advantages or disadvantages, depending on the kind of information that is being collected. Local expertise can prevent misunderstandings and increase participation and retention because at-risk groups are more sensitive to local social norms and the negative connotations that some TB terminology can have (e.g. 'suspects' implies suspicious and untrustworthy in some settings). Involvement of community members in data collection may ensure that the intentions of the proposed research projects are honestly portrayed.

Also if community members are involved in data collection they might be better able to interpret results than external researchers. If problems arise during the course of a study, community members involved in the undertaking of research can troubleshoot the problems at an early stage. In some areas (e.g. slums) security can be an issue. Involving community members might increase the security of the research teams.

Potential disadvantages of involving community members in research can arise when TB research delves into sensitive or stigmatized issues – such as non-compliance with infection control norms, non-use or misuse of medicines, etc. Deductive disclosure can be a problem if community members are not sufficiently aware of research ethical principles. Since community members may share cultural assumptions and terminology of the participants, they might not always probe for the meaning of local concepts when necessary because they may take certain ideas for granted or because the topic is too sensitive.

(vi) Interpretation of findings

Rigorous qualitative methodology involves frequent confirmations with community members to prevent any misinterpretation and enhance the '*groundedness'* (foundation of an argument, a belief, or an action; a basis) of findings. Involving community members in the interpretation of quantitative findings might also result in more applicable conclusions and recommendations.

(vii) Dissemination of findings

The findings of research should be disseminated with and within the communities that participated in the research (see Declaration of Helsinki, above). This could be done using different methods: workshops; radio talk shows; community debriefings; local news media; graphical and visual techniques; during community meetings etc. Community members can have an important role in the dissemination of the research findings. All materials used for dissemination should use non-scientific language that is easily understood by a general audience.

(viii) Translation into action

When communities at-risk have a sense of ownership of research findings, they can play a powerful advocacy role in assuring that research innovations translate into service delivery realities long after researchers have departed. Community activists can use "the facts" to assure that new guidelines and policies are implemented and life-saving technologies are rolled out and brought to scale.

[p] See: http://www.iavi.org/

Other aspects of community involvement

Building capacity for involvement

As interest in community involvement in research has grown, so has the need for research training and support to ensure good practice in this area. At the start of each project the training needs of relevant community members should be assessed: Do they need enabling skills and/or research skills? Examples of training for community involvement in research are available (10). Of note, TB investigators may often need training (e.g. communication skills, cultural competency, participatory methods) as well.

Financing community involvement

Appropriate involvement of the community in research requires an adequate budget, e.g. for community training, and for meetings with the community. Adequate resources for community involvement should be an integral part of the overall research budget. Donors and regulators have key roles to play in fomenting effective collaboration with at-risk communities and have made this a requirement for some research funding streams and regulatory processes.

How to Engage Communities?

The process of involving at-risk communities depends on the type of research intended. Often roles cannot be decided *a priori*, but should be mutually determined with community members. It is most strategic if communities are involved from the outset of research planning, but it is better to reach out during the study than not at all. The more hidden and stigmatized the risk group, the longer it may take to engage them and study timelines should reflect this. Here are the steps followed by TB vaccine trialists in the TBVACSIN network.

Step 1: Initial consultation with stakeholders.

Given that communities are rarely homogenous and may comprise sub-groups with divergent interests and power, it is usually strategic to start with an array of stakeholder meetings where research ideas and priorities can be vetted and the social landscape can be mapped. These consultations are listening exercises to identify sensitivities and to discern the full range of potential collaborators. Be careful not to create unrealistic expectations or make promises you may not be able to fulfil.

Step 2: Recruit and train study staff from communities at-risk

Once the funding and the objectives are confirmed, identify individuals with the local roots to contribute to study design and implementation. Be sensitive to political, social, gender, and ethnic divisions and strive to form a diverse team to bridge divides. Be prepared to invest in human resource development and build this into budgets and study timelines. Staff are the ambassadors of the study and their behaviour reflects on the study when at work and after hours.

Step 3: Form an effective Community Advisory Board

Gather stakeholders from a diverse range of grassroots organizations and community groups and train them in the roles of a CAB. Budget for transportation reimbursement or other modest recognition of the contribution of community members. Beware of creating expectations of payment or allowing misunderstandings and potential accusations of undue inducements or lack of independence. Gather their insights on how to protect participants from social and physical harms, how to ensure high rates of participation without coercion, and how to frame the risks and benefits of participation.

Step 4: Pilot study tools and fidelity to procedures

Conduct focus group discussions with potential participants to ensure that the research tools are understandable as designed. When in doubt, conduct back translations of instruments to avoid misinterpretation. Directly observe role plays of study procedures to ensure that that study staff uphold research ethics principles in their interactions.

Step 5: Do the research together

Invite communities to collect data and involve them in the process to evaluate, refute, refine and interpret the findings.

Step 6: Share the credit for good work

Cultivate a broad ownership and accountability for the dissemination and translation of the results by involving community leaders in the "marketing" of the findings to the wider world.

The United Kingdom National Institute for Health Research describes in a 'How-to guide' the advantages of involving the community (*13*), and provides guidance on how-to involve communities in research.

Conclusion

Involving communities in several or all stages of a research study is not only an ethical requirement reflecting people's right to have a voice in issues that affect them. With adequate capacity-building of both researchers and communities at-risk, these partnerships hold great potential for improving the quality of research and the uptake and sustainability of resulting innovations.

REFERENCES

1. Harrington, M. Community involvement in HIV and tuberculosis research. *Journal of Acquired Immune Deficiency Syndrome*, 2009; 52 Suppl 1: S63-6.
2. Community involvement in tuberculosis care and prevention. Towards partnerships for health. Guiding principles and recommendations based on a WHO review WHO/HTM/TB/2008.397.
3. Report of a WHO consultation on strengthening the active engagement of civil society organizations in the global TB prevention, care and control efforts. WHO/HTM/TB/2010.15
4. UNESCO Searchable Database in Ethics Related Legislations and Guidelines. http://www.unesco.org/shs/ethics/geo/user/index.php?action=select&lng=en&db=
5. Tindana, PO et al. Grand challenges in global health: community engagement in research in developing countries. *PLoS Med*, 2007; 4(9): e273.
6. Staley K. *Exploring Impact: Public involvement in NHS, public health and social care research.* Eastleigh, INVOLVE, 2009.
7. Viswanathan M, et al. Community-Based Participatory Research: Assessing the Evidence. Evidence Report/Technology Assessment No. 99 (Prepared by RTI—University of North Carolina Evidence-based Practice Center under Contract No. 290-02-0016). AHRQ Publication 04-E022-2. Rockville, MD: Agency for Healthcare Research and Quality. July 2004.
8. Abma TA and Broerse JEW. Patient participation as dialogue: setting research agendas. *Health Expectations*, 2010; 13(2): 160-73.
9. Dara M, Grzemska M, Kimerling ME, Reyes H, Zagorskiy A. Guidelines for Control of Tuberculosis in Prisons. TBTCA, The Hague 2009. http://www.tbcta.org//Uploaded_files/Zelf/GuidelineTBPrisons1252321251.pdf
10. Sbongile P et al. Establishment of a Community Advisory Board (CAB) for tuberculosis control and research in the Inanda, Ntuzuma and KwaMashu (INK) area of KwaZulu-Natal, South Africa. *Health Policy,* 2010;95: 211-215.
11. Okelloh D, Mchembere W, Nduba V, Van 't Hoog A, Laserson K, Mitchell EMH. Perceptions of Study Procedures Involving children in Luo Communities. At Symposium 02. "Beyond numbers: the value of qualitative research for TB programme policy and practice"41st World Lung Health Conference, Berlin.
12. Singh JA, Mills EJ: Abandoned trials of Pre-Exposure Prophylaxis for HIV: What went wrong? *PLOS Medicine,* 2005, 2:e234.

13. *Examples of training and support for public involvement in research: Sharing innovative practice workshop.* Eastleigh, INVOLVE, 2010. (http://www.invo.org.uk/pdfs/TrainingSupportWEB140610.pdf, accessed 7 January 2011).

14. *Patient and public involvement: How-to guide.* National Institute for Health Research, London, 2009. (http://www.rds-nw.nihr.ac.uk/resources/PPI/How_to_guide_versionv3.pdf, accessed 7 January 2011).

V. GLOBAL FUND INVESTMENT IN TB OPERATIONAL RESEARCH

The Global Fund, as one of major funders in global health, has been providing financial support for country applicants to conduct operational research. It serves an important role in the grant assessment process. In 2002, the Global Fund Board requested its Portfolio Management and Procurement Committee to provide guidance to the Technical Review Panel (3rd Board meeting, October 2002) on operational research. Later, for Global Fund round 10 proposals, the Board requested the Global Fund Secretariat to "urgently work with partners to adopt measures to identify gaps and to further improve the quality of Global Fund-supported prevention, treatment, care and support including operational research to identify effective scaling up strategies to improve outcomes."[q]

In general, there are no existing policies or requirements for Global Fund country applicants to include operational research in their proposals. However, countries are encouraged to allocate up to 10% of the total budget for activities related to monitoring and evaluation (M&E) including system strengthening, data management, outcome and impact evaluation of programmes and interventions, operational research studies, and any other M&E related activities. Countries receiving Global Fund grants should consider potential research needs in their national planning schedules and earmark sufficient resources to conduct such studies *(1)*. The *Framework for operations and implementation research in health and disease control programs (2)*, published in 2009, was developed in collaboration with other partners to provide a set of comprehensive guidelines for programme managers, implementers, researchers, and policy-makers to plan and carry out research that would inform implementation.

Sound research studies have been proven to be a powerful tool to inform decision-makers on programme implementation issues. However, there has not been a systematic analysis of what type of research is supported by the Global Fund. This section attempts to provide background on the level of the funding and type of research activities undertaken by countries, together with a snapshot of the Global Fund operational research portfolio, through the review of the research studies included in the country grant proposals. It describes the type of research studies proposed by each disease area, and also gives a picture of the research funding allocated from a sample of grants in which financial information was clearly identified.

Methodology

Programmatic indicator data entered as of the end of 2009 from grant rounds 1 to 8 were extracted from the Global Fund Strategic Information Database. Specific indicators of research performance were used to flag particular grants that had a relevant component. Examples of such indicators included "number of operational research studies conducted or funded", "number of training conducted related to operational research". Since there was no strict working definitions of research categories, any studies that countries considered as operational research were flagged. The original proposals of these identified grants were then specifically reviewed, and the level of resources allocated to research and the methodology and the topics of the studies were determined. Grant agreements were also reviewed in the event that final agreed amount was different from the proposal. Individual principal recipients with different studies and budgets were considered as separate grants, even if they were funded in the same round for the same disease area.

Key findings

Overall, 103 grants, among 54 countries, were found to have an operational research component in their proposals; out of which, 26 countries incorporated a TB operational research component. Among the 63 (out of 103) grants with budgetary information, the majority (78%) allocated less than 5% of the total budget towards operational research, with an average of 3.8% (Global Fund Portfolio Analysis

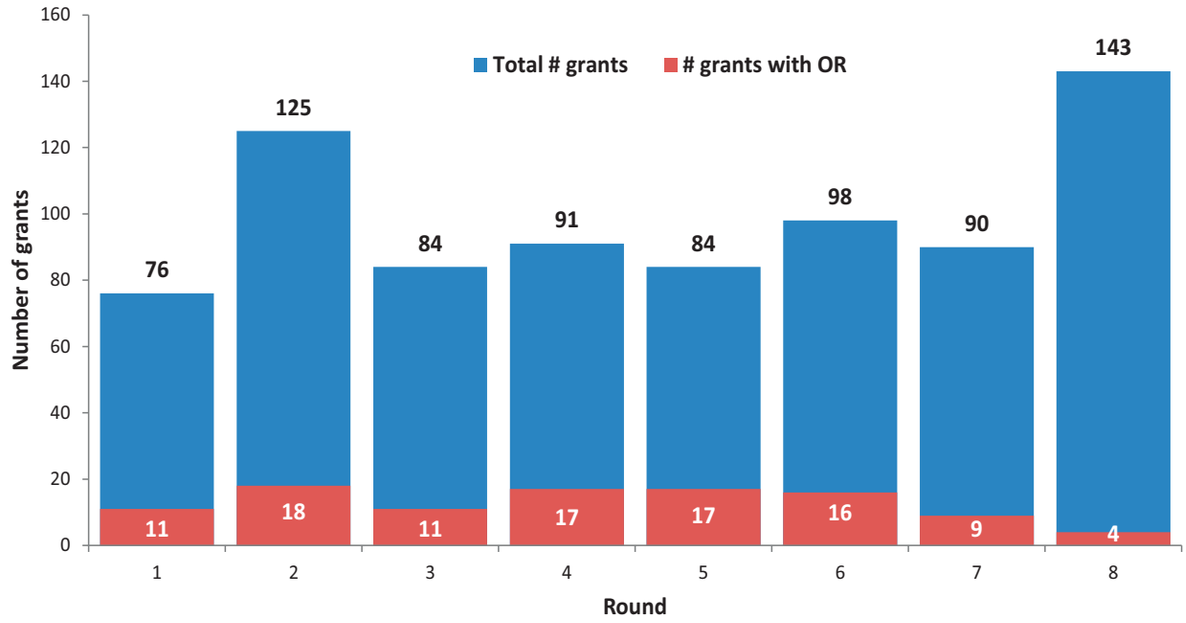
[q] 19th Global Fund Board meeting, 5–6 May 2009.

2010). The disease focus of the identified research was distributed evenly with approximately one-third of grants in each of the Global Fund disease areas, along with 2 HIV/TB, 1 health system strengthening, and 1 integrated grant.

This distribution reflects the general geographic needs of specific disease areas. Approximately half the identified proposals were from sub-Saharan countries: of the 54 relevant originating countries,[r] 27 were from sub-Saharan Africa, (with 11 in West and Central Africa, 9 in southern Africa, and 7 East Africa). Only 4 grants originated in the Eastern Europe and Central Asia region. Rwanda had the highest individual number of relevant grants (5), followed by India (4), Namibia (4) and Sri Lanka (4).

The number of grants with operational research from each Global Fund round is displayed in **Figure 7**. Of note, the low number shown in Round 8 is due to the incomplete performance frameworks of a subset of the grants at the time of the data extraction.

FIGURE 7: NUMBER OF GRANTS BY ROUND



Level of funding

How much effort and resources countries receiving Global Fund financing are devoting to implementation and operational research studies has not been previously clarified. While countries are encouraged to allocate up to 10% of grant funding to M&E and related research, there are no specific guidelines on quality of research design and rigour and appropriate budget levels to allocate to a priority research agenda once identified.

Not all of the grants reviewed provided detailed financial information on their research budget. Of the total 103 grants reviewed, 63 grants had financial information that could be extracted. These grants were distributed among 39 countries and include two multi-country proposals. The review found that the total budget allocated to operational research for these 63 grants was US$ 30 736 854, with an average of US$ 487 887 per grant. The overall amount allocated for operational research was not consistent over grant rounds, as shown in **Table 1**, with a peak occurring in round 6. On average, the reported operational research budget for AIDS, TB and malaria grants amounted to about 2.5% of the total grant budget in each round. While only about a third of the 63 grants were AIDS-related, funding for

[r] Includes three multi-country grants which are considered a separate 'country'.

AIDS-related operational research was about US$ 16.7 million – or approximately 54% of the total budget. The allocation of funding for TB-related operational research was about US$ 5.7 million. Overall, more than half (56%) of the total operational research budget was allocated to grants submitted from sub-Saharan African countries.

TABLE 1: AVERAGE REPORTED IMPLEMENTATION/OPERATIONAL RESEARCH BUDGET BY GRANT, BY ROUND

Rounds	Total number of grants, reviewed by round	Total operational research, US$, by round	Average operational research budget by grant	Allocation to operational research as a percentage of total grant budgets
1	**7**	**4 586 378**	**655 197**	**2.1%**
2	**10**	**2 738 043**	**273 804**	**3.1%**
3	**7**	**4 431 566**	**633 081**	**2.2%**
4	**5**	**1 585 000**	**317 200**	**1.0%**
5	**10**	**3 797 623**	**379 762**	**2.4%**
6	**15**	**9 297 810**	**619 854**	**3.7%**
7	**5**	**2 576 141**	**515 228**	**3.6%**
8	**4**	**1 723 293**	**430 823**	**2.8%**
Total:	**63 grants with financial information**	**Total allocated to operational research 30.7million**	**487 887 (Av.)**	**Av. Of 2.5% of Total Grant Budget**

Types of TB operational research

In general, many programme interventions did not have in-depth epidemiological data on the targeted population in countries and it is probable that they would take this opportunity to establish baseline figures in order to set their targets for the coming years. Very few countries specifically proposed cost-effectiveness studies.

For tuberculosis, portfolio analysis showed that behavioural surveys and disease surveillance were the primary themes of the proposed operational research. Of all the TB- and TB/HIV-related grants reviewed, thirteen countries included studies on MDR-TB; six reported studies on DOTS strategy and implementation; six reported studies on HIV/TB co-infection; and one focused mainly on research capacity building. While no specific study topic was proposed in the area of health system strengthening, a number of indicators related to the numbers of people trained in research and capacity building were reported in all disease areas.

A call to action

The Global Fund has been active engaged in global discussions on the operational research agenda with WHO, UNAIDS, USAID and the US National Institutes of Health (NIH), but has not been directly involved in reviewing study proposals and tracking their progress. One of the limitations of this review

is that currently the Global Fund does not require countries to report on their research studies in detail, and there is not a systematic process to track the progress of studies conducted with funding from the Global Fund other than through grant performance frameworks, which often do not contain detailed information. There is also no consistent mechanism to assess the quality of the research design, methodology, ethics and results, and how the programmes would utilize such results to inform stakeholders or programme design. Because reporting of operational research activities is currently not mandatory, the results presented here are likely to under report the level of funding, scale and type of research currently conducted with Global Fund support.

As clearly described elsewhere in this document, there are several research priorities to improve TB control. The opportunity for developing operational research to TB control starts during the Global Fund proposal development process and should involve in-country researchers and stakeholders. Additional opportunities can be explored through M&E system strengthening efforts, regional meetings, and special country case studies.

The Global Fund is committed to the development of policies on implementation and operational research, and communication to implementing countries to increase awareness of availability of related resources. In partnership with technical partners including the WHO Stop TB department, the Stop TB Partnership and the Special Programme for Research and Training in Tropical Diseases (TDR), this report is an investment towards bridging the information gap on priorities for TB operational research for reference during grant application and implementation to improve performance and impact of Global Fund investments.

REFERENCES

1. Monitoring and evaluation toolkit (3rd edition). Geneva, The Global Fund to Fight AIDS, Tuberculosis and Malaria, 2009.
2. Framework for operations and implementation research in health and disease control programs. Geneva, The Global Fund to Fight AIDS, Tuberculosis and Malaria (with USAID, WHO/TDR, UNAIDS and the World Bank), 2009.

ANNEXES

Annex I

Operational research methods, statistics and definitions

Operational research

There are various definitions of *operational research*. These have been summarized and discussed in detail elsewhere *(1, 2)*. From a TB control perspective, the objective of operational research is generally aimed at helping TB control managers to improve programme operations, and providing policy-makers with evidence-based answers to address service delivery problems. In pragmatic terms, it is generally agreed that operational research can accomplish the following:

- improve programme performance and outcomes by assisting programme managers and staff to understand how the programmes work, identify problems and solve them in a timely manner;
- help programme managers and policy-makers to take evidence-based programmatic /public health decisions;
- assess the feasibility, effectiveness or impact of new strategies or interventions on TB control;
- collect data to guide policy recommendations on specific interventions.

Research approaches

It is essential that the correct methodological design is selected to properly address the stated study objectives. We first describe *quantitative observational designs*, which are often referred to as the 'traditional epidemiological approach'. These classically have a designated study arm which acts as a *control*; this arm usually occurs naturalistically, consisting typically of individuals who either did not experience an outcome of interest (e.g. were diagnosed as TB negative) or who were not exposed to a risk/aetiological factor of interest (e.g. did not have a TB contact in their home or did not smoke).

We then focus on *quantitative interventional designs*. The most robust of these designs is the *randomized controlled trial* (RCT) that is aimed at assessing the efficacy of a test intervention. RCTs require (at least) two designated groups: an *intervention* group in which individuals receive the test intervention, and a *control* or *comparator* group in which individuals receive either a standard intervention (e.g. standard clinical care or a placebo treatment) or, in certain circumstances, no treatment (e.g. if the study objective is to evaluate a community health education programme). Each study participant is placed in one of the study arms, usually through a random process. Traditionally RCTs have not been seen as part of operational research, but in recent years there has been increased recognition of the value of *'pragmatic' randomized controlled trials* (PRCTs) within operational research. These are implementation studies, taking place as much as possible in the context of day-to-day health service provision and practice, and aim at evaluating the effectiveness of an intervention in the real world rather than in the ideal setting created in most conventional RCTs.

In the third section we deal with *qualitative research designs* and approaches. These are invaluable at many stages in the research cycle, and particularly so in the situational analysis stage, where they often help shaping questions or hypotheses to be tested in later stages.

Finally we define some newer research approaches, including those with a particular focus on assessing equity and access, which are important given the links between TB care and control and the Millennium Development Goals.

With all of these approaches, it is important to emphasise that coordination between countries or regions can be advantageous to assist in regionally appropriate decisions. This is not to say that some operational research is, necessarily, context specific and solves local operational issues.

1. Quantitative observational (traditional epidemiological) designs

Cross-sectional study/survey

- A study to measure the distribution of a condition (such as the burden of a disease) at a given point in time *(3)*, or to identify if there is a correlation between a possible determinant(s) and a specific disease at the population level and at a given point in time.

- Data are collected across a selected population within a limited time-frame. For example, the primary objective of a survey might be to estimate the proportion of MDR-TB cases in a population at a given point in time, and to determine if this proportion varies across the population; a secondary objective might be to determine if there is a relationship (correlation) between being an MDR case and having potentially predisposing characteristics such as age, gender or occupation at that same point in time.

- A longitudinal element can be added by repeating the same survey in the same population over a series of time-points to investigate whether the proportion of MDR TB cases is changing in the population, and/or to investigate whether the predisposing factors are changing with time.

Cohort study

- A study of (usually two) groups (*cohorts*) of individuals who are followed-up over a period of time to measure the occurrence of a defined *outcome* (e.g. the diagnosis of TB, or being identified as having MDR-TB).

- Conventionally, one cohort will consist of individuals naturally exposed to a pre-defined aetiological or *risk factor* (i.e. potential determinant of the outcome) while the other cohort will consist of individuals not naturally exposed to the risk factor. If all individuals in a population are exposed to the risk factor to different extents, one cohort may be individuals with a high exposure level and the other cohort individuals with a low exposure level. Typical risk factors for TB include gender, HIV status or smoking.

- The rates of occurrence of the outcome are compared between the cohorts *(3)*. If there is a relationship between the exposure to the risk factor and the outcome, the proportion of individuals in the cohort exposed to the risk factor who experience the outcome will be significantly greater than the proportion of individuals in the cohort not exposed to the risk factor who experience the outcome.

- The cohort study design is the optimal observational design for obtaining evidence of a *causal temporal relationship* between outcome and risk exposure.

- Cohort studies are always longitudinal as they track the individuals in the cohorts over a period of time. They are usually prospective, with the cohorts followed forwards through time, but can also be conducted retrospectively.

- The cohort study design allows multiple outcomes to be related to exposure to a specific risk factor, allows accurate information to be collected on the level of exposure to that risk factor, and enables the study of relatively rare risk factors.

- The results of a cohort study can be reported as *risk ratios* or *risk differences*, which are intuitive and easy to understand. The STROBE statement gives guidelines for the conduct and reporting of observational studies *(5)*.

- The cohort study design is not appropriate if the time between exposure to the risk factor and the occurrence of the outcome may be very long, if the outcome of interest is rare, if the follow-up of cohort members will be difficult, and/or if exposure patterns

may change over time (possibly making the results of the cohort study irrelevant). Cohort studies tend to be expensive to conduct.

Case-control study

- A study of two groups of individuals whose personal histories are examined to determine their levels of exposure to a pre-defined aetiological or risk factor (e.g. HIV status, diabetes).

- One group is defined as the *cases* – these are individuals who have experienced a defined outcome (e.g. an adverse TB treatment outcome such as failure, default or death). The second group is defined as the *controls* – these are individuals who have not experienced the same defined outcome (e.g. have had a favourable TB treatment outcome such as cure or treatment completion).

- The rates (frequencies) of exposure to the pre-defined risk factor are compared between the cases and the controls. If there is a relationship between exposure to the risk factor and the outcome, the proportion of cases exposed to the risk factor will be significantly greater than the proportion of controls similarly exposed.

- Case-control studies are always longitudinal as they track both the cases and controls over a period of time. They are usually retrospective, with the cases and controls followed backwards through time – but can be conducted prospectively.

- The case-control study design allows exposure to multiple risk factors to be related to a defined outcome, enables the (rapid) study of outcomes that are relatively rare or have a long latency between exposure and manifestation, and are often relatively inexpensive to conduct.

- The case-control study design is not appropriate if the information on exposure to the risk factor is poorly remembered (*recall bias*) and/or incomplete (a problem that generally increases as the recall period increases).

- The case-control study design cannot provide correct estimates of either disease prevalence or the probabilities (risks) of the outcome occurring in individuals exposed to and not exposed to the risk factor. Results can only be reported as *odds ratios (3)*, which are not intuitive and can be difficult to understand, and are a proxy measure of the relative risk of being a case when exposed to the given risk factor.

- The choice of an appropriate control group for the cases can be difficult, and the method of selecting individuals into the control group can be contentious. One option is to randomly sample eligible control individuals so that the distribution of characteristics/risk factors in the control group reflects the distribution of these characteristics/risk factors in the general population from which the cases arise.

- Commonly, controls and cases are *matched* according to important (*confounding*) factors such as age, gender and environment (place of residence).

2. Quantitative interventional

Pragmatic randomized controlled trial (PRCT):

- A randomized controlled trial approach with the purpose to inform decisions about routine day-to-day practice *(5).*

- A PRCT differs from an *explanatory* RCT in that it focuses on *effectiveness* (does the intervention work when used in the real world, i.e. under routine normal practice?) rather than on *efficacy* (does the intervention work in ideal and fully controlled conditions?).

- PRCTs are more suited to operational research than explanatory trials. Explanatory trials are necessary and extremely valuable in empirical research to demonstrate the potential usefulness of a new intervention by testing its efficacy and/or safety under strictly controlled conditions, which might then be taken forward into a PRCT.

- Although characterized as being different for the sake of this explanation, it must be recognized that there is a spectrum of trial type with pure pragmatic trials at one end of the spectrum and pure empirical trials at the other *(6)*.

- Quality criteria are well developed and documented for the conduct and reporting of randomized controlled trials in the CONSORT statement *(5)*. An addendum to the CONSORT statement to improve the quality of PRCTs has recently been published *(7)*. This published addendum includes a table describing the key differences between explanatory and pragmatic trials – which is summarised in **Table 2**.

TABLE 2: KEY DIFFERENCES BETWEEN EXPLANATORY AND PRAGMATIC TRIALS

	Explanatory	Pragmatic
Question	Efficacy: does the intervention work?	Effectiveness: does the intervention work when used in normal practice?
Setting	Well resourced, rigorously controlled conditions	Normal clinical or public health practice
Participants	Selected. Participants who are poorly adherent are either considered as having a negative outcome or are not assessed	Little or no selection
Intervention	Strictly enforced and adherence is closely monitored	Applied flexibly within the requirements of normal practice
Outcomes	Often short term surrogates or process measures	Directly relevant to participants, funders, communities and healthcare practitioners
Relevance to practice	Indirect: little effort made to match trial to decision-making needs of those in the usual setting in which the intervention will be applied	Direct: designed to meet the needs of those making decisions about intervention options in the setting in which the intervention will be implemented

Options for units of randomization in PRCTs

Randomization is arguably the most important element of a PRCT. Individuals are allocated to the study groups through a random process in order to prevent bias in the selection of group members. Conventionally, individuals are randomised into the study groups. However, if this is not possible or even desirable (e.g. if the intervention is to be delivered to entire communities), then groups (*clusters*) of individuals are randomized. Individual randomization is more suited to explanatory trials (see above).

- *Individual randomization:* Individual study participants are randomized to receive either a new intervention (e.g. a new diagnostic algorithm or a new treatment/monitoring modality) or an existing/alternative intervention.

- *Cluster randomization:* Clusters (or groups) of study participants are randomised. A cluster can be defined in various ways to suit the design of the particular study in question. Examples, in order of decreasing cluster size, include:

  Administrative Districts (one cluster = all patients attending health services within a given health district)

  Enumeration Area (one cluster = all individuals in a census enumeration area)

  Health Units (one cluster = all patients attending a given health unit, be it a health centre or a hospital – as defined by the study needs)

  Households (one cluster = all individuals in a household)

  Time-defined (one cluster = all patients attending during a defined time period, e.g. one week, two weeks, one month, or several months) this is traditionally

used when an intervention is applied to all health unit attendees for a given time period, and then swapped back to the comparator intervention (so called "week-on, week-off" or "month-on, month-off").

Cluster randomization is more suited to pragmatic trials than individual randomization. Cluster randomized designs permit the capture of health-system effects (both direct and indirect) and often make use of routinely collected data.

Options for control or comparator arm allocation in intervention trials/studies

- ***Before-and-after design***: This design does not include a separate control group to the intervention group. For all participants in this design, data on the outcome measures is collected over a baseline (pre-intervention) time period, the new intervention is then introduced, and finally data on the outcome measures is collected over a follow-up (post-intervention) time period:

In the absence of any other option, this design can provide important information about the possible effectiveness of an intervention. However, the absence of a contemporary control or comparator group means that any difference found between the pre-intervention and post-intervention study periods could be due to factors other than the intervention. Participants act as their own controls, but treatment (control and intervention) is confounded with time – so the interpretation of the results from this design is often problematic. Ideally, the baseline data should be collected prospectively, so that the data collection methods (and hence data quality) are the same in the pre- and post-intervention study periods. Alternatively, however, *historical control* data (data which has already been collected for other purposes) could be used as the baseline.

This is the weakest of the interventional designs. The following designs all include contemporary controls or comparator arms and are amenable to randomization, so the intervention effects detected are considered more robust.

Cluster	Pre-intervention study period →	Post-intervention study period →
1	control	intervention
2	control	intervention

- ***Parallel groups design***: Individuals or clusters are randomly allocated to the intervention and control groups concurrently, and data collection on the outcome measures occurs in both groups concurrently:

This gives a direct comparison of the intervention and control groups. If cluster randomization is to be used, the clusters can be matched in pairs for important characteristics (such as sex distribution, population size, distance from a health facility) and randomized within these pairs.

Cluster	Study period →
1	intervention
2	intervention
3	control
4	control

- ***Cross-over design***: Individuals or clusters are initially randomly allocated to the intervention and control groups concurrently, and data collection on the outcome measures occurs in both groups concurrently over an initial study period. At the end of this period, all participants *cross-over* to the other treatment arm, and data collection on the outcome measures continues in both groups concurrently over a second study period:

This design usually requires fewer participants (and fewer clusters) than a parallel groups design, but a longer time duration is needed to complete the study. There may also be important considerations for a *wash-out time* at the cross-over point so that any effects that occur in the first study period do not carry-over into (contaminate) the second study period.

Cluster	Study period →	Study period →
1	intervention	control
2	intervention	control
3	control	intervention
4	control	intervention

- ***Stepped-wedge design*** *(8):* This is an important and ingenious modification of the cross-over design, and is particularly useful in situations where the only other alternative is the before-and-after study design. This PRCT design makes use of an implementation plan where an intervention is introduced to some areas or health facilities before others over time. Ideally the sequence by which these areas or facilities implement the intervention is randomized. Comparisons are then made over time between those areas of facilities receiving the intervention and those not yet receiving it:

As this is essentially a cross-over design, the stepped-wedge design also generally requires fewer clusters than a parallel groups design, but requires more time to complete. It has the advantage of phasing-in an intervention over time, as would happen with uncontrolled or non-randomized implementation, so no study areas are deprived of a new intervention if this proves to have an added value.

Cluster	Study period				
	1	2	3	4	5
1	control	intervention	intervention	intervention	intervention
2	control	control	intervention	intervention	intervention
3	control	control	control	intervention	intervention
4	control	control	control	control	intervention

3. Qualitative approaches

Qualitative research: aims to gather an in-depth understanding of people's attitudes, behaviours, values, concerns, motivations, aspirations, culture or lifestyles and the reasons that govern such behaviour. Various methods are used, including in-depth individual interviews, participant observation and focus group discussions to saturation point (the point at which no new responses are elicited). Qualitative research is usually conducted on small, focused samples. In-depth individual interviews are usually conducted with fewer than 100 participants – usually patients or service users. A single focus-group discussion is usually conducted with a group

of around 10 people who have been chosen to give a range of views on a particular topic. If several focus-group discussions are held, then there may be more than 100 individuals involved. *Key-informant interviews* tend to focus on service providers or policy makers and use in-depth interview techniques. These may involve fewer than 10 individuals.

Policy analysis: analysis and synthesis of evidence and social values in making clinical and policy decisions. Qualitative research methods are usually employed and can be retrospective or prospective, include semi-structured in-depth interviews of key informants, focus groups, document analysis and process evaluation of implementation. This is sometimes called Policy Transfer (or Policy Transfer Analysis).

Realist review: a model of research synthesis that is designed to work with complex social interventions or programmes, and that is based on the emerging 'realist' approach to evaluation. It provides an explanatory analysis aimed at discerning what works for whom, in what circumstances, in what respects and how *(9)*.

Case studies: these are detailed descriptions of programmatic approaches that achieve a balance between detailed understanding of the context and generic lessons that can be learned *(10)*.

4. Other research approaches

Equity Analysis: This method is still in its infancy in the context of operational research. It is an approach that attempts to answer whether an intervention is fair: principally, does an intervention promote access for all people, especially by the poor and vulnerable? It involves collecting information on patients' socioeconomic profiles and the costs they incur in accessing services. It is a quantitative approach and can be applied in both observational and interventional approaches (see above).

Health systems analysis: This is an approach that captures the health system requirements of a given intervention. It asks what is required in the way of infrastructure, utilities, human resources, quality assurance mechanisms, supply chain, procurement and disposal. Usually these are summarized in economic terms. The economic measures generated are a quantitative outcome (*11*). Health systems analysis can be descriptive in observational studies, but can also be applied in interventional studies, especially cluster-randomized trials, where the analysis can be comparative – for instance comparing health system requirements of an intervention versus a control or comparator.

Mapping: This refers to the use of geographical information systems (GIS) in which digitized, electronic maps are used to illustrate multiple layers of information of interest. For example, health facilities can be plotted onto these maps using geo-satellite referenced personal digitized assistant electronic devices. Such maps can be overlaid with programme-relevant information such as numbers of TB cases notified, sputum specimens examined etc.

Operational modelling: This encompasses a wide range of problem-solving techniques and methods to model optimal or near-optimal solutions to complex decision-making. The potential for scale-up, relative benefits with regard to the new interventions and cost savings are assessed. Different options can be assessed using decision analytical modelling. More detailed assessments, e.g. exploration of strategies managing bottlenecks or barriers in systems can use supply chain and modelling software (such as Witness).[s]

Transmission modelling: Transmission dynamics (i.e. how infections spread) are modelled to identify better strategies for disease control. A number of methodological approaches can be used to investigate these dynamics including observational study and epidemiological analysis, mathematical modelling, and systems analysis.

Systematic review: A systematic review attempts to identify, appraise and synthesize all the empirical evidence that meets pre-specified eligibility criteria to answer a given research question. Researchers conducting systematic reviews use explicit methods aimed at minimizing bias, in order to produce more reliable findings that can be used to inform decision making.[t] *(12)*

Meta-analysis: Meta-analysis is a statistical technique for combining the findings from several independent studies. Meta-analysis can offer an objective unbiased synthesis of the empirical data and assess the consistency of the results of studies included *(12, 13)*.

[s] See: http://www.lanner.com

[t] See: http://www.thecochranelibrary.com/view/0/AboutCochraneSystematicReviews.html

Sample size/statistical power

For both ethical and practical reasons, the optimum number of participants needed to properly address the stated study objectives should always be estimated before the study starts. If too few individuals are studied, important intervention effects or risk factors may go undetected, denying patient's access to effective disease treatments or preventing risk factors being adequately identified, which may call for targeted intervention to reduce disease incidence. If too many individuals are studied, some may be exposed to risk factors or be treated with an inferior intervention for longer than necessary (or even totally unnecessarily).

Predominantly qualitative research approaches require least attention to sample size, as it is often more important to purposively select a sample of individuals who will represent the widest possible range of opinions/experiences across the target study population. However, data saturation principles are often used to ensure that the optimum number of participants are studied (i.e. data collection is stopped when no new information is being obtained from study participants and/or no new themes are emerging). This usually occurs after 10-20 individuals (depending on heterogeneity) have been interviewed. Similarly, focus groups are conducted until saturation is achieved: these typically comprise 6-10 individuals, often grouped by one or two key demographic or health-related factors (e.g. age, sex, patients/suspects, using ART or not).

For quantitative cross-sectional studies/surveys, sample size is based on the precision with which it is desired to estimate the prevalence or incidence of the condition of interest, or the strength of the relationship between two measures. Precision in this context is most often defined as the width of the (95%) *confidence interval* around this estimate. The formula for this confidence interval (and hence for the estimation of sample size) becomes more complex as the study design becomes more complex *(14)*, but the widely used SUDAAN statistical software package[u] calculates sample size for many survey design options.

For comparative quantitative research projects (including cohort studies, case-control studies and PRCTs) sample size is based primarily on the nature of the primary outcome measure and the minimum difference between the groups that it is desired to detect (i.e. minimally important effect size). There are mathematical formulae available for most study design options and most types of outcome measure *(15–18)*. There are also many software packages available, both commercially and (of more variable quality) as shareware on the internet. Many of the formulae are mathematically complex, and the software packages can be difficult to navigate, so it is usually advisable to seek expert statistical assistance. Before doing so, however, the website *Sample size for clinical trials*[v] is worth reviewing, as it provides a very readable introduction to sample size calculation covering most simple situations – and, perhaps more importantly, includes an easy to use graphical method for calculating an initial (approximate) sample size estimate.

For all types of quantitative study design, the complexity of the sample size calculations increases considerably if sampling or randomization is by cluster rather than by individual, as the calculations have to take into account the likely size of the intra-cluster correlation (ICC); unfortunately, few published studies report their ICC so precedent is rare – typically ICC ranges between 0.001 and 0.100, but can be larger. Cluster designs require calculation of the optimum number of clusters *and* the optimum number of individuals per cluster. Many different combinations of these two numbers will give the same level of statistical power, so the final decision is usually a compromise based on practical considerations. In general, a large number of clusters with few individuals per cluster is better than a small number of clusters with a large number of individuals in each cluster *(17)*. Specialist software is available for these designs, but should only be used with expert statistical guidance.

[u] See: http://www.rti.org/sudaan

[v] See: http://www-users.york.ac.uk/~mb55/msc/trials/sampsz.htm

Statistical analysis methods

Most of the information contained in the findings of quantitative research, whether this is a survey, observational comparative study or an interventional study, can be extracted by the proper application of simple summary statistical methods and graphs. For more complex statistical analyses, researchers are strongly advised to seek expert statistical assistance.

In general, greater emphasis is given now to estimating effect sizes with (95%) confidence intervals than to formal significance testing, although both have an important role in the evaluation of clinical research data.

For *simple cross-sectional studies/surveys*, a first analysis should attempt to provide estimates of the main outcome measure (mean values for continuous measures, proportions for categorical measures, correlation coefficients for measures of association) with their 95% confidence intervals, both for the whole study sample and then for important sub-groups. For more complex designs, weightings may need to be applied to sub-groups of participants if, for example, different sampling proportions were used in these sub-groups – this is one point at which expert statistical advice may be required. Time trends in measures such as case-detection rates can be evaluated using contingency table analyses, but Poisson regression modelling methods are usually more sensitive (see below).

For *observational cohort studies*, the proportions of people in each risk exposure group who experience the outcome of interest should be reported. Effect size is best represented by either the ratio of these proportions (*relative risk/risk ratio*) or the difference between the two proportions (*risk difference*), with their 95% confidence intervals. Adjustment for the influence of important confounding factors can be made using stratified analyses, Poisson regression models and/or log-binomial models. Odds ratios are still commonly reported for cohort studies, with adjustment for confounding factors using logistic regression; odds ratios are reasonable estimates of relative risk when disease incidence is rare, but otherwise they overstate the relative risk (often very considerably) – and adjustment using the Zhang and Yu formula can introduce bias *(19)*.

For *observational case-control studies,* the proportions of cases and controls exposed to the risk factor of interest should be reported. Effect size can *only* be represented by odds ratios (with their 95% confidence intervals). As for cohort studies, the odds ratio estimates can be adjusted for the influence of important confounding factors using standard unconditional logistic regression methods if the cases and controls are unmatched. However, if the cases and controls are matched, it may be more appropriate (but not essential) to use conditional logistic regression methods – these can be more difficult to compute and may require expert statistical assistance.

For *standard interventional PRCTs* in which individual participants have been randomized, the method of analysis is straightforward. If the outcome measure is continuous and can be considered to follow a normal (Gaussian) distribution, the mean difference between the intervention groups (with their confidence intervals) is the best representation of effect size; important confounding factors can be adjusted for, using standard regression methods. If the outcome measure is categorical, the PRCT can be regarded as a cohort study in which exposure to the risk factor is determined by a random process, and the same statistical methods apply (risk ratios, risk differences, odds ratios), with standard or logistic regression analysis methods then used to adjust the effect size estimate for important confounding measures.

Standard PRCTs assume that the responses of different participants to an intervention are independent of each other. Cluster-randomized PRCTs, however, assume that the responses of participants within the same cluster will be correlated (ICC) – and so require the effect size to be adjusted for this ICC. This increases the complexity of the statistical analysis, although the summary statistics and analysis methods for both types of PRCT appear similar when presented. Expert statistical advice should be sought when analysing cluster randomized PRCTs.

Most of the statistical methods described above can be carried out by careful application of any good standard statistical software package, although it may be wise to seek expert statistical advice to have a diagnostic check of the assumptions made when applying multiple regression methods.

More complex statistical methods are now widely available in many statistical packages, but these should be used only with the support of expert

statistical advice. A detailed description of these is outside the scope of this document, but the most commonly applied include:

- actuarial methods and Cox regression if the outcome measure is time to an event (e.g. time between start of symptoms and TB diagnosis);
- Poisson or negative binomial regression if the outcome measure is a count (e.g. the number of new TB cases in an area in a given month);

Finally, *Qualitative research* requires content analysis and the drawing out of key themes. This should always be done by researchers who have experience in qualitative, social science research and who must be involved in the research from the beginning. For all quantitative study designs, if data is collected at different levels (e.g. if some data is collected from individual participants while other data is aggregated across communities and/or regions), multi-level modelling methods are required. These methods require more complex analytical methods than are readily available in most statistical packages, so again expert statistical advice should be sought.

REFERENCES

1. Lienhardt C and Cobelens F. Operational research for improved TB control: the scope, the needs and the way forward. *International Journal of Tuberculosis and Lung Disease*, 2011; 15(1):6–13.
2. Zachariah R et al. Operational research in low-income countries: what, why, and how? *Lancet Infectious Diseases*, 2009; 9: 711–17
3. Enarson DA, Kennedy SM and Miller DL. Choosing a research study design and selecting a population to study. *International Journal of Tuberculosis and Lung Disease*, 2004; 8(9): 1151–6.
4. von Elm E et al (for the STROBE Initiative). The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. *Journal of Clinical Epidemiology*, 2008; 61(4):344–9.
5. Zwarenstein M et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ,* 2008; 337:a2390.
6. Thorpe KE, Zwarenstein M, Oxman AD, Treweek S, Furberg CD, Altman DG, et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *J Clin Epidemiol*. 2009;62(5):464-75.
7. *CONSORT statement*. London, CONSORT group, 2007. (http://www.consort-statement.org, accessed 7 January 2011).
8. Hussey MA and Hughes JP. Design and analysis of stepped wedge cluster randomized trials. *Contemporary Clinical Trials*, 2007; 28(2):182–91.
9. Pawson R, Greenhalgh T, Harvey G and Walshe K. Realist review: a new method of systematic review designed for complex policy interventions. *Journal of Health Service Research Policy*, 2005; 10 Suppl 1:21–34.
10. Theobald S et al. Towards building equitable health systems in Sub-Saharan Africa: lessons from case studies on operational research. *Health Research Policy and Systems*, 2009; 7:26.
11. Mann GH et al. The role of health economics research in implementation research for health systems strengthening. *International Journal of Tuberculosis and Lung Disease*, 2011; 15(6):715–21.
12. Menzies D. Systematic reviews and meta-analyses. *International Journal of Tuberculosis and Lung Disease*, 2011; 15(5):582–93.
13. Egger M and Smith GD. Meta-Analysis. Potentials and promise. *BMJ*, 1997; 315(7119): 1371–4.
14. Bartlett JE II and Higgins C. Organizational research: Determining appropriate sample size for survey research. *Information Technology, Learning, and Performance Journal*, 2001; 19(1):43–50.

15. Donner A. Approaches to sample size estimation in the design of clinical trials: a review. *Statistics in Medicine*, 1984; 3(3):199–214.

16. Machin D CM, Fayers PM and Pinol A. *Sample Size Tables for Clinical Studies (3rd edition)*. London, BMJ Books, 1997.

17. Hayes RJ and Bennett S. Simple sample size calculation for cluster-randomized trials. *International Journal of Epidemiology*, 1999; 28(2):319–26.

18. Moulton LH et al. Statistical design of THRio: a phased implementation clinic-randomized study of a tuberculosis preventive therapy intervention. *Clinical Trials*, 2007; 4(2):190–9.

19. McNutt L-A, Wu C, Xue X, Hafner JP. Estimating the relative risk in cohort studies and clinical trials of common outcomes. *American Journal of Epidemiology*, 2003; 157(10):940–3.

Annex II

Research methods to address priority areas

Note: In this annex, indications for expected duration and budget of the proposed research studies are being provided as follows:

Expected duration/timeline of study
Short-term: 6 to 18 months
Medium-term: 18 months to 4 years
Long-term: > 4 years

Budget
Low: <US$ 10 000
Medium: US$ 10 000 to 250 000
High: US$ > 250 000 to 10 000 000 (NB. large multi-country pragmatic randomized controlled trials would be an example of the type of study requiring the upper limit of this budget range.)

1. Improving access, screening and diagnosis of TB

1.1 Situation analysis

Objective: To identify local barriers in accessing diagnosis of all forms of TB in various populations (including difficult-to-reach populations) and specific risk groups (e.g. TB suspects, DS-TB, re-treatment, DR-TB, HIV, children).

Illustrative designs:
a) Audit of existing programme registers (suspect or chronic cough registers, laboratory and treatment registers).

b) Cross-sectional descriptive qualitative studies with patients (including assessment of health-seeking behaviour, delay to diagnosis, etc.).

c) Mapping of services provision and availability (e.g. health facilities) and diagnostic modalities in relation to population characteristics (e.g. density, socioeconomic factors).

d) Patient costing studies.

Setting/study population: Patients with all forms of TB recruited from health facilities in different settings (rural, urban, prison, refugee camp, etc).

Methods:
a) Retrospective review of facility registers to track cases from recognition as a suspect (in the chronic cough or TB suspect register) to submission of diagnostic specimens (in the laboratory register) to starting treatment (in the treatment register).

b) Qualitative research methods to identify financial, geographical, socio-cultural and health-system barriers. Approaches with patients, including symptomatic patients identified through TB prevalence surveys, will include critical incident narrative interview techniques to map patient pathways in seeking diagnosis, structured interviews, focus group discussions, gender analysis and case studies. Costs incurred and time spent by the patient for each care-seeking visit, provide useful additional data. Key informant interviews and focus group discussions with people in the community and health providers are important for context, triangulation and validation of findings from patients.

c) Mapping of facilities and services onto geographical information systems (GIS).

d) Costing surveys using questionnaires informed by an understanding of barriers identified through cross-sectional studies described in (a) above. An example of a patient costing tool is given in **Annex IV** (NB: this is a tool for capturing costs of diagnosis, relevant for this section, and also for documenting patient costs of accessing treatment).

Expected outcomes:
a) Estimates of proportion of patients who drop out of the diagnostic pathway at different stages.

b) Ranked lists of barriers faced by different patient groups.

c) GIS maps illustrating the relationship between patients, communities and services.

d) Key barriers expressed in economic terms.

Analysis:
a) Description of numbers of patients recorded in each register for a given time-period (eg. six months) along with cross-checking for duplicate entries. Documentation of the numbers of patients failing to proceed from one stage of the diagnostic algorithm to the next.

b) Evaluation of TB patient care-seeking pathways, drawing out major themes and describing barriers.

c) Mapping poverty indicators on geographical information systems in relation to additional data, where available, including TB notification, population density, health facilities.

d) Analysis of patient costs incurred per TB diagnosis made. Further analysis can include breakdown of

major cost categories incurred by patients during diagnosis (e.g. transport, food, user fees). Where possible, this can be further broken down into costs per health-seeking visit, and comparisons between costs incurred and assets available can be made.

Guidance for sample size calculation:
a) No formal sample size calculations are required for record reviews or audits, but a rough guide is to analyse the ultimate fate of approximately 500 TB suspects. The number of months of register data required to reach this number will depend on the size and throughput of the facility or facilities being studied.

b) No formal sample size calculations are needed for purely qualitative research. Individuals are conventionally recruited using purposive sampling (i.e. individuals are recruited to provide a cohort with the widest possible range of opinions/experiences across the target study population) until saturation is achieved (i.e. until no new themes are emerging); this usually occurs after 10-20 individuals have been interviewed (depending on heterogeneity). Similarly focus groups are conducted until saturation is achieved: these typically comprise 6-10 individuals, often grouped by one or two key demographic or health-related factors (e.g. age, sex, patients/suspects, using ART or not etc.); saturation is usually achieved after 3-4 focus group discussions with each stratum.

c) No formal sample size calculations are required for GIS mapping. It is usual to map a district or a given catchment area – usually covering a population of about 500,000.

d) For quantitative cost estimates it is usual to interview at least 100 to 200 patients.

Expected duration/timeline: Short-term.

Suitable scale:
a) Local.
b) to d): Local, regional or national.

Estimated budget range:
a) Low.
b) to d) inclusive: Medium.

Note: As a result of the situation analysis, practical steps must be taken to address identified obstacles in accessing TB diagnosis. Suggested methods for identification of new approaches are described below.

Illustrative references: *(1–6)*.

1.2 Identifying new programmatic approaches

Objective(s):
a) To understand factors facilitating or hindering effectiveness of existing diagnostic algorithms.
b) To document accuracy of new diagnostic tests or packages of tests.
c) To project a number of suitable options of diagnostic approaches or packages for potential implementation.
d) To select, at the international or country level, new or revised diagnostic approaches to be further piloted (see **section 1.3**).

Illustrative Designs:
a) Realist review of local experience (national or regional) by researchers.
b) Systematic review of new tool(s) or approaches (carried out at international level).
c) Operational and transmission modelling.
d) Expert group review meetings at international, national or regional level (informed by results of (a), (b) and (c) above) to examine the results of individual studies, systematic reviews and realist reviews in relation to existing tool(s) and approaches that have been nationally or internationally endorsed. Where possible, results of locally-conducted studies (national or regional) should be included in order to identify the new approaches that are most suitable to the local context.

Setting/study population: National programme and local academic organizations – with international partnership as appropriate.

Methods:
a) Realist reviews: Synthesis of programmatic implementation experiences alongside any available informative research findings from work carried out in the situational analysis (see **section 1.1**).

b) Systematic reviews: These usually synthesise global evidence, but a country may identify an implementation question about a new or local programmatic approach that has not already been the subject of a systematic review. In this case a national body may wish to partner with an international organization with systematic reviewing expertise.

c) Operational and transmission modelling: Operational modelling can be used for virtual testing of new programmatic approaches, predicting likely health system requirements for

given workloads of different options or packages. If linked with transmission modelling, impacts on TB epidemiology may also be predicted. This can help in making rational choices of programmatic approaches to pilot, phase-in or scale-up.

d) Expert group review[w] meetings convened at national or international level to examine the results of individual studies, systematic reviews, realist reviews, and modelling outputs (if available) in relation to existing nationally and internationally endorsed tool(s) and approaches. Where possible, results of locally conducted studies (national or regional) should be included in order to identify the new approaches or algorithms that are most appropriate to the local context.

Expected outcome(s):
a) Compilation of lessons learnt about factors facilitating or hindering effectiveness of existing diagnostic algorithms.

b) Synthesised data on accuracy of potential new diagnostic tests or combinations/packages of tests.

c) A number of suitable options of diagnostic approaches or packages giving estimates of resource requirements (e.g. infrastructure, human resources, procurement) and likely epidemiological impact (on TB transmission).

d) Local evidence-based formulation of a new programmatic approach (or diagnostic package) based on internationally and nationally endorsed tools or approaches. Ideally this approach would then feed into research activities in **section 1.3** (see below).

Analysis:
a) Realist review: Reflective, flexible and participatory analysis of both qualitative and quantitative evidence to explain how complex interventions work (or why they fail) in particular contexts and settings.

b) Systematic review: Comprehensive review of published and unpublished materials; meta-analysis and synthesis.

c) Operational Modelling: Creating national-level models of health systems and testing different potential placement of tests (or packages) or algorithms using software (e.g. Witness) and modelling expertise.

d) National or regional expert group meeting: typically convened at international level (e.g. by WHO) using the Grading of Recommendations Assessment, Development and Evaluation (GRADE)[x] approach to generate evidence-based guidance. For some larger countries or a number of countries in a region, it may be appropriate that national or regional expert group meetings are convened.

Guidance for sample size calculation:
Formal sample size calculations are not required for the above approaches to synthesising existing evidence.

Expected duration/timeline: a) to d) inclusive: Short-term.

Suitable scale: a) to d): Local, national and international.

Estimated budget range: a) to d): Medium.

Note: On the basis of the above process, new diagnostic algorithms could be developed by international organizations (e.g. WHO) that include new diagnostics tools or approaches. These could include, for example:

i) Development of improved clinical algorithms for smear-negative TB in high, medium and low HIV prevalence settings with and without access to (digital or conventional) chest X-ray.

ii) Identification of specific risk factor profiles for MDR-TB in different settings, that could be used for the presumptive identification of suspect MDR-TB cases or identification of MDR-TB risk groups.

Another recent example is the development of new diagnostic algorithms that is following the endorsement by WHO of the new fully automated NAAT, Xpert MTB/RIF; these algorithms will have to be tested and validated in various settings (see **section 1.3** below), as part of a roadmap to roll-out the use of this new diagnostic tool.

Illustrative references: *(7–10)*

[w] At international level, the Strategic and Technical Advisory Group for TB (STAG-TB) of the WHO provides technical endorsement and guidance on new tools and approaches.

[x] See: http://www.gradeworkinggroup.org

1.3 Piloting implementation of a new diagnostic tool or package of tools in different settings

1.3.1 Through existing diagnostic services (routine health service provision)

Objectives:
a) To optimize implementation of a new diagnostic tool(s)/package(s) of tools (identified in **section 1.2**, above) within the health care system structure of a given country using available resources.

b) To determine required resources to facilitate equitable access and optimize patient diagnosis and outcome.

Illustrative Designs:
a) Operational and transmission modelling: If sufficient data exist on operating characteristics of new tests or approaches in terms of accuracy, health system requirements, patient-important outcomes and costs, it may be possible to predict best choices for a given national setting, rather than invest in a full pragmatic randomised controlled trial (PRCT) (see below). In some circumstance, this may have already been achieved in **section 1.2** (see above)

b) PRCT (cluster)[y]: If sufficient data on accuracy are available, but there are insufficient data on operational requirements, patient-important outcomes, costs etc., it may be appropriate to carry out a PRCT (cluster) of the chosen approach or package, comparing against an existing package or an alternative package. This can be achieved through a 'before-and-after' design (see **Annex I**) for simplicity, but concurrency of observations and randomization is important to minimize bias and confounding and to obtain robust evidence on impact to inform the next phase of scale-up. A PRCT (cluster) can also be carried out as part of scale-up or phased implementation, employing a stepped-wedge design.

Setting /study population/data sources:
a) National health-system data and predicted operating characteristics and performance of tests.

b) Within the health-care system, at the level predicted to be optimal by operational modelling, in different epidemiological situations (e.g. high or low MDR-TB, high or low HIV prevalence, high/low density populations, urban/rural etc.).

Methods: recruitment of subjects (eligibility criteria); intervention (as appropriate):
a) Operational and Transmission Modelling (as described above in 1.2).

b) PRCT (cluster) as described above and in **Annex I**. Clusters could be health units (e.g. hospitals or health centres) or whole health districts, depending on how extensive the algorithm being phased in is. For example, if a simple point-of-care test is being tested at the health centre level, then the health centre is the appropriate unit of randomization. If, however, an algorithm includes a new point-of-care diagnostic test at the health centre level, and a laboratory-based diagnostic machine at district level, for example, then the district is the appropriate unit of randomization.

Expected outcomes:
a) Projected estimates of the effectiveness of the intervention and likely health system requirements.

b) Direct evidence on effectiveness, equity of access, acceptability to patients and providers, and likely health system requirements.

Analysis:
a) Comparison of different projected configurations of a new intervention or comparison of a projected new intervention with an existing system in terms of likely effects on chosen outcome measures (such as numbers of patients starting and completing therapy, numbers of laboratory staff, etc.). The IAF (see **Annex III**) can serve as a check list for generating a relevant list of outcome measures.

b) Direct comparison of selected outcome measures resulting from a new diagnostic algorithm or package with an existing diagnostic algorithm or package. Here also, the IAF can serve as a check list for generating a relevant list of outcome measures prior to field implementation of the PRCT (cluster).

Guidance for sample size calculation:
The sample size is determined by the nature of primary outcome measure, the effect size to be detected *(11)*, and whether individual participants or clusters of individuals are to be randomised *(12)*.

[y] Here, and in the rest of the annex, PRCT (cluster) refers to a pragmatic randomised controlled trial using the cluster design (as opposed to individual randomization).

Sample size considerations are also slightly different for step-wedge designs *(13)*. **Annex I** provides an overview of these methods.

Expected duration/timeline:
a) Short-term.
b) Medium-term.

Suitable scale:
a) Local or national.
b) Regional, national or international.

Estimated budget range:
a) Low or medium.
b) High.

Illustrative references: *(14–16)*.

1.3.2 Through active TB case-finding[z]

a) Developing and evaluating ACF approaches:

Objectives: To identify the most effective and affordable approach to ACF in different settings.

Design:
Several approaches are possible, from simpler to more elaborate options, and are not mutually exclusive.

i. Review of published studies on ACF in the general community and targeted populations.

ii. 'Rapid surveys' or more formal prevalence surveys for undiagnosed TB (providing assessment of likely impact of ACF) *(17–20)*. Ideally combined with assessment of whether or not ACF can be effectively targeted on individual risk factors (e.g. age, HIV status, diabetes, etc.) through estimate of the population-attributable fraction of risk factors. This can be one selected target population being considered for ACF, or a random cluster sampling of the general population *(21)* to give more generalizable results.

iii. Prospective observational (one ACF strategy) or comparative study (two different ACF strategies) in a few high risk groups and communities known to have a high incidence or prevalence of TB.

iv. More elaborate studies (cluster-randomized trials) will give better quality evidence, but require considerably more resources and should be preceded by pilot studies (as in ii, above) wherever possible. Cluster-randomized trials require considerable planning and data analysis, but can be delivered pragmatically (through community worker routine services) and are by far the best design for case-finding interventions against infectious diseases (owing to their ability to capture the secondary benefits from reduced transmission rates). However, the number of cases found will not provide an appropriate end-point in long-term interventions, because a highly successful intervention will lead to a *fall* in new incident cases. Examples of cluster randomized trials include references at district and sub-urban levels *(22–24)*.

v. Most questions around the impact of ACF are *not* ideally tackled through individually randomized trials (because TB diagnosis in one individual provides benefit to others from prevention of TB transmission). Such trials may, however, provide evidence of reduced mortality from early TB case-detection *(25)*.

Setting /study population: High risk communities, such as high density urban communities known to have high TB case-notification rates and/or access problems.

Methods: Cross-sectional surveys, with or without random sampling (using satellite maps, community health worker catchment areas and census data to provide the sampling frame), analysis of interventional research, ideally combined with qualitative research, will require analysis of the cumulative yield of cases, ideally including cases diagnosed from the same populations through the routine health services during the same time period.

Expected outcomes: Estimated prevalence of undiagnosed symptomatic TB, patient diagnostic rate, cost-effectiveness (as measured by cost-per-case found) and yield (cases per 1000 population, and percentage of cases found through ACF and routine services during a fixed time period). Yield could be stated in terms of the 'number needed to screen' (NNS) to detect one case.

[z] Further background and resources about active case-finding (ACF) and measuring effects of ACF on TB incidence are available in **Annex V**.

Analysis: Cross-sectional analysis (with adjustment for design effect if cluster-sampling or household sampling is used), cumulative yield of TB from ACF adjusted for population size, economic analysis (cost-benefit and cost-effectiveness).

Expected duration/timeframe: Medium-term.

Suitable scale: Sub-national; single defined settings (e.g. slums, prisons)

Estimated budget: Low to high depending on approach chosen and study design.

Illustrative references: *(17–25)*.

b) Assessing the effect of sustained ACF on TB incidence:

Measuring the effect on TB incidence of ACF strategies that have been chosen and supported for sustained implementation is essential for long-term TB control strategies. Reduced duration of symptoms before diagnosis (patient delay), substantial numbers of TB cases diagnosed through ACF, and an increase in the total numbers of TB cases diagnosed (active plus routine) from the whole community can be used to show success *initially*. However, one-off interventions will not have any lasting impact on TB control, so long-term success requires a sustained intervention. So, after an initial peak of cases (corresponding to improved case-finding), the aim is to lead altogether to reduced numbers of total TB cases identified through ACF plus routine services. A falling yield from ACF could mean *either* successful TB control *or* implementation/intervention failure, so it is important to ensure that delivery of the intervention is monitored and that trends in *total* case-notifications are reliably measured. Alternatively, a 'before-after' evaluation of undiagnosed TB can be used to measure effect on TB incidence.

Objectives: To identify impact of sustained ACF interventions.

Design:

(i) Data analysis from TB control registers, so as to examine trends in overall TB case-notification rates.

(ii) Cross-sectional prevalence surveys (before and after ACF implementation) to assess trends in undiagnosed TB in the community.

Setting /study population: Communities in which sustained ACF interventions are being implemented.

Methods:
(i) Trends in TB case-notifications with evaluation to confirm that trends provide a proxy indicator of true TB incidence.

(ii) 'Before-after' intervention cross-sectional surveys for undiagnosed TB, looking for a substantial decline during the course of the intervention.

Expected outcomes: Evaluation of the impact of ACF interventions on TB control (case-notifications or prevalence of undiagnosed disease).

Analysis: Time-trend analysis of TB case-notification rates (combined ACF and routine services), or cross-sectional 'before-after' prevalence survey methods, adjusted for confounding variables (including age, gender, HIV infection or other locally important individual risk factors for TB disease, and factors such as household crowding). Process evaluation of implementation of sustained ACF programmes.

Guidance for sample size calculation: for (ii), sample size will be very large and highly dependent on baseline incidence. Outcome evaluation will be facilitated by taking a unit that is already used for TB monitoring and evaluation (e.g. a district).

Expected duration/timeframe: Medium- to long-term – needs to be nested into ongoing ACF interventions.

Suitable scale: Sub-national.

Estimated budget: ACF intervention costs will determine overall budget. Additional time-trend evaluation costs are low; before-after survey for undiagnosed TB costs are high.

Illustrative *(26)*.

1.4 Evaluating the impact of scale-up of a new test or new package of tests

1.4.1 Modelling expected impact and implications of scale-up

Objective:
a) To model and forecast the operational requirements, full economic costs and the clinical and epidemiological effects of going to scale with

a new test or intervention package (including ACF approaches) that has been trialled in 1.3 (above) from the health system, patient, and societal perspectives.

b) To critically appraise a new intervention or algorithm (e.g. those piloted in Section 1.3) against other interventions available internationally or that may become available for uptake in the short- to medium-term.

Design:
a) Operational and transmission modelling.

b) Use operational and transmission modelling based on available data on new interventions at earlier stages in the diagnostic pipeline.

Setting/study population:
a) Population targeted for new diagnostic test(s) or intervention package in the health facility provision of the region or country.

b) As for a) but assessed against alternative diagnostic intervention options.

Methods:
a) Transmission and operational modelling as already described and defined, but now using actual data from an approach that has been trialled in section 1.3.

b) As a) but using existing data on alternative diagnostic intervention options.

Expected outcome:
a) Projection of operational requirements and costs along with expected impacts on TB transmission.

b) Facilitated decisions on uptake into policy and practice.

Analysis:
a) Operational and forecasting modelling along with health economic projections.

b) Comparison of available information on alternative diagnostics (including those that are still at early stages of development and not yet endorsed by WHO) with the test currently being considered for adoption (note: this may be a test that has recently been piloted through the process described in **section 1.3** and is now being projected for national scale-up).

Guidance for sample size calculation:
a) N/A. Requires primary data from implementation research and demonstration studies carried out in **section 1.3**.

b) as for a).

Expected duration/timeline: a) and b): Short-term.

Suitable scale: a) and b) National or international.

Estimated budget range: a) and b) Medium.

Illustrative references: *(27–29)*.

1.4.2 Assessing the impact of a new test or new diagnostics package

Objective: To document the effects of going to scale with new intervention or package.

Design: Audit of NTCP diagnostic and treatment registers to examine numbers of patients starting treatment and achieving favourable treatment outcomes before and after going to scale with new diagnostic intervention.

Setting /study population: Full national datasets.

Methods: Cohort analysis: comparison of case-finding (case notification primarily, but case-detection rate if possible), and treatment outcome indicators before and after going to scale with the new diagnostic intervention.

Expected outcome: Evaluation of case-finding (case notification primarily, but case-detection rate if possible) and treatment outcome indicators against those expected or predicted in forecasting (see **section 1.4.1**).

Analysis: Comparison of patient cohorts before and after intervention.

Guidance for sample size calculation: Full national datasets could be used for this if raw data in electronic form are available, comparing at least a year before implementation with at least a year afterwards. If not, then randomized selection of registers from at least 10% of the diagnostic and treatment centres could be undertaken, but still comparing at least a year's worth of data either side of an implementation date. This is very similar to usual programme evaluation.

Expected duration/timeline: Short-term.

Suitable scale: National or international.

Estimated budget range: Medium.

TABLE 1: SUMMARY

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
1.1 Situation analysis					
Identify local barriers to accessing TB diagnosis	a) Audit and Retrospective review	Estimates of proportion of "drop-out" patients	Short-term	Local	Low
	b) Qualitative research	Ranked lists of barriers		Local, regional or national	Medium
	c) Mapping of facilities	GIS maps		Local, regional or national	Medium
	d) Costing surveys	Economic barriers		Local, regional or national	Medium
1.2 Identifying new programmatic approaches					
a) To understand factors facilitating or hindering effectiveness of existing diagnostic algorithms	a) Realist review	Lessons learned	Short-term	Local, national and international	Medium
b) To document accuracy of new diagnostic tests or packages of tests	b) Systematic review	Synthesised data			
c) To project a number of suitable options of diagnostic approaches or packages for potential implementation.	c) Operational and transmission modelling	A number of suitable options of diagnostic approaches or packages giving estimates of resource requirements			
d) To select, at international or country level, new or revised diagnostic approaches to be piloted	d) National or International level Expert Group Review meeting	Local evidence-based programmatic approach			
1.3 Piloting implementation of a new diagnostic tool or package of tools in different settings					
1.3.1 Through existing diagnostic services					
a) Optimize implementation of a new diagnostic tool/package	a) Operational and transmission modelling	Projected estimates of effectiveness and health system requirements	Short-term	Local or national	Low or medium
b) Determine required resources	b) Pragmatic cluster-randomised controlled trial	Direct evidence on effectiveness, equity of access, acceptability and health system requirements	Medium-term	Regional, national or international	High

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
1.3.2 Through active case finding (ACF)					
a) Identify the most effective and affordable approach to ACF in different settings	a) Cross-sectional surveys, analysis of interventional research including qualitative research	Estimated prevalence of undiagnosed symptomatic TB, patient diagnostic rate, cost-effectiveness and yield	Medium-term	Subnational	low to high
b) Identify impact of sustained ACF interventions	b) Trends in TB case-notifications and before-after intervention cross-sectional surveys	Evaluation of the impact of ACF interventions on TB control	Medium to long-term	Subnational	low to high
1.4 Evaluating the scale-up impact of a new test or new package of tests					
1.4.1 Modelling expected impact and implications of scale-up					
a) Forecast operational requirements including costs and impact on transmission	a) Simulation and transmission modelling	Projection operational requirements including costs and impact on transmission	Short-term	National or international	Medium
b) Critically appraise a new intervention or algorithm against other interventions	b) As for (a) but using existing data on alternative diagnostic intervention options	Facilitated decisions on uptake into policy and practice			
1.4.2 Assessing the impact of a new test or new package of tests					
Document effects of going to scale	Cohort analysis	Evaluation of case-fining and treatment outcome indicators	Short-term	National or international	Medium

2. Developing sustainable collaboration with all care-providers for TB care and control

2.1 Improving and scaling up existing approaches to engaging all providers

Objective(s): Develop an evidence base of different PPM models and approaches to scale-up that includes contextualized analyses of reasons for success/failure as well as mechanisms to create demand for quality services.

Illustrative designs:

a) Realist reviews to consolidate evidence on the factors contributing to the success (or failure) of different PPM models in meeting PPM objectives and/or achieving scale-up (international).

b) Case studies to describe how and why scale-up of different models was or was not achieved.

Setting /study population:
For realist reviews, the setting will be PPM programmes.

Methods:

a) Realist review: Synthesis of evidence from peer reviewed literature, national reports and other documents detailing the implementation and outcomes from various PPM programmes. The evidence base will be expanded through national-level studies outlined below.

b) Case studies that can describe the type of PPM model implemented with information on the partners involved; the costs of the programmes; the successes and challenges encountered, together with the specific experiences with scale-up. Programmatic data from TB registers will help to quantify the effectiveness of the programme, but qualitative methods will also be required (e.g. key informant interviews with national TB programme managers, PPM focal points and private or other public sector partners; focus groups or individual interviews with patients).

Expected outcomes: Broader knowledge for different models of PPM with understanding of factors leading to success (on which to build) and of factors leading to failure (which should be avoided).

Analysis:
a. Evidence synthesis (e.g. in partnership with members of the PPM subgroup of the Stop TB Partnership).

b. Case studies analysed by thematic area (e.g. use of private providers in case-finding, diagnosis, treatment, DOT; management of PPM programmes; quality assurance; training; issues concerning access to care etc.).

Guidance for sample size calculation: N/A

Expected duration/timeline: Short-term.

Suitable scale: International; national/provincial.

Estimated budget range: Medium.

Note: Areas of enquiry which feed into this objective include: i) assessing enablers and incentives for different care-providers; ii) assessing different mechanisms to fund scale-up; iii) identifying locally appropriate approaches to scale-up, including PPM/PAL integration; iv) assessing mechanisms to create demand for quality services; and v) assessing quality of TB care and control, using the International Standards for TB Care as the yardstick, as initiatives are taken to scale.

Illustrative references: *(7, 30)*.

2.2 Measuring the contributions of different provider groups to TB care and control

2.2.1 Contribution of different care-providers to TB control

Objective: To assess contributions of different care-providers to TB control.

Design: Time series or cross-sectional analyses of existing data sources including the following: chronic cough, outpatient or other registers recording sources (types of provider) of referral at public and private sites together with laboratory and TB registers, TB prevalence survey datasets.

Setting: TB referral centres, diagnosis and treatment registration sites; NTCP records database (if sufficiently detailed).

Methods: User and provider surveys; surveys of drug sales; analysis of information from pharmaceutical companies; review of registers review of existing data in TB prevalence surveys indicating from which type of provider prevalent cases received diagnosis and are receiving TB treatment.

Expected outcomes: An understanding of relative sizes of contribution of different types of care-provider to case detection indicators and treatment outcome measures.

Analysis:
a) Description of relative numbers of patients cared for by different types of provider.

b) Cohort analysis of patient groups originating from or cared for by different types of provider.

Guidance for sample size calculation: Whole population studies (i.e. 100% of patient data that has already been recorded in a given implementation area or prevalence survey for a set time frame (e.g. 1 year)).

Expected duration/timeline: Short-term.

Suitable scale: National.

Estimated budget range: Low.

Illustrative references: (none yet identified).

2.2.2 Ability of different providers to improve access to care for underserved groups

Objectives:
a) To assess the abilities of different providers to improve user access, case detection and outcomes for underserved groups, and reduce diagnostic delays and costs of care.

b) To understand resource requirements for scale up.

Design:
a) Prospective cohort studies of patients detected or registered for treatment by different types of provider.

b) Economic studies of patient and provider costs.

Setting (study population):
Patients accessing TB services through different types of providers (public, private, for- and not-for-profit, formal, informal).

Methods:

a) Follow-up of patients in laboratory and TB registers (and a chronic cough register if existing) to assess drop-out between stages of care and to assess outcomes from different types of providers.

b) Assess equity through patient questionnaire surveys using, among others, the costing tool (described in **Annex IV**) with patients on intensive phase of treatment (to ensure recall around referral and diagnostic processes), ideally followed-up to treatment completion. These will provide data on socioeconomic status, care-seeking pathways, costs and benefits of provider choice. Studies to assess the resource requirements of involving each provider-type can feed into economic models of the costs of scaling-up PPM activities.

Expected outcomes:
a) Documentation of the extent to which different providers promote access by underserved patients (e.g. the poor and those with other characteristics of vulnerability or facing barriers to accessing services).

b) Documentation of the resource requirements of both underserved groups and different providers for improving access.

Analysis:
a) For cohort studies, analysis will include estimates (with 95% confidence intervals) of various treatment outcomes disaggregated by provider type, significance tests of group comparisons.

b) Cost-effectiveness analysis comparing resources used by NTCPs/the public health system in engaging different provider types; evaluation of patient care-seeking pathways through different types of care-provider and assessment of patient costs incurred along care-seeking pathways.

Guidance for sample size calculation: Data on case notifications will be from registers and will include 100% of all patients seen in the intervention and control areas over the relevant period. For patient costs sample sizes of 100-200 for each provider type are usually sufficient (depending on the homogeneity of the populations).

Expected duration/timeline: Short-term.

Suitable scale: National/local.

Estimated budget range: Medium.

Illustrative references: *(5, 31)*.

2.3 Encouraging involvement of as yet unengaged providers

This research can be conducted in two successive steps:

2.3.1 Assessing potential models and approaches involving as yet unengaged providers

Design:
a) Literature or realist review to learn from other country experiences of working with a broader range of providers. Learning from the outcomes of different approaches outlined in section 2.1 (Refer to **section 2.1** above for methods, costs etc).

b) Expert group meeting at national level to discuss results from a) in relation to local experience and suggest rational approaches for engaging additional provider groups or types.

Expected outcome: Ranked priority list of types of as yet unengaged providers to include in 2.3.2.

2.3.2 Evaluating effectiveness of models involving new providers

Objectives:
a) To locate potential new providers that could provide accessible and effective services.

b) To assess effectiveness of partnerships involving new providers.

Design:
a) Brainstorming with key stakeholders +/- GIS mapping of new providers.

b) Implementation study; potential for PRCT (cluster).

Setting/study population: Population groups accessing TB care via selected new providers and populations accessing the public services directly.

Methods:
a) Mapping of as yet unengaged providers, such as informal providers, or employers of large workforces and other organizations. Mapping may be a precise process using GIS tools, or less precise, with a brainstorming process with key stakeholders to think of potential providers and marking approximate geographical areas where they may be located.

b) PRCT (cluster), to evaluate the extent to which changes in health care provision will be effective in the usual conditions under which these changes will be applied. As indicated in **Annex I**, the PRCT (cluster) design in this context may require expert statistical and epidemiological input *(32)*. However, if health-care providers are reluctant to engage in a study where they may be randomly allocated to not receive the new changes in health provision, the option of using a 'before-and-after' design (see **Annex I**) with either a single or multiple clusters may have to be considered on practical grounds – but the disadvantages of not having a contemporary control should be carefully considered and discussed with the participating health providers.

Expected outcomes:
a) Compilation of potential additional providers.

b) Information on effectiveness of new providers in terms of case-detection indicators (e.g. case notifications; case referral rates) and treatment outcome measures (treatment success rates, mortality etc); as well as cost implications for the service.

Analysis:
a) Gap analysis in mapping exercises.

b) Comparison of case-detection indicators and treatment outcome measures among population groups accessing TB care via new providers against the same indicators and measures among the population groups accessing public services directly. Then, conduct cost-effectiveness analysis comparing resources used by NTCPs/the public health system in engaging different provider types; evaluation of patient care-seeking pathways through different types of care-provider and assessment of patient costs incurred along care-seeking pathways.

Guidance for sample size calculation:
a) For qualitative research and mapping, conduct interviews and focus group discussions to saturation point.

b) For implementation studies, including PRCT (cluster) – see previous sections.

c) For PRCT (cluster) or 'before-and-after' designs, a minimum of 100 cases/respondents per arm (where an arm may be an area with newly engaged providers or one without) is usually required to compare patient costs. This is usually only a proportion of the total number of participants in each arm of the study.

Expected duration/timeline: Short- to medium-term.

Suitable scale: National/local.

Estimated budget range: Medium.

Illustrative reference: *(33)*.

2.4 Encouraging involvement of non-public sector in MDR-TB management and TB/HIV collaborative activities

Objectives: To develop an evidence base of different models and approaches through the identification of potential providers that could provide accessible and effective services for MDR-TB and TB/HIV management; and to assess the effectiveness of models for PPM for MDR-TB and TB/HIV.

Design:
a) Realist review (international level or national with international collaboration).

b) Observational cohort study; potential for PRCT (cluster).

Setting/study population:
For realist reviews the population will be PPM programmes.

For cohort studies, the study population will be MDR-TB or TB/HIV co-infected patients in study areas pre- and post- implementation.

Methods:
a) Realist reviews from **section 2.1** (see above) may be used to assess which providers would be best suited for engagement in MDR-TB or TB/HIV PPM programmes. Specific challenges related to MDR-TB or TB/HIV service delivery should be considered against reported challenges of the PPM programmes under consideration.

b) Implementation cohort studies will follow methods described in earlier sections (particularly **section 2.3**) and should include assessments of patient and health-system costs. The cohorts will be defined by clinical status (e.g. MDR-TB or TB/HIV co-infection) and types of provider.

Expected outcomes:
a) Compilation of potential PPM models suited for management of MDR-TB or TB/HIV co-infection, along with key facilitating factors.

b) Information on effectiveness of new providers in terms of case-detection indicators (e.g. case-notifications; case-referral rates etc.) and treatment outcome measures (treatment success rates, mortality etc.).

Analysis:
a) Capturing key features of success or lack of it in existing PPM models when appraised for suitability for TB/HIV or MDR-TB.

b) Descriptive statistics with their (95%) confidence intervals for case-detection and treatment outcome indicators, and appropriate for comparing rates between patient and provider groups (e.g. Fisher exact test for simple group comparisons, logistic regression for adjustment for important confounding factors); and cost-effectiveness analysis.

Guidance for sample size calculation:
a) N/A

b) For implementation studies, including PRCT (cluster) – see previous sections.

Within PRCT (cluster) or 'before-and-after' designs a minimum of 100 cases/respondents per arm (where an arm may be an area with newly engaged providers or one without) is usually required to compare patient costs. This is usually only a proportion of the total number of participants in each arm of the study.

Expected duration/timeline: Short- to medium-term.

Suitable scale: National/local.

Estimated budget range: Medium.

Illustrative reference: *(33)*.

2.5 Developing and assessing responses to changing involvement of diverse providers in TB care and control

Objective: To identify and assess ways to ensure rational use of new diagnostics and drugs in the private sector.

Design: Realist reviews/evaluations; qualitative studies.

Setting (study population): Providers, drugs manufactures/distributors.

Methods: Methods include structured evaluation of existing approaches; qualitative research to better understand providers' practices; 'mystery client' surveys[aa] to assess use of inappropriate techniques; cost-effectiveness studies with a sample of anonymized private providers.

Expected outcome: Promotion of rational use of appropriate diagnostics and drugs in the private sector.

Analysis:
Evidence synthesis, cost-effectiveness analysis for individual studies; quality of care analysis from mystery client surveys.

Guidance for sample size calculation: For qualitative research key informant interviews, focus groups and mystery client studies will be conducted to point of new knowledge saturation (typically 20-30 interviews); for cost-effectiveness studies number of providers interviewed will depend on range of services offered, but for homogenous types of providers at similar levels of service provision 4-6 providers should be sufficient.

Expected duration/timeline: Short-term.

Suitable scale: Local, national and international.

Estimated budget range: Mostly low.

Note: This area will be difficult to undertake comprehensively due to a potential lack of willingness to share commercially-sensitive information and to reveal practices of questionable quality. Social marketing organizations (such as Population Services International)[bb] have developed good techniques for assessing quality of care in private providers and so could support this area. Identifying the scale of the problem will be a particular challenge and may require some projections made from discussions of sales of diagnostics and drugs with company representatives or local pharmacists.

Illustrative references: (see footnote below)

2.6 Encouraging introduction of novel regulatory approaches to collaborating care-providers

2.6.1 Developing an evidence base of regulatory approaches that includes contextualized analyses of reasons for success/failure

Design: Realist review.

Setting/study population: National regulatory and enforcement bodies.

Methods: Includes structured evaluation of existing approaches (such as mandatory TB case notification, certification and accreditation) through realist review; qualitative research to better

[aa] A mystery client survey involves an actor presenting him or herself as a TB suspect to a provider to assess aspects of the quality of care.

[bb] See: http:// www.psi.org

understand providers' practices. Lessons may be drawn from experiences in other countries or in other sectors in the same country (e.g. control of pesticide use by the ministry of agriculture).

Expected outcomes: Inventories of regulatory approaches that are present and functioning a) locally, and b) in other settings.

Analysis: Analysis of benefits of different regulatory approaches, challenges and costs associated with implementation and enforcement.

Guidance for sample size calculation: N/A

Expected duration/timeline: Medium-term.

Suitable scale: International and national.

Estimated budget range: Medium.

Note: Full attention must be paid to contextual factors that support or inhibit the success of regulation implementation in other countries or sectors.

2.6.2 *Developing locally-appropriate regulatory approaches*

Design: Qualitative, consultative studies.

Setting/study population: Regulatory authorities, enforcers and service providers.

Methods: Qualitative research to understand providers' practices, including participant observation, to understand practices and willingness to comply with regulation; policy analysis to assess ability of the state/others to encourage and ensure compliance with regulations.

Expected outcomes: Enforceable regulations that improve quality of care and public health outcomes.

Analysis: Political, social, cultural and legal.

Guidance for sample size calculation: N/A

Expected duration/timeline: Medium-term.

Suitable scale: National.

Estimated budget range: Medium.

Note: While regulations will have to apply on a national scale, different options for enforcement may be piloted**.**

TABLE 2: SUMMARY

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
2.1 Improving and scaling up existing approaches to engaging all providers					
Develop evidence-base of different PPM models and identfy factors that result in success or failure of PPM	a) Realist review b) Case studies	Broader knowledge of different models of PPM and understanding of factors which lead to success or failure	Short term	International or national/ provincial	Medium
2.2 Measure the contributions of different provider groups to TB care and control					
2.2.1 Contribution of different care providers to TB control					
Assess contributions of different care providers	Surveys, review of exisiting data	Understanding of contribution of differrent care providers	Short-term	National	Low
2.2.2 Ability of different providers to improve access to care for underserved groups					
a) Assess ability of different providers to improve access to care	a) Prospective cohort study	Documentation of the extent to which different providers promote access by underserved patients	Short-term	Local, national	Medium
b) Understand resource requirements for scale up	b) Costing questionnaires	Documentation of the resource requirements of both underserved groups and different providers for improving access			
2.3 Encouraging involvement of as yet unengaged providers					
2.3.1 Assessing the potential models and approaches involving as yet unengaged providers					
Assess the potential models and approaches involving as yet unengaged providers	a) Literature or realist review b) Expert Group meeting	Ranked priority of types of as yet unengaged providers	Short to medium-term	National	Medium
2.3.2 Evaluating effectiveness of models involving new providers					
a) Locate potential new providers	a) Brainstorming with key stakeholders +/- GIS mapping of new providers	Compilation of potential additional providers	Short to medium-term	Local, National	Medium
b) Assess effectiveness of partnerships involving new providers	b) Pragmatic cluster-randomised controlled trial	Effectivess of new providers			

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
2.4 Encouraginig involvement of non-public sector in MDR-TB management and TB/HIV collaborative activities					
a) Identify potential non-public sector providers	a) Realist review	Compilation of potential PPM models suited for management of MDR-TB or TB/HIV co-infection, along with key facilitating factors	Short to medium-term	Local, national	Medium
b) Assess effectiveness of PPM	b) Observational cohort study or potential for pragmatic randomised controlled trial	Effectiveness of new non-public sector providers			
2.5 Develop and assess responses to changing involvement of diverse providers in TB care and control					
Identify and assess ways to ensure rational use of new diagnostics and drugs in private sector	a) Structured evaluation of exisiting approaches b) Qualitative research c) Cost-effectiveness studies	Promotion of rational use of new diagnostics and drugs in private sector	Short-term	Local, national and international	Low to medium
2.6 Encouraging introduction of novel regulatory approaches to collaborating care providers					
2.6.1 Developing an evidence base of regulatory approaches that includes contextualised analyses of reasons for success/failure					
Develop an evidence base of regulatory approaches	Realist review	Inventories of regulatory approaches locally and in other settings	Medium-term	National, international	Medium
2.6.2 Developing locally appropriate regulatory approaches					
Develop locally appropriate regulatory approaches	Qualitative, consultative studies	Enforceable regulations that improve quality of care and public health outcomes	Medium-term	National	Medium

3. Prevention and treatment of TB in persons living with HIV

3.1 Optimizing linkages between TB and HIV programmes

Objectives: For HIV-infected TB patients in different groups and epidemiological settings (adults, children, families, and special at-risk populations (e.g. injection drug users (IDUs), men who have sex with men (MSM), prisoners):

a) To determine optimal strategies to integrate and deliver joint TB/HIV interventions.

b) To determine optimal models of community participation for enhanced TB case-finding and early HIV detection.

Design:
For both a) and b) above:

i) Realist review of existing models and studies, extracting, where possible cost-effectiveness data.

ii) Prospective observational and case-studies with 'before-and-after' assessment of key indicators (including HIV/TB-related detection and treatment outcome measures, costs etc.);

iii) Qualitative studies among providers and patients.

Setting/study population: General population; HIV clinics; special services for at-risk populations (e.g. IDUs, prisoners).

Methods: i) Realist review; ii) Observational case studies; iii) Focus group discussions, key-informant interviews, and in-depth interviews.

Expected outcomes: Documentation of successful and less successful models at health system and community level in different settings.

Analysis:
(i) and (iii) : qualitative analysis

(ii) compare effectiveness estimates from before/after assessment of HIV/TB-related case-detection and treatment outcomes (capturing patient/community costs, and health-system costs).

Guidance for sample size calculation (and estimated number of participants): N/A

Expected duration/timeline: Short-term for (i) and (iii), and medium-term for (ii).

Suitable scale: National/regional.

Estimated budget range: Low for (i) and (iii), and medium for (ii).

Illustrative references *(34)*.

3.2 Assessing the validity of TB screening algorithms in different settings

Objective: To evaluate the effect of implementing the WHO-recommended algorithm (or other algorithms) for TB screening compared with current policy on TB screening in different groups of PLHIV.

Illustrative Designs:
a) Observational (before/after) case studies.

b) PRCT (cluster) +/- stepped-wedge design.

Qualitative studies of patient and provider acceptability may be nested within a) and b).

Setting/study population: PLHIVs in different settings (these may be within health units or clusters for the purposes of the research design) where TB screening may occur including HIV counselling and testing, HIV clinics, community-based case-finding and household contact investigations.

Methods:
Comparison of HIV- and TB-related case-finding indicators (notifications, time to starting intervention) and treatment outcome measures (including mortality):

a) At baseline and after implementation in observational studies.

b) Between different algorithms in PRCT +/- step-wedge design.

Expected outcomes: Evidence of relative effectiveness of different algorithms, along with evidence of impact on access, acceptability and health system requirements.

Analysis:
a) Primary effectiveness analysis: number of additional cases starting TB treatment or IPT; change in time to starting TB treatment or IPT.

b) PRCT (cluster), to evaluate the extent to which changes in health care provision will be effective in the usual conditions under which these changes will be applied – the use of this design may require expert statistical and epidemiological input *(32)*.

c) Equity analysis: Stratified analysis of patients being screened (gender, age, socioeconomic status) – particularly examining high-risk or vulnerable groups; acceptability of screening algorithm to patients (qualitative); health system analysis (e.g. human resources, logistics, infrastructure and other cost-incurring requirements for implementation of alternative algorithms).

Guidance for sample size calculation:
Sample size calculation will have to take account of the likely intra-cluster correlation (ICC) if cluster sampling/randomization is used; few published studies report their ICC so precedent is rare – typically ICC ranges between 0.001 and 0.100, but can be larger. Sample size varies considerably according to nature of primary outcome measure, effect size to be detected, ICC and number of available clusters – it is better to have many small clusters than a few large clusters *(35)*.

Expected duration/timeline: Medium-term.

Suitable scale: Regional, national.

Estimated budget range: Medium.

3.3 Optimal timing of IPT in relation to ART

Objective: To define the optimal time to start, duration, safety, efficacy and cost-effectiveness of IPT in patients eligible for IPT and ART.

Illustrative designs:
a) Observational cohort study.

b) PRCT (individual) or PRCT (cluster).

Setting (study population): Patients eligible for both IPT and ART.

Methods: In both a) and b) this would be a comparison between patients on ART given IPT for varying lengths of time. In a) there would be no actual intervention and health provider-patient interactions would determine the duration of IPT based on local recommendations. In b) there would be a randomized assignment of patients or health units to the different IPT duration strategies. Both will require long-term follow-up to determine timing of adverse events.

Expected outcomes: Evidence on effectiveness and cost-effectiveness of different durations of IPT in patients receiving ART, as determined by the occurrence of pre-defined end-points (e.g. death, development of TB, etc.).

Analysis:
a) Observational design (note: adjustment for unmeasured confounding will be required).

b) PRCT (cluster) design: effectiveness analysis (including actuarial methods/Cox regression models) – mortality as the main outcome measure. Equity analysis and health system analysis analogous to that outlined in **section 3.3** above could also be added.

Guidance for sample size calculation:
a) Sample size for rate(s) of occurrence of pre-defined end-points (e.g. death, development of TB) based on either precision required for estimate of rate(s) or on size of difference between patient sub-groups that would be considered clinically significant.

b) Sample size for 'time to' measures based on difference in median time to end-point (e.g. death, development of TB) between groups that would be considered clinically significant.

Expected duration/timeline: Long-term.

Suitable scale: Regional, national, international.

Estimated budget range: High.

3.4 Models to improve adherence to IPT

3.4.1 Identifying optimal models

Objective: To identify optimal medication delivery, community support and clinical monitoring of IPT in PLHIV to maximize adherence.

Illustrative Design: PRCT (cluster) with factorial design.

Setting/study population: PLHIV initiating IPT.

Methods: The cluster here is a health unit. The choice of adherence support mechanisms should be based on existing synthesised evidence *(36)*:

- clusters in Arm 1: receive no additional adherence support between routine clinic visits;

- clusters in Arm 2: receive community-peer support for adherence through patient-support groups;

- clusters in Arm 3: other support mechanisms (for instance, weekly text reminders via mobile phones).

Expected outcome: Evidence on relative effectiveness (most feasibly in terms of IPT default rates, but can also include pills count, self-reported adherence, or measurement of INH metabolites in urine), patient acceptability, health system requirements of different adherence support mechanisms.

Analysis:
- Effectiveness analysis (suggested outcome measure = IPT default rate. Incidence of TB disease is relatively rare and would likely require a very large sample size).

- Equity and health systems analysis – analogous to that described in Section 3.4 above.

Guidance for sample size calculation:
A large number of clusters (areas) and number of default patients is required to ensure generalizable results; many small clusters are preferable to a few large clusters. Stratification by urban/rural areas may be appropriate.

Expected duration/timeline: Medium.

Suitable scale: Regional, national.

Estimated budget range: Medium.

3.4.2 Identifying operational requirements

Objective(s): To model and forecast the operational requirements and full economic costs of going to scale in administering IPT in HIV care settings.

Design: Modelling study.

Setting/study population: PLHIV initiating IPT.

Methods: Operational research modelling to include optimal frequency of symptom screening, monitoring tools and measures to maintain high adherence. This approach corresponds to layer 4 of the Impact Assessment Framework (see **Annex V**) and requires empirical data (including data on costs) from studies outlined in section 3.5.1.

Expected outcome: Assessment of operational requirements and costs.

Analysis: Operational modelling.

Guidance for sample size calculation: N/A

Expected duration/timeline: Medium-term.

Suitable scale: National.

Estimated budget range: Medium.

3.5 Optimizing infection control to reduce TB transmission

Objective: To determine the impact on nosocomial, congregate and household TB transmission in HIV-prevalent settings with the introduction of TB infection control measures based on recommendations from WHO.

Design: PRCT (cluster) trial with stepped-wedge design, ensuring that all units ultimately receive the WHO-recommended TB infection control measures.

Setting (study population): Health-care staff working at facilities that provide chronic HIV care; staff working at congregate settings with long-term (e.g. prisons) and short-term (e.g. jails and homeless shelters) duration of stay of dwellers and residents of households in which at least one HIV-positive person resides.

Methods:
- *Intervention clusters*: WHO recommendations on TB infection control.

- *Control clusters*: current policy and practice of TB infection control, until such time as the WHO recommendations can be rolled out to these.

- Information on known TB exposure, if any, should be obtained using interviews, log books and staff rotas. TB infection rates should be monitored through serial testing in the control and intervention arms using either the tuberculin skin test (TST) or the interferon gamma release assay (IGRA).[cc] The ultimate end-point is the number of TB cases among exposed populations over a defined period of time.

Expected outcomes: Evidence on effectiveness of standardized pack of infection control interventions, along with acceptability to health-care workers and impact (costs) on health system.

Analysis:

- Primary effectiveness analysis: Mean numbers of cases of TB in exposed and unexposed populations, rate ratio or rate difference (all with 95% confidence intervals). Appropriate regression methods to adjust rates for important confounding factors. Analyses may have to be adjusted also for clustering effects. For details, see **Annex I**.

- Qualitative research with health-care workers/other staff to understand acceptability of IC interventions.

- Health system analysis, including cost-effectiveness analysis (infrastructure, utilities, etc requirements.

Guidance for sample size calculation (and estimated number of participants):

Depends on size of clusters, value of ICC, background rate of TB incidence among health care-workers/other staff, and size of reduction in incidence of TB infection and disease that will constitute a clinically significant effect.

Expected duration/timeline: Medium-term.

Suitable scale: Regional, national, multi-country.

Estimated budget range: High.

Illustrative references: *(37, 38)*.

3.6 Reducing mortality in TB/HIV co-infected patients

Objective(s): To identify the risk factors associated with death and the causes of death in PLHIV being treated for TB.

Design: Prospective (routine clinical surveillance) survey.

Setting (study population): HIV-positive patients pre-ART and on ART commencing treatment for TB (programmatically defined, or, if resources allow – microbiologically confirmed).

Methods: Prospective clinical assessment (routine surveillance). If routine follow-up of patients is adequate and if records are sufficiently complete, it may be possible to use retrospective case notes or health register data.

Expected outcomes: Proportions of patients surviving fixed time periods (with 95% confidence intervals); hazard ratios/medians (with 95% confidence intervals) for survival times, stratified on CD4 count results. List of major risk factors associated with mortality in PLHIV treated for TB (examples to consider should include objectively measurable indicators of clinical condition at the start of treatment, such as body mass index, pulse rate and respiratory rate, undiagnosed MDR-TB, as well as general performance indicator) *(39)*.

Analysis: Actuarial/Kaplan-Meier plots of survival times for all patients and for important sub-groups. Fisher exact tests to compare early death rates (e.g. during the first two weeks or during the intensive phase) between sub-groups and logistic regression methods to identify risk factors associated with mortality rates at fixed time points. Log-rank tests

[cc] The issue with IGRAs on serial testing is whether an increment threshold is required to determine a negative to a positive test result. The advantage over TST is that there is no boosting effect as it is an ex-vivo test, but variability around the threshold could be an issue when conversion results are around the threshold. TST is not a great tool for serial testing because of the boosting effect and the fact that people are not keen on a regular in vivo test. If neither test is being used, transmission assessment will then be made only on incident cases, so the impact on infection rates will not be assessed.

to compare (median) times to death between sub-groups and Cox regression methods to identify factors associated with survival time (e.g. during the continuation phase); analyses should be terminated when numbers at risk fall below 30 *(40)*. These methods may require expert statistical advice.

Guidance for sample size calculation:

Depends on completeness and reliability of information on deaths and causes of deaths, current death rate, and the risk factors on which data is being recorded. Sample size for mortality rates at fixed time points based on size of difference that would be considered clinically significant; sample size for times to death based on difference in hazard rates that would be considered clinically significant and requires use of specialized formulae/software *(41)*.

Expected duration/timeline: Medium-term.

Suitable scale: Local, regional, national.

Estimated budget range: Medium.

Illustrative references: *(42–45)*.

TABLE 3: SUMMARY

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
3.1 Optimizing of linkages between TB and HIV programmes					
a) Determine optimal strategies to integrate and deliver joint TB/HIV interventions b) Determine optimal models of community participation for enhanced TB case finding and early HIV detection	(for both objectives) i) Realist review, ii) observational case studies and iii) qualitative research	Documentation of successful and less successful models at health system and community level in different settings	(i) and (iii) short term, (ii) medium term	National, regional	(i) and (iii) low, (ii) medium
3.2 Assessing the validity of TB screening algorithms in different settings					
Evaluate the effect of implementing the WHO recommended algorithms for TB screening in PLHIV compared to current policy	a) Observational studies including qualitative studies of patient and provider acceptability b) Pragmatic randomised controlled trial including qualitative studies of patient and provider acceptability	Evidence of effectiveness and impact of different algorithms	Medium-term	Regional, national	Medium
3.3 Optimal timing of IPT in relation to ART					
Define optimal duration, safety, efficacy and cost-effectiveness of IPT in patients eligible for IPT and ART	a) Observational cohort study b) Pragmatic randomised (individual or cluster) controlled trial	Evidence on effectiveness and cost-effectiveness of different durations of IPT in patients receiving ART	Long-term	Regional, national, international	High

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
3.4 Models to improve adherence to IPT					
3.4.1 Identifying optimal models					
Identify optimal medication delivery, community support and clinical monitoring of IPT in PLHIV to maximise adherence	Pragmatic cluster-randomised controlled trial with factorial design	Evidence on relative effectiveness, patient acceptability and health system requirements of different adherence support mechanisms	Medium-term	Regional, national	Medium
3.4.2 Identifying of operational requirements					
Forecast operational requirements and full economic costs of going to scale in giving IPT in HIV care settings	Operational modeling	Operational requirements and costs	Medium-term	National	Medium
3.5 Optimizing infection control to reduce TB transmission					
Determine impact on TB transmission in HIV prevalent settings with introduction of WHO TB infection control policy	Pragmatic cluster-randomised controlled trial with step-wedge design	Evidence on effectiveness, acceptability and cost of infection control interventions	Medium-term	Regional, national, international	High
3.6 Reducing mortality in TB/HIV co-infected patients					
Identify risk factors associated with death in PLHIV and being treated for TB	Prospective clinical survey	Identification of modifiable risk factors	Medium-term	Local, regional, national	Medium

4. Treatment of drug-susceptible and M/XDR-TB: optimal access, delivery and community participation

4.1 Identifying reporting gaps

Objective:
To determine the validity of reported treatment outcomes of new smear-positive pulmonary TB, and re-treatment pulmonary TB.

Design:
Comparison of routinely reported treatment outcomes from programme cohort analysis with in-depth audit of registers and records. For example, within a given diagnostic and treatment centre, detailed treatment outcomes can be re-calculated from the following primary data sources:

a) Individual patient treatment cards – a random selection of treatment cards can be extracted and treatment outcomes in the standard outcome categories defined so that outcomes can be estimated and compared with the overall cohort report from the same time period.

b) laboratory registers can be compared with treatment registers (for a selected time-frame) to identify patients who are documented as smear-positive by the laboratory services but who do not start treatment (primary defaulters). This kind of analysis can highlight areas in the patient pathway to and through treatment that either needs stronger documentation, programme strengthening or both.

c) Reference laboratory registers can also be scrutinized for data on drug resistance.

Setting/study population:
Patients' records: individual treatment cards, laboratory registers, treatment registers, and cohort reports.

Methods:
Record and register reviews. Analysis of treatment outcomes can be disaggregated by different patients' populations: e.g. new/re-treatment, male/female, urban/rural residence etc.

Expected outcomes: In-depth understanding of gaps in recording and reporting, improved accuracy of treatment outcome data, indicative areas requiring additional programme support.

Analysis:
a) and b): Comparison of existing treatment outcome reports from routine cohort reporting with re-calculated treatment outcomes derived from in-depth analysis of different data sources (laboratory registers, treatment registers, patient treatment cards).

c): Descriptive statistics of primary and secondary drug resistance rates and patterns.

Guidance for sample size calculation:
A sampling frame of all eligible health units will be needed. Random samples of records/registry entries selected for audit and a decision made about what time-span the audit will cover. An appropriate proportion of treatment cards will need to be drawn from all of the registers in each selected health unit, using either random or systematic sampling methods; sample size determined by precision required for estimate of prevalence rates for different treatment outcomes.

Expected duration/timeline: Short term.

Suitable scale: National.

Estimated budget range: Low.

4.2 Carrying out investigational studies to address deficiencies in first-line management of TB within national TB control programmes

Objective: To understand reasons behind default (both primary and secondary), poor adherence, missed doses, and drug stock-outs.

Illustrative Designs:
a) Cross-sectional descriptive qualitative studies with patients and providers of first-line TB treatment within NTCPs.

b) Mapping of service provision for treatment initiation and support in relation to population characteristics.

Setting/study population:
a) Patients currently on TB treatment, and NTCP staff or first-line providers within the public health service who procure, distribute stock, prescribe, and dispense TB drugs. The primary focus would be on districts or health units with higher than usual adverse treatment outcomes, but identical studies in better functioning districts or health units are also worthwhile.

b) Clusters (health units or districts) with different treatment success rates.

Methods:
a) Qualitative research methods to identify factors behind known deficiencies in NTCP function (default, drug stock-outs etc). Themes will be expected to emerge from both patient and health system perspectives.

b) Mapping of facilities and services onto GIS.

Expected outcomes:
a) Context-specific understanding of reasons behind known deficiencies in NTCP function.

b) GIS maps illustrating relationship between patients, communities and treatment services.

Analysis:
a) Selection of modifiable processes behind known deficiencies in NTCP function.

b) Identification of geographical and other modifiable gaps in provision of treatment initiation and support.

Guidance for sample size calculation (and estimated number of participants):
a) No formal sample size calculations required (see details described in **section 1.1**, above).

b) N/A

Expected duration/timeline: Short-term.

Suitable scale: Local or national.

Estimated budget range: Low.

4.3 Improving management of drug-sensitive TB

Objective: To test effects of strategies emerging from studies in section 4.2 (above) to improve first-line management of drug-sensitive TB.

Design:
a) Before-and-after studies.

b) PRCT (cluster) or PRCT (individual) with stepped-wedge design.

Setting/study population:
a) and b): districts or health units with high rates of adverse treatment outcomes.

Methods:
a) Compare treatment outcomes at baseline (may be retrospectively) to changes in treatment outcomes over time after initiation of improved management strategy (e.g. improved recording of patient locators and strengthened drug procurement and distribution mechanisms).

b) Compare treatment outcomes between control districts that have not yet received improved management strategy with intervention districts that receive improved management strategy first.

Expected outcome: Estimates of effectiveness of different strategies for improved first-line management of TB.

Analysis:
If routine recording data are being used and observation periods are all equal, comparison of mean numbers of adverse outcome events between clusters and/or observation periods is feasible. If observation periods are unequal and/or times to events are known, arctuarial/ log-rank methods may be possible.

Guidance for sample size calculation:
Sample size for rate of occurrence of pre-defined outcome measures based on either precision required for estimate or on size of differences considered clinically significant.

Expected duration/timeline: Long-term, to allow time for capture of treatment outcomes.

Suitable scale: National.

Estimated budget range: Medium.

Illustrative reference: *(46)*.

4.4 Identifying major drivers of drug-resistant TB

It is important to identify the major drivers of the DR-TB epidemic so as to effectively target control interventions at regional or local level and in specific settings, and in order to monitor interventions for their impact on the DR-TB situation. Programme managers need to know the scale of the problem (prevalence, incidence), particularly among specific geographical or socioeconomic groups, the interaction with other diseases (e.g. HIV, diabetes) and other risk factors (e.g. malnutrition), and what is the relative contribution of transmission and acquisition of DR-TB. They need this information for the optimal planning and implementation of the TB control programme. Therefore, depending on the situation, the major drivers of DR-TB should be investigated either through the existence of routine drug resistance surveillance data, or through the conduct of drug resistance surveys. In both situations, data should be used to identify those drivers of the DR-TB problem that are relevant at the local or national level. These can be primarily of individual (e.g. treatment adherence, transmission), or of health service or programmatic nature (e.g. drug stock-outs or drug quality), and require different study approaches.

4.4.1 At the individual level

Objective: To identify risk factors for development of drug-resistant TB at the individual level.

Design:
a) Case-control study comparing new cases with drug-resistant isolates and new cases with drug-sensitive isolates. (Note: this would primarily yield insight into risk factors for transmission of drug-resistant *M. tuberculosis*).

b) Case-control study comparing re-treatment cases with drug-resistant isolates with re-treatment cases with drug-sensitive isolates. (Note: this would yield insight into risk factors for both transmission and acquisition of drug-resistant *M. tuberculosis*, e.g. through low treatment adherence).

Setting/study population:
- Cases (a and b): patients whose specimens grow *M. tuberculosis* and which are tested for first-line DST and found to have drug-resistant patterns.

- Controls (a and b): patients whose *M. tuberculosis* isolates are fully sensitive to first-line drugs.

Note: for the second case-control study, case and control patients should be matched by diagnostic category (i.e. treatment failure, return after default, relapse).

Methods: Clinical, demographic, socioeconomic variables should be collected in cases and controls. In particular, these should try to investigate risk factors for acquisition of drug-resistance (e.g. adherence to treatment, use of rifampin before current treatment, previous treatments – where, what, interrupted, switched, access to care), and potential determinants for transmission of drug-resistance (e.g. HIV status, history of hospitalization, frequent outpatient clinic attendance, incarceration, drug use, homelessness, alcohol use, country/region of origin, etc.).

Two approaches are possible for these case-control comparisons:

a) Within the context of a national MDR-TB survey, an investigational case-control study, when a prospective sampling plan is used to collect specimens from representative populations of new and re-treatment TB patients. In such a case, clinical, demographic, and socioeconomic variables are collected at the time of sample collection from all patients included in the survey. Alternatively, detailed data can be obtained by interviewing identified case and control patients after the survey DST results are available. A comparison of these variables between those identified with DR isolates and those with DS isolates is then possible.

b) Outside the context of a national MDR-TB survey, new patients who are identified with DR isolates during the course of treatment or when treatment is identified as having failed can be interviewed at the time that they are identified as

having drug resistance and are switched to a new treatment regimen. The interview can then capture personal, demographic, clinical and socioeconomic variables. The treatment register can then be used to find a case with drug-sensitive isolates who matches the index DR case. This same procedure can be followed for re-treatment cases identified with DR isolates.[dd]

Expected outcome: Identification of major individual risk factors for DR-TB operating at the individual patient level.

Analysis: Odds ratios for individual risk factors with their (95%) confidence intervals; logistic regression adjustment of odds ratios for important confounding variables. When large studies are done or data from various smaller studies are combined in order to explain regional/international differences, the effect of risk factors operating at different levels (e.g. individual, village/community, region, etc.) can be taken into account using multi-level modelling methods (see Annex I).

Guidance for sample size calculation:
Depends on background rates of MDR-TB, and prevalence of expected risk factors. Study power should be based on the detection of risk factors that are either strong or have a high prevalence in the study population (i.e. are strong in terms of attributable fraction). In general terms, risk factors with odds ratios of less than two will have limited clinical value; 150-200 cases and controls are needed to detect factors with odds ratios of this magnitude with 80–90% power.

Expected duration/timeline: Depending on the design: medium (surveillance-based) to high (survey-based).

Suitable scale: Local, national, international.

Estimated budget range: Low to Medium.

Illustrative reference: *(47)*.

4.4.2 Programmatic risk factors for drug-resistant TB

Objective: To identify risk factors for development of drug-resistant TB at the programmatic level.

Design:
Case-control study comparing health units or districts with high rates of (M)DR-TB, with units or districts with low rates of (M)DR-TB.

Setting/study population: Patients and programme processes in health units or districts with differing (M)DR-TB rates.

Methods: Comparison of programmatic issues in health units or districts with differing (M)DR-TB rates. Example issues requiring attention are: the use of the WHO re-treatment regimen, frequency of drug stock-outs, changes in drug suppliers, changes/lapses in treatment regimens, use of fixed-dose combinations, quality of reporting and recording (including patient-locator details (such as addresses) in diagnostic and treatment registers), hospitalization practices.

Expected outcome: Identification of major programmatic risk factors for DR-TB.

Analysis: Odds ratios for individual risk factors with their (95%) confidence intervals; logistic regression adjustment of odds ratios for important confounding variables.

Guidance for sample size calculation:
Depends on prevalence of expected programmatic risk factors. Study power should be based on the detection of high-impact factors, which by definition will have medium to high prevalence and clinically desirable impact. In general terms, risk factors with odds ratios of less than two will have limited clinical value; 150-200 cases and controls are needed to detect factors with odds ratios of this magnitude with 80–90% power. Study power may be increased by using two to four control health units/districts (with low rates of MDR-TB) for each district with a high MDR-TB rate.

Expected duration/timeline: Short-term.

Suitable scale: Local or National.

Estimated budget range: Low to Medium.

Illustrative reference: *(47)*.

[dd] The limitation of a case-control approach outside of the context of an MDR-TB survey is that many countries are currently not offering DST to new TB cases. This may change as drug sensitivity testing modalities are more widely offered.

4.5 Defining and evaluating strategies for integration/scale-up of DR-TB management within TB control programmes

Objectives:
a) To develop algorithms for selecting patients eligible for DST and second-line treatment in different settings (to identify MDR-TB early).

b) To develop strategies for provision of second-line treatment (including adherence and use of incentives and enablers, community-based ambulatory care and support).

c) To evaluate the effectiveness of existing infection control measures and strategies for implementing recommended infection control measures at community, household and health facility levels.

Design:
a) For development of appropriate diagnostic algorithms, please refer to Section 1.2 (above), where the approaches outlined for diagnostic approaches can be followed, but with specific attention to early identification of MDR-TB.

b) Case studies of community-based provision of complex, long-duration antibiotic therapies including injectable drugs. Lessons could be extracted from case studies of community-based provision of treatment in areas of the health sector other than TB care and control, and from novel pilots of community-based provision of TB care.

c) For evaluation of existing infection control measures in relation to new and recommended control measures, refer to **section 3.6**, above.

Setting/study population: Health-care facilities, congregate settings and households.

Methods:
a) See **section 1.2**.

b) Case-studies: giving detailed programme description and extracting key features of success from examples in other parts of the health sector as well as novel community-based MDR treatment programmes.

c) See **section 3.6**.

Expected outcomes: Options for provision of scaled-up MDR provision, which include: i) improved diagnostic algorithms for early detection of MDR-TB, ii) accessible and practical community-based MDR treatment arrangements, and iii) adequate infection control procedures.

Analysis:
a) (See **section 1.2**).

b) Case studies.

c) (See **section 3.6**).

Guidance for sample size calculation:
a) (See **section 1.2**)

b) A single case study may be useful in drawing out context-specific lessons for going to scale with a given model in a given country. If case studies in different countries of different models are compared, there is the possibility to draw out more generic lessons for scale-up which may be more widely applicable.

c) (See **section 3.6**)

Expected duration/timeline: Medium- to long-term.

Suitable scale: National, but with international support for case-study comparisons.

Estimated budget range: Medium to high.

4.4 PPM collaboration studies

Please refer to **section 2**.

4.5 Improving decentralized and fully integrated access to TB and ART treatment

Objective: To understand factors that facilitate fully integrated service provision at the point-of-care for HIV-infected TB patients.

Design: Case studies giving detailed programme description and extracting key features of success.

Setting (study population): Settings with high HIV prevalence providing high levels of integration of TB and HIV service provision at the point-of-care (model programmes).

Methods:
Integrated TB and HIV pilot programmes in countries with high dual burden of TB and HIV.

Expected outcomes: List of factors promoting integrated TB and ART service provision at the point-of-care.

Analysis: Case study methodology.

Guidance for sample size calculation (and estimated number of participants)
A single case study may be useful in drawing out context-specific lessons for going to scale with a given model in a given country. If case studies in different countries of different models are compared, there is the possibility to draw out more generic lessons for scale up which may be more widely applicable.

Expected duration/timeline: Medium.

Suitable scale: National.

Estimated budget range: Medium.

TABLE 4: SUMMARY

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
4.1 Identifying reporting gaps					
Determine validity of reported treatment outcomes of new smear positive PTB and re-treatment PTB	Record and register reviews	Understanding gaps in recording and reporting and identifying areas for programme support	Short-term	National	Low
4.2 Carry out investigational studies to address deficiencies in first-line management of TB within NTCPs					
Understand reasons for default, poor adherence, missed doses and drug stock-outs	a) Qualitative research	Context specific understanding of reasons behind deficiencies in NTP programme	Short-term	Local or national	Low
	b) Mapping of facilities	GIS maps			
4.3 Improving management of drug-sensitive TB					
Test effects of strategies emerging from studies in 4.2 to improve first line management	a) Before and after study	Estimates of effectiveness of different strategies for improved first line management of TB	Long-term	National	Medium
	b) Pragmatic randomised (individual or cluster) controlled trial with step-wedge design				
4.4 Identifying major drivers of drug-resistant TB					
4.4.1 Individual risk factors for drug resistant TB					
Identify risk factors for development of drug resistant TB at the individual level	Case-control study	Identification of major risk factors for DR-TB	Medium to long-term	Local, national, international	Low to medium
4.4.2 Programmatic risk factors for drug resistant TB					
Identify risk factors for development of drug resistant TB at the programmatic level	Case-control study	Identification of major programmatic risk factors for DR-TB	Short-term	Local or national	Low to medium

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
4.5 Defining and evaluating strategies for integration/scale-up of DR-TB management within TB control programs					
a) Develop algorithms for selecting patients eligible for DST and 2nd line treatment in different settings	a) Refer to Section 1.2	Improved diagnostic algorithms for early detection of MDR-TB	Medium to long-term	National with international support	Medium to high
b) Develop strategies for provision of 2nd line treatment	b) Case studies	Accessible and practical community based MDR treatment arrangements			
c) Evaluate effectiveness of existing control measures and strategies	c) Refer to Section 3.6	Adequate infection control procedures			
4.6 PPM collaboration studies					
Refer to **Section 2**					
4.7 Improving decentralised and fully integrated access to TB and ART treatment (refer to Section 3 also)					
Understand factors which facilitate fully integrated service provision at POC for HIV infected TB patients	Case studies: extract key success features from integrated TB and HIV programmes	List factors promoting integrated TB and ART service provision at POC	Medium	National	Medium

5. Capacity building for operational research

5.1 What is the impact of existing training courses in terms of products and measurable outcomes?

Although there is a large literature on different approaches to educational evaluations, few interventions for work-based education of health professionals in developing countries have been adequately evaluated *(48, 49)*.

Objective: To assess the impact of existing training in terms of products/outputs and outcomes.

Design: Evaluations of training courses should be rigorous but also simple and feasible in a resource-constrained setting. There are no specific evaluation tools available for evaluating training courses in developing countries, so it is important to design evaluations using published frameworks that incorporate criteria derived from different perspectives *(50)*. Evaluations should assess whether or not course outcomes have been achieved and should include indicators of process (i.e. how the course was delivered), content (i.e. what was delivered) and outcomes (i.e. completed assignments and projects; improved competence and confidence in applying research skills into practice)*(51)*. Criteria for selecting the evaluation tools for different types of training will vary but in principle they should be selected from those published in peer-reviewed journals and be relevant to the course outcomes and for evaluating innovative educational interventions based on social learning *(48, 49)*, and they could be applied within the available time and resource constraints. Ideally two different methods should be used to triangulate the assessment of each of the learning outcomes of the training.

Setting (study population): Existing training courses in operational research offered by international agencies, NGOs or academic institutions.

Methods: Subjects should include students, tutors, institutional managers and other stakeholders including users of the course outputs. Examples of evaluation tools include: course assignments, questionnaires, surveys, self-efficacy scale which asks learners to score a series of 11 statements about their research skills from 1 (= not at all able) to 10 (= very able) (this has good internal consistency and face validity across a range of professional programmes); 'stages of change' tool to assess progress in changing learners' attitudes, intentions and actions in relation to research. Outcomes should also include operational research publications, mainly those leading to policy change at local, national or international level.

It is also suggested that an alumni association of participants trained in operational research is set up by training institution that should monitor on an annual basis whether these participants are still engaged in operational research, have completed studies, have written-up papers and are actively involved in training/monitoring others.

Suitable scale: Individual training courses.

Estimated budget range: Low to medium (course evaluation approximately 10-15% of the total course budget for a rigorous, possibly publishable, evaluation).

5.2 Models of operational health research capacity building

Objective(s): Assessment of existing models of operational research capacity building and lessons learnt.

Design: Operational health research capacity building (OHRCB) is more than the provision and evaluation of training and evaluation as has been described in **section 5.1**. It is defined as an "ability of individuals, organizations, or systems to perform and utilise health research effectively, efficiently, and sustainably" (see section on definitions in **Annex I**). A review of published and grey literature to collate information about models for OHRCB is required, or countries may wish to review publications if already in existence *(52)*.

Setting (study population): All settings in which individual and/or institutional capacity building have been described should be included.

Methods: Studies identified from: i) formal searches of health, education and management literature using pre-defined search terms; ii) websites of funders and users of health research, health consultancy organizations and academic institutions; iii) personal contacts involved in doing or evaluating health research capacity building; iv) review of institutional research repositories. The types of studies will include cases studies, realist reviews, expert opinions and theoretical frameworks for designing and/or evaluating health research capacity building.

Expected outcome: Different documented models for OHRCB which have been rigorously evaluated and shown to work in a variety of settings and which provide examples from which country-level TB programmes and partners can choose.

Analysis: Studies that describe real-life health research capacity building projects and which include a formal evaluation using a pre-designed evaluation framework/indicators will be analysed. The analysis will involve developing a framework in collaboration with researchers and research users to ensure that the outputs meet the needs of a range of players. Key indicators in this framework, as indicated in the definition, include uptake and use of research and sustainability. Sustainability is usefully framed in terms of both funding sources and retention of trained staff (see also **section 5.3**). The framework will be populated with information from each of the eligible studies independently by two individuals and common lessons will be extracted by comparing across studies.

Guidance for sample size calculation (and estimated number of participants) N/A.

Expected duration/timeline: Short-term.

Suitable scale: International.

Estimated budget range: Medium (full-time post doctoral researcher, consumables and one meeting with stakeholders).

Illustrative references: *(53-57)*.

5.3 Sustaining operational research capacity at the national level

5.3.1 Identifying funding mechanisms for operational research capacity building at national level

Objective: To identify funding mechanisms to sustain operational research capacity at national level.

Design Some international funders are increasingly investing in capacity building (e.g. the European Commission, Wellcome Trust, International Development Research Centre) and have formed a collaboration to focus on how to evaluate capacity building *(58)*. One element of operational health research capacity strengthening is securing of national 'buy-in' to the value of operational research for programme delivery, which may be expressed in allocation of national budgets.

5.3.2 What are possible ways of sustaining and retaining trained operational research staff within programmes?

Along with sustained funding, sustained staff retention is another indicator of successful OHRCB, and lessons for retaining and sustaining trained staff should be extracted from the OHRCB review described in **section 5.2**.

TABLE 5: SUMMARY

Objective(s)	Methods	Expected outcomes	Expected Duration	Suitable Scale	Estimated budget range
5.1 What is the impact of existing training courses in terms of products/outputs and outcomes?					
Assess impact of existing training	Evaluation	Determine impact of existing training	Short-term	Individual training courses	Low to medium
5.2 Models of Operational Health Research Capacity Building (OHRCB)					
Determine existing OHRCB and lessons learnt	Review of published and grey literature (including case studies, realist review, expert opinion and theorectical frameworks)	Documented relevant models for OHRCB which have been shown to work in a variety of settings	Short-term	International	Medium
5.3 Sustaining OR capacity at national level					
5.3.1 Identifying funding mechanisms for OR capacity building at national level					
Identify funding mechanisms to sustain OR capacity at national level	Search for funding opportunities, advocacy for funding for capacity building in national budgets	Identification of funding for capacity building	Short-term	International	Low
5.3.2 What are possible ways of sustaining and retaining trained operational research staff within programmes?					
Identify mechanisms for sustained staff retention	Extract lessons for retaining and sustaining trained staff from OHRCB review described in **Section 5.2**	Identification of mechanisms for sustainable retention of staff	Short-term	International	Medium

REFERENCES

1. Long NH et al. Longer delays in tuberculosis diagnosis among women in Vietnam. *International Journal of Tuberculosis and Lung Disease*, 1999; 3(5):388–93.
2. Wandwalo ER and Morkve O. Delay in tuberculosis case-finding and treatment in Mwanza, Tanzania. *International Journal of Tuberculosis and Lung Disease*, 2000; 4(2):133–8.
3. Rajeswari R et al. Socio-economic impact of tuberculosis on patients and family in India. *International Journal of Tuberculosis and Lung Disease*, 1999; 3(10):869–77.
4. Squire SB et al. 'Lost' smear-positive pulmonary tuberculosis cases: where are they and why did we lose them? *International Journal of Tuberculosis and Lung Disease*, 2005; 9(1): 25–31.
5. Kemp JR et al. Can Malawi's poor afford free tuberculosis services? Patient and household costs associated with a tuberculosis diagnosis in Lilongwe. *Bulletin of the World Health Organization*, 2007; 85(8):580–5.
6. Squire SB and Thomson R. Tuberculosis and Poverty. In: Schaaf HS and Zumla A, eds. *Tuberculosis: A Comprehensive Clinical Reference.* Oxford, Saunders–Elsevier, 2009.
7. Pawson R, Greenhalgh T, Harvey G and Walshe K. Realist review: a new method of systematic review designed for complex policy interventions. *Journal of Health Service Research Policy*, 2005; 10 Suppl 1: 21–34.
8. Steingart KR et al. Fluorescence versus conventional sputum smear microscopy for tuberculosis: a systematic review. *Lancet Infectious Diseases*, 2006; 6(9):570–81.
9. Leeman J et al. Implementation of antiretroviral therapy adherence interventions: a realist synthesis of evidence. *Journal of Advanced Nursing*, 2010; 66(9):1915–30.
10. Dieleman M, Gerretsen B and van der Wilt GJ. Human resource management interventions to improve health workers' performance in low and middle income countries: a realist review. *Health Research Policy and Systems*, 2009; 7:7.
11. Donner A. Approaches to sample size estimation in the design of clinical trials--a review. *Statistics in Medicine*, 1984; 3(3):199–214.
12. Machin D CM, Fayers PM and Pinol A. *Sample Size Tables for Clinical Studies (3rd edition)*. London, BMJ Books, 1997.
13. Moulton LH et al. Statistical design of THRio: a phased implementation clinic-randomized study of a tuberculosis preventive therapy intervention. *Clinical Trials*, 2007; 4(2):190–9.
14. Zwarenstein M et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. *BMJ*, 2008; 337:a2390.
15. Khan MS et al. Improvement of tuberculosis case detection and reduction of discrepancies between men and women by simple sputum-submission instructions: a pragmatic randomised controlled trial. *The Lancet*, 2007; 369(9577):1955–60.
16. Fairall LR et al. Effect of educational outreach to nurses on tuberculosis case detection and primary care of respiratory illness: pragmatic cluster randomised controlled trial. *BMJ*, 2005; 331(7519):750–4.
17. Elink Schuurman MW et al. The rapid village survey in tuberculosis control. *Tubercle and Lung Disease*, 1996; 77(6):549–54.
18. Demissie M et al. A rapid survey to determine the prevalence of smear-positive tuberculosis in Addis Ababa. *International Journal of Tuberculosis and Lung Disease*, 2002; 6:580–4.
19. Jittimanee SX et al. A prevalence survey for smear-positive tuberculosis in Thai prisons. *International Journal of Tuberculosis and Lung Disease*, 2007; 11(5):556–61.
20. Sekandi JN et al. Active case finding of undetected tuberculosis among chronic coughers in a slum setting in Kampala, Uganda. *International Journal of Tuberculosis and Lung Disease*, 2009; 13(4):508–13.

21. Williams B et al. The design effect and cluster samples: optimising tuberculosis prevalence surveys. *International Journal of Tuberculosis and Lung Disease*, 2008; 12(10):1110–5.

22. Datiko DG and Lindtjorn B. Health extension workers improve tuberculosis case detection and treatment success in southern Ethiopia: a community randomized trial. *PLoS ONE*, 2009; 4(5):e5443.

23. Corbett EL et al. Comparison of two active case-finding strategies for community-based diagnosis of symptomatic smear-positive tuberculosis and control of infectious tuberculosis in Harare, Zimbabwe (DETECTB): a cluster-randomised trial. *The Lancet*, 2010; 376(9748):1244–53.

24. Miller AC et al. Controlled trial of active tuberculosis case finding in a Brazilian favela. *International Journal of Tuberculosis and Lung Disease*, 2010; 14(6):720–6.

25. Churchyard et al. Twelve versus six-monthly radiological screening for the active case-finding of tuberculosis: a randomised controlled trial. *Thorax*, 2011; 66:134-139.

26. *TB Impact Measurement policy and recommendations for how to assess the epidemiological burden of TB and the impact of TB control.* Stop TB policy paper, no. 2 (WHO/HTM/TB/2009.416). Geneva, World Health Organization, 2009.

27. Baggaley RF and Fraser C. Modelling sexual transmission of HIV: testing the assumptions, validating the predictions. *Current Opinion in HIV/AIDS*, 2010; 5(4):269–76.

28. Cohen T etal. Mathematical models of the epidemiology and control of drug-resistant TB. *Expert Review of Respiratory Medicine*, 2009; 3(1):67–79.

29. *Case studies*. Redditch, Lanner, 2008–2011. (http://www.lanner.com/en/case-studies.cfm, accessed 7 January 2011).

30. Floyd K et al. Cost and cost-effectiveness of PPM-DOTS for tuberculosis control: evidence from India. *Bulletin of the World Health Organization*, 2006; 84(6): 437-45.

31. Kelkar-Khambete A et al. India's Revised National Tuberculosis Control Programme: looking beyond detection and cure. *International Journal of Tuberculosis and Lung Disease*, 2008; 12(1): 87-92.

32. Thorpe KE et al. A pragmatic-explanatory continuum indicator summary (PRECIS): a tool to help trial designers. *Journal of Clinical Epidemiology*, 2009; 62(5): 464-75.

33. Malmborg R, Mann G, Thomson R an d Squire SB. Can public-private collaboration promote tuberculosis case detection among the poor and vulnerable? *Bulletin of the World Health Organization*, 2006; 84(9): 752-8.

34. Theobald S et al. Towards building equitable health systems in Sub-Saharan Africa: lessons from case studies on operational research. *Health Research Policy and Systems*, 2009; 7: 26.

35. Hayes RJ and Bennett S. Simple sample size calculation for cluster-randomized trials. *International Journal of Epidemiology,* 1999; 28(2):319–26.

36. Munro SA et al. Patient adherence to tuberculosis treatment: a systematic review of qualitative research. *PLoS Medicine*, 2007; 4(7): e238.

37. World Health Organization. *WHO policy on TB infection control in health-care facilities, congregate settings and households* (WHO/HTM/TB/2009.419; 2009). Geneva, World Health Organization, 2009.

38. Oxlade O et al. Developing a Tuberculosis Transmission Model That Accounts for Changes in Population Health. *Medical Decision Making*, 2010; 31:53-68.

39. De Valliere S, Barker RD. Poor performance status is associated with early death in patients with pulmonary tuberculosis. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, 2006 Jul;100(7):681–6.

40. *Survival/Failure Time Analysis (software)*. Tulsa, Statsoft Inc., 2010. (http://www.statsoft.com/textbook/survival-failure-time-analysis, accessed 7 January 2011).

41. Vaeth M and Skovlund E. A simple approach to power and sample size calculations in logistic regression and Cox regression models. *Statistics in Medicine*, 2004; 23(11): 1781-92.

42. Waitt CJ, Squire SB. A systematic review of risk factors for death in adults during and after tuberculosis treatment. *International Journal of Tuberculosis and Lung Disease*, 2011; 15(7):871-885 dio:10.5588/ijtld/10.0352.

43. Lawn SD et al. Early mortality among adults accessing antiretroviral treatment programmes in sub-Saharan Africa. *AIDS*, 2008; 22(15): 1897-908.

44. Braitstein P et al. Mortality of HIV-1-infected patients in the first year of antiretroviral therapy: comparison between low-income and high-income countries. *The Lancet*, 2006; 367(9513): 817-24.

45. Tuboi SH et al. Mortality during the first year of potent antiretroviral therapy in HIV-1-infected patients in 7 sites throughout Latin America and the Caribbean. *Journal of Acquired Immune Deficiency Syndrome*, 2009; 51(5): 615-23.

46. Thiam S, et al. Improving adherence to tuberculosis treatment in a resource-poor setting: A Cluster Randomised Controlled Trial. *JAMA*, 2007; 297(4):380-6.

47. Balaji V et al. Risk factors for MDR and XDR-TB in a tertiary referral hospital in India. *PLoS One*, 2010; 5(3): e9527.

48. Cantillon P. Evaluation: beyond the rhetoric. *Journal of Evaluation in Clinical Practice*, 1999; 5(3): 265-8.

49. Prideaux D. Researching the outcomes of educational interventions: a matter of design. RCTs have important limitations in evaluating educational interventions. *British Medical Journal*, 2002; 324(7330): 126-7.

50. Eraut M. Handling values issues. In: *The politics and ethics of evaluation*. London, Croom Helm, 1984.

51. Cervero RM. Continuing professional education and behavioral change: a model for research and evaluation. *Journal of Continuing Education in Nursing*, 1985; 16(3): 85-8.

52. Kakuma et al. *Evaluating Capacity Development in Global Health Research – is there evidence?* Abstract. First Global Symposium on Health Systems Research, Montreux, Switzerland, 2010.

53. *Accra Agenda for Action* (2008). 3rd High Level Forum on Aid Effectiveness. Accra, Ghana.

54. Cooke J. A framework to evaluate research capacity building in health care. *BMC Family Practice*, 2005; 6: 44.

55. Neilson S and Lusthaus C. *IDRC-Supported Capacity Building: Developing a Framework for Capturing Capacity Changes.* Ottawa, Universalia, 2007. (http://www.idrc.ca/uploads/user-S/11762347991CB-Developing_Framework_Capturing_Capacity_Changes_FINAL.pdf, accessed 7 January 2011).

56. Capacity Building In Research: 'How To' Note. London, Department for International Development, 2010.

57. *WHO's role and responsibilities in health research: Bamako Global Ministerial Forum on Research for Health* (Background paper to WHO, Executive Board, 124th Session). Geneva, World Health Organization, 2009.

58. *Funders meet in new initiative on research capacity strengthening: ESSENCE is launched*. Geneva, Special Programme for Research and Training in Tropical Diseases, 2008. (http://apps.who.int/tdr/svc/news-events/news/essence, accessed 7 January 2011).

Annex III

The Impact Assessment Framework for new diagnostics

The Impact Assessment Framework (IAF) *(1)* has been developed by a multidisciplinary team at the Liverpool School of Tropical Medicine and collaborators, including clinicians, laboratory specialists, health economists, social scientists and health systems analysts. It is based on a range of prior research activities in different countries that supported various elements of the evidence base *(2–14)*. These elements have been combined to provide an overarching framework (the IAF) to indicate how sufficient information for policy decisions could be collected in a systematic manner for all new diagnostic tools and approaches. The sufficiency of information has been considered in line with the international targets of the Global plan to Stop TB and the Millennium Development Goals (MDGs) *(15)*. The IAF, with references relating to different types of evidence, is shown in **Table 3**.

The IAF comprises five interconnected layers:

1. Layer 1: Effectiveness analysis
2. Layer 2: Equity analysis
3. Layer 3: Health systems analysis
4. Layer 4: Scale-up analysis
5. Layer 5: Policy analysis.

Layer 1: Effectiveness analysis
This layer requires evidence about the accuracy (sensitivity and specificity) of new tools and approaches, but also flags the need to go further than this, and build evidence on effectiveness. Data on sensitivity and specificity are universally provided by developers of new diagnostics and their positive and negative predictive values have been suggested by GRADE as proxies for patient important outcomes in the assessments of new tools. However, estimations of the number of patients who might start and complete 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 (accuracy), which may not always translate into appropriate clinical or public health management decisions for patients within the context of health services (effectiveness).

Layer 2: Equity analysis
This layer examines who benefits from the new intervention. The Global plan to Stop TB highlights the need to "prioritize the needs of the poor and vulnerable" recognizing that the greatest burden of TB is found among poor people and they face the greatest barriers in access to care *(8)*. Typically, however, the systematic measurement of equity in health and health interventions is either absent or sporadic. Although the first MDG is expressed in terms of an equitable outcome, the health and other goals that are intended to contribute to this make no specific reference to equity or distributional issues *(16)*.

Layer 3: Health systems analysis
This layer examines the health systems requirements of a new intervention, for example human resources, infrastructure, operating procedures, quality assurance, procurement and maintenance.

These data are sometimes collected during the demonstration studies (see **Figure 1**, in **section I**) – in optimized operational settings - of new diagnostics, but not in all cases. Even where they are collected, the improvement to operations necessarily provided through the demonstration study may mask issues that become apparent in implementation studies. This layer provides crucial data for assessing the feasibility of implementation and for identifying where key constraints, or bottlenecks, in the system may occur.

Layer 4: Scale-up analysis
This layer projects and models the full economic costs as well as the clinical and epidemiological effects of going from demonstration or implementation studies to full scale (national or regional) with a new

intervention. Health system, patient, and societal perspectives are all important here. Modelling techniques can provide information concerning the epidemiological benefits of scaling up and, when combined with patient costs from Layer 2, total additional costs or savings for patients. At the same time mathematical systems analysis techniques can outline the potential constraints to and resources required for scale-up. When combined with cost analyses from Layer 3, these can give an indication of total resources required as well as identify and quantify likely resource gaps.

Layer 5: Policy analysis

This layer critically appraises the new intervention studied in layers 1-4 against other interventions that are available or may become available for uptake in the short- to medium-term. An important part of this layer is a scoping of the risk that a given new diagnostic test may be supplanted by newer technology within a short period of time. It requires a rapid assessment of data on other pipeline diagnostics from the previous four layers and a review of whether changes made for one diagnostic may provide a better platform for the next technology, or alternatively whether the new technology is 'disruptive' *(17)*, or 'market transformational' *(18)*, both terms are used to describe a technology that could radically alter the way in which TB diagnosis is achieved.

Using the IAF

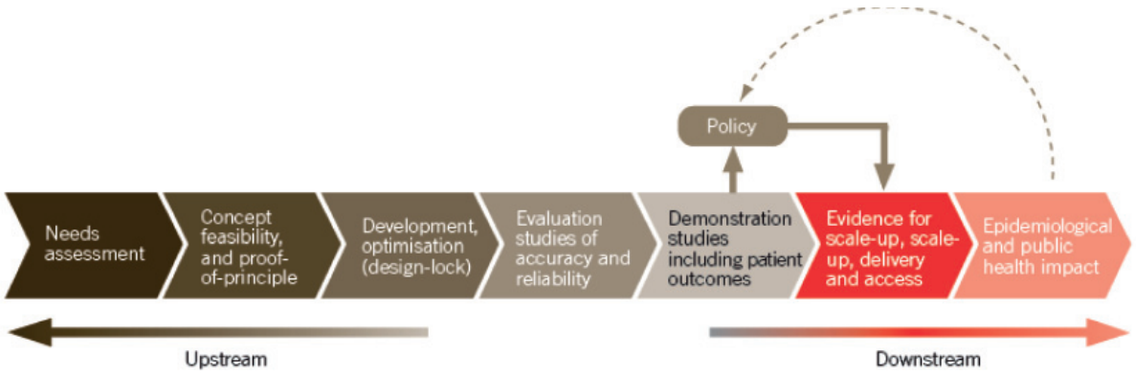
The IAF can be used by diagnostics research teams during the 'demonstration' and 'evidence for scale-up, delivery and access' phases of development shown in **Figure 8**. The latter may take the form of field evaluation, or implementation studies in non-optimized settings, or other operational research activities. The framework can also be used by international policy-makers during the policy development process to systematically assess a broader range of evidence and by national policy-makers to support adoption, implementation and scale-up decisions.

The IAF has already been used for the development of protocols for a multi-country research programme to study the implementation of line probe assays (LPA), which were recommended by WHO STAG-TB in 2008 *(19)*. Clinicians and laboratory specialists from three countries (Russia, Brazil and South Africa), along with other members of the core group of the TREAT TB Diagnostic Tools Initiative,[ee] discussed the priority research questions they would like to answer with regard to the use of LPAs and mapped these questions to each layer of the IAF. All the questions that were raised mapped to one layer of the framework and all layers were addressed; the resulting framework is shown in **Table 3**. Each of these teams now has a different protocol for collecting the evidence, due to the stage at which their NTCPs are with regard to rolling out LPA. Nevertheless, each will provide data against the same set of outcome indicators, facilitating comparisons across different epidemiological settings.

The recommended methodology to feed robust evidence into Layers 1 to 3 is prospective randomised, controlled trials (RCT). This design permits comparisons between the existing technology and approach (control) and the new (intervention) as follows:

For Layer 1: A comparison of effects on a) numbers of patients achieving important outcomes (including

FIGURE 8: DIAGNOSTICS DEVELOPMENT PATHWAY



Source: Stop TB Partnership's New Diagnostic's Working Group. *Pathways to better diagnostics for tuberculosis: a blueprint for the development of TB diagnostics (2009)*. Geneva, World Health Organization, 2010.

[ee] See: http://www.treattb.org

diagnosis, start of treatment, and treatment completion); and b) time to achieving these outcomes.

For Layer 2: A comparison of effects on different patient sub-groups (e.g poor compared with less poor, adults compared with children). Equity may be assessed based on outcome indicators among different groups, in terms of morbidity or mortality measures, or process indicators such as health service use *(16)*. Analysis of socioeconomic status may use asset-based measures to define different socioeconomic groups *(20)*. Demographic and health surveys, and more recent TB prevalence surveys, are increasingly using these methods *21, 22)*.

For Layer 3: A comparison of the health system inputs required. Data for this may be gained through economic analyses of standard versus new diagnostic interventions, focussing on the health system not just the tool, and through interviews with health systems personnel.

Data for these comparisons can be obtained across all study participants in both intervention and control arms, or through nested sub-studies on more limited numbers. For example, in-depth qualitative and quantitative research on patient costs incurred during a diagnostic process (either control or intervention) is time consuming and thus only collected for a sub-group of study participants. Data from Layers 1 to 3 can then be fed into the modelling and other methodologies required in Layers 4 and 5.

The type of randomized trial employed will depend on the stage of diagnostic development to which the IAF is being applied. During demonstration studies (which may be prior to STAG-TB approval) an explanatory RCT with well-controlled study conditions and data collection instruments is appropriate. During subsequent implementation or operational research a pragmatic RCT approach using existing health system data will be more suitable (for a fuller description of the difference between explanatory and pragmatic RCTs, see reference *23*). There are concerns that RCT designs deny some patients (those in the control arm) the assumed benefits of a new technology – especially in implementation research of STAG-approved technologies. Such ethical concerns need to be addressed, for example by ensuring that the PRCT includes a scale-up plan, for example through a stepped-wedge approach in which all sites access the technology but in a phased manner to allow for comparisons between those with and without the technology.

TABLE 3: A FRAMEWORK FOR IMPACT ASSESSMENT FOR NEW DIAGNOSTICS.

Layer of Assessment	Kinds of question(s) being addressed	References to work addressing these questions
Layer 1: Effectiveness analysis	How well does new tool work in terms of accuracy? How many additional cases will be identified who would otherwise not have been identified? How many additional cases will actually (start and complete) treatment as a result of using new tool?	*2* *6* *7*
Layer 2: Equity analysis	Who benefits from new tool? (ambulant or hospitalized, poor/less poor, men/women, adults/children) Why do these benefits accrue? (health system level in which new diagnostic is deployed, change time to issue of results, change in patient costs)	*8* *13*
Layer 3: Health system analysis	What are the human resource implications of introducing the new tool? (training, number and cadre of staff) What are the infrastructure implications? (equipment, laboratory layout, safety installations) What are the procurement implications? (reagents, consumables, documentation) What are the implications for quality assurance? (internal and external)	*3* *11* *14*
Layer 4: Scale-up analysis	What are the projected impacts of going to scale with new tool? e.g. a) cost savings to patients in relation to income b) cost savings to health providers/the health system d) Effects on transmission of improved infection control as a result of new tool	*4* *11*
Layer 5: Policy analysis	What other similar technologies are available or likely to become available? How do similar existing or emerging technologies compare in their projected performance within each of the layers above?	*11*

REFERENCES

1. Mann G et al. Beyond accuracy: creating a comprehensive evidence base for TB diagnostic tools. International Journal of Tuberculosis and Lung Disease, 2010, 14(12):1–7.
2. Wilkinson D et al. Trial-of-antibiotic algorithm for the diagnosis of tuberculosis in a district hospital in a developing country with high HIV prevalence. International Journal of Tuberculosis and Lung Disease, 2000, 4:513-8.
3. Mundy CJF et al. Quality assessment of sputum transportation, smear preparation and AFB microscopy in a rural district in Malawi. International Journal of Tuberculosis and Lung Disease, 2002, 6:47-54.
4. Mundy CJF et al. The operation, quality and costs of a district hospital laboratory service in Malawi. Transactions of the Royal Society of Tropical Medicine and Hygiene, 2003, 97:403-8.
5. Squire SB et al. “Lost” smear positive pulmonary tuberculosis cases; where are they and why did we lose them? International Journal of Tuberculosis and Lung Disease, 2005, 9:25-31.
6. Ramsay A et al. The bleach microscopy method and case detection for TB control. International Journal of Tuberculosis and Lung Disease, 2006, 10:256-8.
7. Cambanis A et al. Rural poverty and delayed presentation to tuberculosis services in Ethiopia. Tropical Medicine & International Health, 2005, 10:330-335.
8. Kemp JR et al. Can Malawi’s poor afford free TB services? Patient and household costs associated with a TB diagnosis in Lilongwe. Bulletin of the World Health Organisation, 2007, 85:580-585.
9. Liu X et al. How affordable are TB diagnosis and treatment in rural China? An analysis from community and TB patient perspectives. Tropical Medicine and International Health, 2007, 12:1464-1471.
10. Theobald S et al. Towards building equitable health systems in Sub-Saharan Africa: What can we learn from case studies on operational research? Health Research Policy and Systems, 2009, 7:26 doi:10.1186/1478-4505-7-26.
11. Ramsay A et al. New Policies, New Technologies: Modelling the Potential for Improved Smear Microscopy Services in Malawi. Public Library of Science ONE, 2009, 4(11): e7760. doi:10.1371/journal.pone.0007760.
12. Yagui M et al. Timely diagnosis of MDR-TB under program conditions: is rapid drug susceptibility testing sufficient? International Journal of Tuberculosis and Lung Disease, 2006, 10:838–843.
13. Ramsay A et al. Sputum sex and scanty smears: new case definition may reduce sex disparities in smear-positive tuberculosis. International Journal of Tuberculosis and Lung Disease, 2009, 13:613–619.
14. Harries AD et al. Using a bus service for transporting sputum specimens to the Central Reference Laboratory: effect on the routine TB culture service in Malawi. International Journal of Tuberculosis and Lung Disease, 2004, 8:204–210.
15. World Health Organization, The Stop TB Partnership. The Stop TB strategy: building on and enhancing DOTS to meet the TB-related Millennium Development Goals. World Health Organisation, 2006.
16. Gwatkin D R. Health inequalities and the health of the poor: what do we know? What can we do? Bulletin of the World Health Organization, 2000, 78:3–1.
17. Christensen CM. The innovator’s dilemma: when new technologies cause great firms to fail. Boston, MA, USA: Harvard Business School Press, 1997.
18. Blumstein C, Goldstone S, Lutzenhiser L. A theory-based approach to market transformation. Energy Policy, 2000, 28:137–144.
19. World Health Organization. New WHO policies. World Health Organisation, 2010. http://www.who.int/tb/dots/laboratory/ policy/en/index.html Accessed September 2010.

20. Wagstaff A. Socio-economic inequalities in child mortality: comparisons across developing countries. Bulletin of the World Health Organization, 2000, 78:19–29.

21. Lönnroth K et al. Inclusion of information on risk factors, socio-economic status and health seeking in a tuberculosis prevalence survey. International Journal of Tuberculosis and Lung Disease, 2009, 13:171–176.

22. Muniyandi M, Rajeswari R. Socio-economic inequalities of tuberculosis in India. Expert Opin Pharmacother, 2008, 9:1623–1628.

23. Zwarenstein M et al. Improving the reporting of pragmatic trials: an extension of the CONSORT statement. BMJ, 2008; 337:a2390.

Annex IV
Tool for capturing patient costs

The Tool to Estimate Patients' Costs can help to assess the impact of TB on the welfare of households and individuals. More specifically, it helps to estimate the costs that TB patients incur before and during diagnosis and during treatment. It also helps to gather information on health-seeking behaviour patterns, individual and household income, coping behaviour, socioeconomic situation, gender-related issues and the social impact of TB.

The tool consists of a generic questionnaire designed to be adapted to local settings, a literature review, as well as guidelines on adaptation of the tool, interpretation of findings, socioeconomic indicators, methods and sampling and possible interventions. In addition, an Epi-Info data entry template, an Excel-based summary sheet and an example Powerpoint showing how to present results are also included.

The tool can be downloaded as a pdf file. It is available at: http://www.stoptb.org/wg/dots_expansion/tbandpoverty/spotlight.asp.

Please note that the data entry template is not provided in MS Access format, but in Epi Info format. Epi Info can be downloaded for free on the CDC Atlanta website.[ff]

[ff] See: http://wwwn.cdc.gov/epiinfo/

Annex V

Resources for developing and evaluating active case-finding approaches

Typically, about half the burden of undiagnosed culture-positive TB will be found in the 5 to 10% of individuals who report chronic cough in any given community (this does not vary much by HIV status) *(1–4)*. Many options exist for delivering active case-finding (ACF) services (see **Figure 9** and **Table 4**). On this basis, different models should be developed and tested in various situations (urban and rural settings, high and low burdens of TB, presence or absence of strong community health worker programmes, etc.), as well as in important risk groups, such as PLHIV, contacts of TB cases, persons with diabetes and drug users. For instance, options include interventions delivered by multipurpose community workers (both successful and unpromising results have been obtained; see **Table 4**) and dedicated TB teams. In order to identify the best ACF delivery model, these would have to be evaluated and compared with regards to effectiveness (number needed to screen and cases diagnosed) and costs of the different approaches in deciding (i) *whom* to screen for undiagnosed TB in the general community (whole populations, all adults, adults who are symptomatic on direct questioning, adults who respond to the offer to report their symptoms, or adults known to have risk factors for TB); and (ii) *how* to screen them (different symptoms, smear microscopy, chest radiography, newer TB diagnostics). In addition, the effectiveness and costs of different approaches to delivering community-based services should be assessed.

FIGURE 9: VARYING LEVELS OF EFFECTIVENESS AND SUSTAINABILITY OF VARIOUS ACF SERVICES ACCORDING TO THE CHARACTERISTICS OF THE TARGETED GROUPS

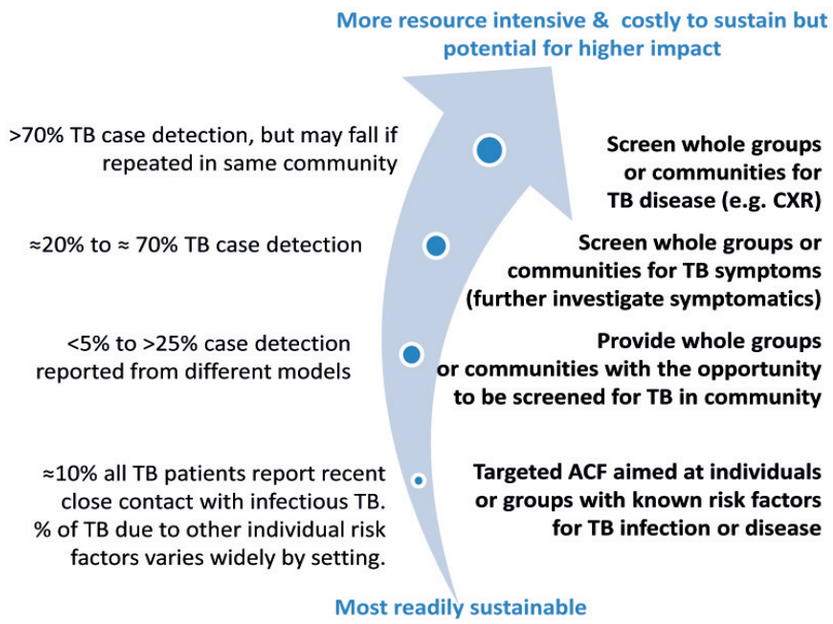


TABLE 4: BROAD STRATEGIES AND REPRESENTATIVE EXAMPLES OF DIFFERENT APPROACHES TO PROVIDING ACF (SEE ALSO **FIGURE 9**)

Broad Strategy	Representative examples	
Screen whole groups or communities for TB disease	Most prevalence surveys *(5)* Mobile CXR in USA and Europe during last century *(6)* Annual CXR in S. African miners and in Korean civil servants *(7–10)*.	Most commonly screening all adults with CXR, with or without a symptom screen. One-off provides estimate of prevalence. If repeated provides trends *(8,21)* or ACF if repeated frequently *(6)*. Residual burden of radiologically undetectable TB may limit impact of repeat CXR screening *(22)*.
Screen whole groups or communities for TB symptoms (further investigate ymptomatic)	"Rapid surveys" for TB in urban/rural pops & prisons *(11–15)*. Individual interview with all household members *(11–14)* or door-to-door enquiry for chronic cough in the household *(7)*.	CXR or smear used to investigate people reporting symptoms. Provides low cost estimate of prevalence (one-off implementation, symptomatic disease only) *(11–15)*, and if repeated periodically provides ACF *(7)*. More resource intensive than providing opportunity to report symptoms to outreach services and, in one study, less effective (≈20% case-detection per round *(7)*).
Provide whole groups or communities with the opportunity to be screened for TB in community	Outreach or community services delivered to all wanting screen, or symptomatics only, delivered through - Periodic outreach clinic/van *(7,16)* - CHWs or lay volunteers providing continuous services *(17–19)* - community-directed *(20)*.	Relies on community mobilization/ dissemination, with community-level services providing sputum collection in the community. High impact on case-detection in some *(7,17,19,20)* but not all *(16,18)* examples. High impact on undiagnosed prevalent TB in one urban population *(7)*.
Targeted ACF aimed at individuals or groups with known risk factors for TB infection or disease	Screening in those with known TB risk factors (e.g. HIV/diabetes /IVDU / immigrants) and Contact tracing (typically screen for TB infection & disease)	Can be linked to IPT. Screening for TB linked to risk factors will be most effective in communities where there is a high population attributable fraction, good diagnostic and chronic care services. Reaching the very poor will require community-level services to identify "at risk" individuals and provide screening

REFERENCES

1. Tupasi TE et al. Bacillary disease and health seeking behavior among Filipinos with symptoms of tuberculosis: implications for control. *International Journal of Tuberculosis and Lung Disease,* 2000;4(12):1126-32.
2. Ayles H et al. Prevalence of tuberculosis, HIV and respiratory symptoms in two Zambian communities: implications for tuberculosis control in the era of HIV. *PLoS ONE,* 2009;4(5):e5602.
3. Corbett EL et al. Provider-initiated symptom screening for tuberculosis: diagnostic value and the impact of HIV. *Bulletin of the World Health Organization,* 2010; 88(1):13-21.
4. National Tuberculosis Institute B. Tuberculosis in a rural population of South India: a five-year epidemiological study. *Bulletin of the World Health Organization,* 1974;51(5):473-88.
5. Glaziou P et al. Tuberculosis prevalence surveys: rationale and cost. *International Journal of Tuberculosis and Lung Disease,* 2008;12(9):1003-8.
6. Golub JE, Mohan C, Comstock GW, Chaisson RE. Active case-finding of tuberculosis: historical perspective and future prospects. *International Journal of Tuberculosis and Lung Disease,* 2005;9:1183-203.
7. Corbett EL et al. Comparison of two active case-finding strategies for community-based diagnosis of symptomatic smear-positive tuberculosis and control of infectious tuberculosis in Harare, Zimbabwe (DETECTB): a cluster-randomised trial. *Lancet*, 2010;376:1244-53.
8. Bai GH, Kim SJ, Lee EK, Lew WJ. Incidence of pulmonary tuberculosis in Korean civil servants: second study, 1992-1994. *International Journal of Tuberculosis and Lung Disease,* 2001;5(4):346-53.
9. Churchyard GJ, Kleinschmidt I, Corbett EL, Mulder D, De Cock KM. Mycobacterial disease in South African gold miners in the era of HIV infection. *International Journal of Tuberculosis and Lung Disease,* 1999;3:791-8.
10. Churchyard GJ et al. 12 versus 6-monthly radiological screening for the active case-finding of tuberculosis: a randomised controlled trial. *Thorax,* 2011;66(2):134-9.
11. Sekandi JN, Neuhauser D, Smyth K, Whalen CC. Active case finding of undetected tuberculosis among chronic coughers in a slum setting in Kampala, Uganda. *International Journal of Tuberculosis and Lung Disease,* 2009;13(4):508-13.
12. Schuurman EMW et al. The rapid village survey in tuberculosis control. *Tuberculosis and Lung Disease,* 1996;77(6):549-54.
13. Jittimanee SX, Ngamtrairai N, White MC, Jittimanee S. A prevalence survey for smear-positive tuberculosis in Thai prisons. *International Journal of Tuberculosis and Lung Disease,* 2007;11:556-61.
14. Demissie M, Zenebere B, Berhane Y, Lindtjorn B. A rapid survey to determine the prevalence of smear-positive tuberculosis in Addis Ababa. *International Journal of Tuberculosis and Lung Disease,* 2002;6:580-4.
15. Shargie EB, Yassin MA, Lindtjorn B. Prevalence of smear-positive pulmonary tuberculosis in a rural district of Ethiopia. *International Journal of Tuberculosis and Lung Disease*, 2006;10:87-92.
16. Harper I, Fryatt R, White A. Tuberculosis case finding in remote mountainous areas--are microscopy camps of any value? Experience from Nepal. *Tuberculosis and Lung Disease,*1996;77(4):384-8.
17. Datiko DG, Lindtjorn B. Health extension workers improve tuberculosis case detection and treatment success in southern Ethiopia: a community randomized trial. *PLoS ONE*, 2009;4(5):e5443.
18. Shargie EB, Morkve O, Lindtjorn B. Tuberculosis case-finding through a village outreach programme in a rural setting in southern Ethiopia: community randomized trial. *Bulletin of the World Health Organization,* 2006;84:11-119.

19. Miller AC et al. Controlled trial of active tuberculosis case finding in a Brazilian favela. *International Journal of Tuberculosis and Lung Disease,* 2010;14(6):720-6.

20. Getahun H, Maher D. Contribution of 'TB clubs' to tuberculosis control in a rural district in Ethiopia. *International Journal of Tuberculosis and Lung Disease,* 2000;4(2):174-8.

21. Hong YP, Kim SJ, Lew WJ, Lee EK, Han YC. The seventh nation-wide tuberculosis prevalence survey in Korea, 1995. *International Journal of Tuberculosis and Lung Disease,* 1998;2:27-36.

22. Lewis JJ et al. HIV Infection does not Affect Active Case Finding of Tuberculosis in South African Gold Miners. *American Journal of Respiratory & Critical Care Medicine,* 2009;180:1271-8.

Annex VI

Capacity building for implementing operational research in a national tuberculosis control programme: where, what and how ?

No sustainable operational research can be conducted within NTCPs unless suitable capacity is being established. From an NTCP perspective, this section focuses on: a) where exactly this operational research capacity should be located? b) what sort of capacity is needed? and c) how capacity building should be done? It will be illustrated through various case studies, based on experiences in capacity development and implementation of TB operational research in Malawi, Brazil, and Indonesia.

1. Where should operational research capacity be located?

Should this be within the NTCP, a government research institute, universities (local or international) or nongovernmental organization working with the NTCP? Each of these options creates potential opportunities and limitations, which are discussed below.

1.1 Embedding operational research within the NTCP structure will promote research as an important health service delivery activity and could lead to improvements both in the monitoring of programme activities and the development and use of a suitable information system. Importantly, it could contribute to empowerment of programme managers to translate research findings into policy and practice. However, spearheading operational research within the NTCP often faces a number of challenges *(1,2)*. First, there is a lack of dedicated time to conduct operational research and often this is coupled with limited capacity. Second, programme managers might not recognize the relevance of operational research because they are concentrated on programme implementation and research questions might not be perceived as being of priority, or there might be no mechanism(s) for deciding on research priorities at the national level. Thirdly, individuals who study abroad and obtain Masters or other higher degrees are usually appointed to senior-level management posts and they often have no time or opportunity for research. These challenges are often worsened by lack of office infrastructure and implementation support.

There are, however, a number of enabling factors that could help overcome these problems. First, and foremost, a coordination mechanism should be established at national level to provide a clear strategy for setting-up (and revising) research priorities, which ensures that research questions are relevant to programme implementation and connected to service delivery. It is necessary to have a full-time competent research officer with suitable programmatic skills who can support the programme manager, and develop strategies to

motivate programme staff (e.g. on-the-job training and supervision, research workshops). All this must be coupled with the provision of suitable infrastructure to support research, such as office space, internet access, data management and analysis capacities, stationary and transport. **Box 3** gives an example of embedding research capacity and implementation in the NTCP in Malawi, and outlines the enabling factors that contributed to developing TB operational research in that country.

BOX 3: MALAWI CASE STUDY

Building capacity and implementation of TB operational research in Malawi

Technical assistance and dedicated funding for operational research

In 1996, the United Kingdom Department for International Development (DFID) began providing technical and financial support for the Malawi National Tuberculosis Control Programme (NTCP) to develop, implement, write up and disseminate operational research. Between 1996 and 2003, in addition to existing support provided by The Union, other development agencies and international organizations entered into partnership with DFID to support the NTCP implementation of operational research. These included the Norwegian Agency for Technical Cooperation and Development (NORAD), the Royal Dutch TB Association (KNCV), WHO and the United States Agency for International Development (USAID). Resources were allocated to training that included in-service training during supervision, an annual research training workshop, an annual writing skills workshop and an annual review meeting to disseminate research findings to national and international stakeholders.

Multistakeholder partnership

A multistakeholder partnership was set up whereby research ideas from within the NTCP, from local institutions (e.g. the Malawi Medical School, the National AIDS Programme) and from international organizations and NGOs (e.g. WHO, MSF and The Union) and academia (e.g. the Liverpool School of Tropical Medicine) were discussed and endorsed at the six-weekly meetings of the Malawi TB Programme Management Group. After priorities were established, research activities were implemented by the various stakeholders. A good relationship was established with the Malawi National Health Science Research Committee that received and approved the annual research plan and programme before the start of the forthcoming year, and in turn expected an end-of-year report and copies of any published papers. At the end of every year, a report was written on research undertaken, studies completed, studies published and the effect that these studies had on influencing policy and practice (*2*–*9*).

The appointment of a Central Unit Officer, who worked closely with the Director of the NTCP, and was responsible for leadership, organization and implementation of operational research, was an important enabling factor in the development of an effective partnership and TB operational research programme in Malawi (*1*).

Guiding principles to build a sustainable well-functioning TB programme with country-wide, standardized case-finding, treatment and monitoring systems

From the outset, the NTCP developed guiding principles that underpinned the integration and implementation of research. The top priority was to have a well-run programme that incorporated all the essential elements of the DOTS strategy, and in particular to ensure that registers (particularly the sputum laboratory register and the district/hospital TB register) and TB treatment cards were well maintained. It was anticipated that these registers and treatment cards would form an important part of the data collection for operational research questions.

Assessment and prioritization of research studies targeted at constraints to national and local setting TB control objectives, and integral to established TB systems

With the NTCP developing initial 3-year, and ultimately 5-year plans, the goal and purpose of national

TB control were served by a number of key objectives. Constraints were identified that prevented these objectives being met, and research questions were asked to help address these constraints, by either clarifying or finding solutions to them. Research questions tended to be based around three main themes: i) Is there a lack of knowledge about the issue in question? For example, what is the prevalence of tuberculosis among prisoners?; ii) Is there a lack of a suitable tool or can a better tool be used? For example, will a package of HIV testing, counselling and co-trimoxazole preventive therapy for TB patients reduce case fatality rates?; iii) Are the tools used inefficiently or are the tools ineffective? For example, is it more cost effective for the health services to screen pulmonary tuberculosis suspects using two rather than three sputum smears?

Thinking ahead

An annual research programme detailing research activities planned within the NTCP was included in the annual workplan and approved each year by the TB Programme Steering Group. A large number of studies were undertaken, completed and published between 1996 and 2003. The success of the operational research was judged in various ways: 1) Whether proposed annual targets in terms of projects initiated, projects completed, papers written and papers published were met; 2) Whether the research findings influenced policy and practice; and 3) Whether the research helped to improve programme performance.

Strong emphasis on dissemination, both nationally and internationally

Once research studies were completed, they were quickly translated into reports and papers, many of which were subsequently published in international peer-reviewed journals. It was felt that this was the best way to disseminate knowledge within Malawi and to a global audience, and it was felt that such publication would enable TB and TB-HIV operational research to obtain scientific credibility and respect. In this seven-year period, the NTCP wrote over 100 original research articles and about 30 review/opinion and policy papers. Research publications from the Malawi TB Control Programme were collated each year into an annual report that was printed and disseminated to all districts in the country.

A focus on results, translation into policy and practice

The operational research studies conducted led to key changes in national policy and practice. These included the creation of a prison tuberculosis control programme, which continues to this day (*10,11*); improved recording and reporting of patients with previously treated TB (*12*); a change of treatment regimens from hospital-based, 2-months intensive phase therapy centred around daily injections of streptomycin – to oral, ambulatory therapy given from health facilities or from family-based guardians (*13*); better management of patients who transfer between facilities (*14*); and routine HIV testing and counselling for all TB patients, and provision of co-trimoxazole preventive therapy to those found to be HIV-positive (*15–18*). At the international level, this research contributed to the development or modification of guidelines on health-care worker safety, prison TB control, decentralized TB treatment from hospitals to health centres and beyond, management of HIV-associated TB and the eventual adoption by WHO of the recommendation to base diagnosis of TB on two rather than three sputum smears.

Challenges and lessons learnt

Despite the achievements, not all operational research was successful. There were too few well-trained research officers with too much to do, and attempts to increase these numbers were only partially successful. Several projects started and implemented with NTCP funding were never completed because of poor study design, poor or unreliable data collection and poor supervision, and there was occasionally a failure to translate a few completed but complex projects into understandable and readable papers. Sometimes the research was completed and published showing that an intervention was feasible and beneficial to the programme (*19*), yet for various reasons, policy and practice remained unchanged. A key lesson that emerged was that it is essential to put learning systems in place to capture emerging lessons from what works, where, why and how.

1.2 Locating operational research capacity at a government research institute. This is the case, for instance in Kenya, for example, with the Kenya Medical Research Institute (KEMRI) taking a lead role in operational research. The advantages are that it acts as the first point of reference for the Ministry of Health for public health research, serving other key national programmes and not just TB. Thus, it is well placed to stimulate government ownership and responsibility for research at a wider level; and it could also provide career opportunities for trained national researchers. However, there is at present limited practical experience on how to make this work in terms of functional links with the TB programme, and it relies on the need to obtain additional funding. In situations where such a structure does not already exist, it might be wiser to build a structure within the NTCP initially, with the long-term goal of integrating this into a cross-programme research centre.

1.3 Universities provide potential operational research opportunities. They have skilled resources that can be particularly useful for specific research (e.g. qualitative, economics and social sciences), they have the methodology and publication skills and their involvement can stimulate a scientific culture within the TB programme. There is a risk, however, that outsourcing research to academic institutions may draw away researchers from programmes, thereby handicapping any existing capacity. Also, as implementation is not in the academic mandate, findings may be handed to busy programme managers who do not have a sense of ownership. As a result, there may be limited or no impact in terms of getting policy and practice outputs. Many academic researchers also lack practical programme skills, which may make research rather academic and distanced from programme realities. However, examples exist of a systematic process of linking research institutes to the NTCP, and strengthening research capacity outside the NTCP, such as in Indonesia - see **Box 4**

BOX 4: INDONESIA CASE STUDY

Building operational research capacity in Indonesia: linking research institutes to NTCP and strengthening research capacity outside of the NTCP

Multistakeholder partnership

The TB Operational Research Group (TORG) was established in 2003 with membership including researchers from leading universities in Indonesia, the National Centre for Health Research and Development (NCHRD), the NTCP and key funding agencies. Terms of reference for this group included: i) Development of a plan for operational research in Indonesia; ii) Promotion of information exchange between researchers and the TB programme; iii) Support to operational research at the national, provincial and district levels; iv) Promote research co-ordination at National level; v) Capacity building for TB operational research by actively supporting young researchers from universities and regional health offices; vi) Assessment of the relevance and quality of operational research proposals submitted to the NTCP either for endorsement or funding support.

Strategy for research capacity development: Decentralized training targeted at young researchers

The TORG developed a plan for operational research in Indonesia in 2008 linked to research institutes and a process of strengthening research capacity outside the NTCP. It subsequently organized operational research courses including:

Designing and Conducting Health Systems Research Projects

- Volume I: Proposal Development and Fieldwork
- Volume II: Data analysis and report writing;

The training is organized around four groups per course of five persons (three from university and two from provincial or district health office). Participants are selected from national/provincial/district TB programmes and local universities. 23 out of 33 provinces have already participated.

Promotion of information exchange between researchers and the NTCP

A key feature of the programme is regular meetings and contact between the chair of the TORG and NTCP. The TORG also organizes national operational research dissemination conferences and sharing of results among the provinces participating in the TORG workshops.

Relevance, quality and prioritization of operational research

Over a two year period, TORG developed Guidelines for evaluation of research proposals that have been used to review 36 operational research proposals submitted to the NTCP; with 12 projects approved for funding.

Strong emphasis on publication, both nationally and internationally

Since it was established, the TORG has yielded three publications in international peer-reviewed journals. 15 provinces completed operational research projects with dissemination of recommendations; 10 provinces completed related projects, with recommendations implemented; 60% of participants (40 people) have gone on to engage in other research projects after the course.

Challenges and lessons learnt

Building capacity outside the NTCP is not without challenges. A number of researchers at universities do not have adequate knowledge or experience in TB programme implementation. This poses a potential risk of bias in the statement/formulation of research questions, which may be driven by interests of researchers, and not aligned to NTCP programme priorities. Externally-driven research may have limited stakeholder engagement, and as a result, implementation of research recommendations might be more difficult. Funding may be difficult to sustain given that the research outcomes may not be translated into policy and practice.

However, linking research institutes to the NTCP and strengthening research capacity demonstrated unique advantages. Researchers embedded in a scientific environment were well positioned to bring rigour and robust methodology to the research process. Research is the main activity of universities and research institutes, which were therefore able to dedicate human resources to development and completion of the research projects, while researchers in the NTCP often had (many) other duties. In addition, research undertaken by those outside the NTCP was perceived to be more independent, and this in turn allowed the NTCP to tap into research resources that were already available in-country. Given the commitment to research, building partnerships between research institutions and the NTCP forged TB research capacity that was sustainable, with a high feasibility of involvement of other disciplines (multidisciplinary research). This enabled the NTCP to have permanent access to problem-solving expertise at an affordable cost.

1.4 Nongovernmental organisations (NGOs) working with programmes can provide a number of opportunities. They work in specific settings and with vulnerable groups (e.g. prisoners, sex workers, MDR-TB patients etc) where researchers often have little or no access. An advantage is that those NGOs conducting operational research projects are usually involved in translating research findings into policy and practice on a local scale, which is of added programme benefit. NGOs that are relatively well resourced can bring in complementary human and material resources for implementation. In particular, NGOs can be strong in advocacy, which is of vital importance in raising the profile of operational research and catalysing changes to policy and practice *(3–5)*. A common problem with NGOs is that they may lack training and capacity in the research field, as well as the culture and skills for interacting with national programs. **Box 5** gives an example of an NGO in Brazil spearheading the process of promoting OR and research capacity development in an integrated manner (integrating implementers, academic institutions and community) to contribute to TB and TB/HIV control.

BOX 5: BRAZIL CASE STUDY

Experience of *Rede-TB* in Brazil: Development and impact evaluation of new health system intervention tools for TB control, with particular focus on TB/HIV and DR-TB

Background

During the period of 1980–2000, there were no NGOs working on TB control in Brazil, and almost no interaction between the NTCP, academia and the private sector. Health professionals and researchers did not perceive one another as partners. Development of policies for TB control was not driven by scientific evidence, and efficiency and impact of such policies was not monitored. The lack of investment in efforts to build research capacity was reflected in the limited range and type of research conducted during this period.

Multistakeholder partnership

In response to the lack of a systematic approach to development of research capacity in Brazil, a 'Foresight seminar on TB research and control' was organised in Rio de Janeiro in March 2001. The aim of the seminar was to identify strategies that would enable the development and evaluation of new products, technologies and strategies for TB control. A number of stakeholders were invited in order to mobilize political commitment and establish a systematic, sustainable system for research capacity development. These included representatives from national and provincial TB and AIDS programmes; public and private laboratories; research/education institutions; biomedical associations; the national regulatory agency and health Council; and relevant NGOs and companies. In this seminar, a consensus was reached regarding the need for a Tuberculosis Research Network as an innovative strategy to facilitate coverage of all gaps.

Promotion of information exchange between researchers and the NTCP

The Brazilian TB Research Network *(Rede-TB)* was created in April 2001 to spearhead the effort of developing TB control research capacity at the national level. The main objective of Rede-TB was to promote research and educational activities in an integrated manner in order to contribute to TB and TB/HIV control.

Rede-TB is an NGO initially constituted by an interdisciplinary group of researchers and students working on health sciences, engineering and education, who were later joined by civil society partners and health service (TB and AIDS) representatives (from the federal, state, and municipality levels). *Rede-TB* is a membership organization of 160 members from 47 institutions, including researchers, policy-makers, and AIDS and TB managers. The strategy was to establish a network with self-organizing nodes called 'Coordination Areas' or 'Working Groups' according to specific areas of research interest, linked to a common vision that formed the basis of a platform for membership engagement, identification of gaps and research partnership building.

Rede-TB played a key role in the creation of the Brazilian Partnership Against Tuberculosis (BPAT) in 2004. *Rede-TB* researchers were invited by the Ministry of Health (MoH) to help in the definition of a National TB Research Agenda in 2004, 2007, and 2010. It also led to the establishment of a number of research initiatives including the scientific and technological platform in TB diagnosis developed together with the Latin America Network for TB Control and WHO. Research institutes and universities are now engaged in the implementation of TB operational research, with good interaction between basic and clinical research using several laboratories with standardized procedures. This has led to prominent interaction between governmental institutions (NTCP, Central Laboratories, Sanitary Supervision Institutions) and industry.

Strategy for research capacity development: Decentralized multilevel training in health systems research methods and operational research

With an award from the Science Technology Department of the MoH, Rede-TB piloted research training in Rio de Janeiro Province (2004). Building on this experience, the International Collaborative Operational and Health System Research on TB and HIV/AIDS (ICOHRTA) Research Capacity project expanded training to five additional provinces using funding from United States National Institutes of Health and the MoH to implement research training on clinical, operational and health sector aspects. Health care workers and community leaders are invited to participate in research methodology courses of various duration and levels, according to their scientific background, and to develop their own research projects, based on their service experiences and issues. The best projects are supervised by Rede-TB researchers and funded by ICOHRTA and the MoH.

Strong emphasis on expansion of academic output

An evaluation of the trends of scientific articles focusing on TB in Brazil published between 1986 and 2006 showed that among 1054 publications, only 6.8% were on operational research and 3.5% included qualitative evaluations (*20*). Rede-TB contributed significantly to the expansion of the academic output in TB in Brazil over recent years. Analysis of CAPES (the Brazilian agency for graduate studies) databank on theses and dissertations on TB in 2004–2008 revealed that 42% of PhD theses and 37.4% of MSc dissertations were mentored by Rede-TB members.

A focus on results, translation into policy and practice

Following development of broad consensus that any guidelines modification should be seen by guideline developers and policy-makers as a new opportunity for technology incorporation activity, TB and AIDS Programme coordinators (at federal, state and municipality levels) prioritized operational research to evaluate the effectiveness of existing tools used in TB control, and the impact of new tools before introduction into practice. Working in an integrated way, *Rede-TB* researchers have received significant national and international funding for basic, clinical and operational research. Recently, Rede TB researchers took part in the development of protocols to evaluate the impact of the introduction in the public health system of new diagnostic tests (e.g. Xpert MTB/RIF, line-probe assays), and received funding to investigate the accuracy of a new molecular test for pulmonary TB diagnosis (DETECT TB) developed by Brazilian scientists in collaboration with national companies.

In addition, *Rede-TB* conducted a nationwide qualitative study to map performance of health services in the diagnosis of tuberculosis in 10 metropolitan areas of Brazil from the perspectives of patients, health professionals, managers of health units and civil society organizations. The investigation found low effectiveness of TB care, verified by the low suspicion of TB by professionals (even those with high knowledge of TB disease), high number of referrals to hospital/emergency and primary care services, and a deficiency in communication flow and referring activities between services at all levels (including the laboratory network) (*21–23*). Moreover, TB diagnostic tests were not performed routinely in the health facility initially visited by patients. Patient travel requirements until effective TB diagnosis were a factor that increased the time between symptom onset and TB diagnosis. This scenario occurred even in those cities with more decentralized health systems, where healthy family programmes have received high priority.

TB operational research led to key changes in national policy and practice described in the latest National Guidelines, released in March 2010. This included the creation of a TB control programme for vulnerable citizens (i.e. indigenous, homeless, prison inmates); the development of tools to be used for TB infection control in health settings (hospitals, prisons, and primary health services, etc); the adoption of routine TB culture for all HIV-infected TB suspects, drug susceptibility testing for drug-resistant TB suspects; and HIV

testing for all TB suspects; a change of treatment regimens for newly diagnosed TB patients (using a fixed-dose combination with four drugs); and standardized drug-resistant TB treatment.

Challenges and lessons learnt

The creation of *Rede-TB* helped to create renewed bridges between academia, the health system and civil society. Engagement of academia strengthened the capacity of TB and AIDS managers, health professionals, health-system users and the national Industry in the production of scientific knowledge that responds to local demands, through operational and health system research approaches.

Rede-TB is a unique example of NGO leadership and coordination of a national research capacity development process that has demonstrated significant, measureable results. Challenges remain, however. Sustainability of achievements will require continued investments in research capacity development; funding research through the established research coordination areas; training health-care workers, managers, and community leaders towards a change in attitude and practices to improve conditions of access, the continuity of care, and communication flow between TB services (at all levels), as well as with the civil society. In addition, ongoing confusion with monitoring and evaluation will need to be resolved, and processes strengthened to incorporate research into the national TB programme.

In summary, there are often many locations and partners intervening at national level, and each of them has comparative advantages. The best operational research option would be a partnership model led by national programmes that would promote better involvement, co-ownership and responsibility of programme staff with researchers. It is particularly important to integrate funding and resources for operational research into a national programme so as to get decisional power for setting the national research agenda on TB, evaluate how and when new technologies should be incorporated and decide on implementation. The latter has often been the monopoly of foreign institutions and this imbalance needs correction.

2. What sort of research capacity is needed?

The ultimate goal of research *capacity building* is to provide the ability to individuals, organizations or systems to perform and utilize health research effectively, efficiently, and sustainably *(24)*. It is distinct from *training* (which is an organized activity aimed at imparting information and/or instructions to improve the recipient's performance or to help him/her attain a required level of knowledge or skill). At the programme level, the following are required in the process of capacity building:

a) ability to define the right research question(s);

b) knowledge of the steps needed to conduct a methodologically rigorous study;

c) capacity to collect, store, quality-assure, manage, and analyse data (both qualitative and quantitative);

d) capacity to write up the results for publication in peer-reviewed journals;

e) engagement of the appropriate range of stakeholders (including community and policy-makers) in the whole process of research to facilitate ultimate utilization of the results.

While writing skills for publishing present a challenge widely acknowledged across a number of NTCPs, fundamental capacity gaps in defining critical research questions and in research methodology have been recognized. In this regard, this document attempts to provide a reference tool for NTCP managers and researchers to conduct operational research projects through brief synopses of suitable research methodologies for the set priorities (see **Annex II**).

3. How should capacity be built?

3.1 Training courses:

There has been much investment in operational research training by various organisations such as WHO, the US Centers for Disease Control (CDC), Médecins sans Frontières (MSF), Japan Anti-Tuberculosis Association (JATA), The International Union Against TB and Lung Disease (the Union) and many others (this list is not exhaustive), but there has been limited evaluation of these models and of their impact on research and programme implementation. A recent publication from the Research Institute of TB in Japan showed that, of all participants attending an international training course in Japan between 2001 and 2007, only 40% started operational research projects, and none wrote a scientific paper *(6)*. The main cited reasons for failure to implement and complete studies were: lack of time, lack of funds, disapproval by supervisors and lack of writing skills. In addition, weak or non-existent retention strategies and long-term career opportunities further contribute to attrition of trained researchers. Training can, therefore, contribute to capacity building, but unless it is sufficiently integrated with development programmes or collaborations, or supported by other longer term capacity development initiatives, such as mentorship programmes, it may not result in changes in programme implementation such as improved patient care and public health.

3.2 Integrated and targeted capacity development approaches:

Capacity development can be integrated within development programmes or research collaborations as illustrated in **Boxes 2 (section III) 3, 4 and 5 (Annex VI)**, resulting in research networks that support TB Programmes. Such integrated, country-based approaches can be supported by targeted and novel processes of learning, such as the one developed by The Union and MSF along with partners *(7)*. The purpose of this training is to develop the practical skills for conducting operational research among doctors, nurses, data analysts and other programme staff, and ensure publication of results as appropriate. This model is based upon a hands-on modular approach that is practical, with strict criteria for candidate selection. Strong and sustained mentorship is provided by experienced facilitators, and candidates are offered long-term retention opportunities, such as operational research fellowships. Evaluation of training is performance-based, and candidates are required to fulfil set milestones after each module in order to continue. Preliminary results of this approach are very encouraging, with the first twelve candidates who started the course in 2009 having submitted a total of 14 papers for publication, of which, 11 were already published or in press within the twelve months of the course finishing.

3.3 Indicators for evaluation of capacity development:

Ideally, research capacity development should result in researchers who are able to independently develop, manage, and obtain funding for their own research programmes and ensure that their findings are being disseminated and used as appropriate. Publication of research results in peer-reviewed journals is a crucial output of any research process, allowing dissemination of research findings, and a good indicator of a successfully completed research study. In addition, the evaluation of capacity development can be carried out through a series of quantitative and qualitative indicators, as suggested in **Table 5**, although these would still need to be validated.

ISBN 978 92 4 154825 0

TABLE 5: SUGGESTED INDICATORS FOR EVALUATION OF OPERATIONAL RESEARCH CAPACITY DEVELOPMENT WITHIN TB PROGRAMMES (FROM REFERENCE 25)

Quantitative indicators	Qualitative indicators
• Presence of an operational research office • Presence of a research officer/focal point • Ear-marked institutional or programme funding for operational research • Evidence that operational research is in the annual budget and plans • Number of publications in national and/or international peer-reviewed journals • Number of conference presentations • Number of degrees awarded (e.g. Masters, PhD, Post-doctoral) • Number staff starting/completing graduate programme • Number of research grants/other funding secured • Number of trained staff employed for operational research with existing organisations and collaborators (e.g. government agencies, NGOs, etc.) • Number of research/support staff • Average time for PhD completion • Career trajectories – number of promotions.	• Access to mentors • Academic support (security of employment/ length of tenure) • Programme quality (as judged by management quality indicators e.g. on leadership, team-working, decision-making) • Teaching quality indicators (teaching efficacy, teaching methods) • Learning quality indicators (learning attitude, attendance rate) • Learners' confidence and competence in research outcomes (i.e. attitudes, intentions).

Conclusion

There is a need to rapidly and consistently develop capacity to conduct sustainable operational research at the programme level. The importance of operational research has been emphasized in the recent revision of the *Global Plan to Stop TB 2011-2015*, that gives directions about the operational research needs towards the 2015 MDG targets and Stop TB Partnership goals. Development of capacity must be rapid as training models will need to deliver within the time frame of these targets. Approaches to training must be practical and target-oriented, with defined products and outputs that can influence policy and practice. The long-term vision is the development of leadership in operational research through collaborative networks in various parts of the world.

As it stands, capacity building is a paramount element of strengthening operational research at the programme level. In addition, there are a number of other enabling factors that foster a favourable environment for research within TB programmes.

REFERENCES

1. Harries AD. Integration of operational research into National Tuberculosis Control Programmes. *Tuberculosis* (Edinb), 2003; 83(1-3):143–7.
2. TB Research: Putting research into policy and practice: the experience of the Malawi National Tuberculosis Programme. Geneva, World Health Organization, 1999.
3. Zachariah R et al. Voluntary counselling, HIV testing and adjunctive cotrimoxazole reduces mortality in tuberculosis patients in Thyolo, Malawi. *AIDS*, 2003; 17(7):1053–61.
4. Zachariah R et al. Payment for antiretroviral drugs is associated with a higher rate of patients lost to follow-up than those offered free-of-charge therapy in Nairobi, Kenya. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, 2008; 102(3):288–93.
5. Zachariah R NF, Draguez B, Yun O, Reid T. Conducting operational research within a non-governmental organisation: the example of Médecins Sans Frontières. *International Health*, 2010; 2:1–8.
6. Ohkado A et al. Evaluation of an international training course to build programmatic capacity for tuberculosis control. *International Journal of Tuberculosis and Lung Disease*, 2010; 14(3):371–3.
7. The Union/MSF approach to operational research training. Paris, International Union Against TB and Lung Disease, 2009.
8. United Nations Millennium Development Goals. New York, United Nations, 2001.
9. Harries AD et al. Tuberculosis research in Malawi: making it count. *Recent Advances and Research Updates*, 2001; **2**:103–118.
10. Nyangulu DS et al. Tuberculosis in a prison population in Malawi. *The Lancet*, 1997; **350**:1284–1287.
11. Harries AD, Nyirenda TE, Yadidi AE, Gondwe MK, Kwanjana JH, Salaniponi FM. Tuberculosis control in Malawian prisons: from research to policy and practice. *International Journal of Tuberculosis and Lung Disease*, 2004; **8**:614–617.
12. Harries AD, Salaniponi FML. Recurrent tuberculosis in Malawi: improved diagnosis and management following operational research. *Transactions of the Royal Society for Tropical Medicine and Hygiene*, 2001; 95:503–504.
13. Salaniponi FM et al. Decentralisation of treatment for patients with tuberculosis in Malawi: moving from research to policy and practice. *International Journal of Tuberculosis and Lung Disease,* 2003; 7:S38–S47.
14. Meijnen S et al. Outcome of patients with tuberculosis who transfer between reporting units in Malawi. *International Journal of Tuberculosis and Lung Disease,* 2002; 6:666–671.
15. Zachariah R, Spielmann ML, Chinji C, et al. Voluntary counselling, HIV testing and adjunctive cotrimoxazole reduces mortality in tuberculosis patients in Thyolo, Malawi. *AIDS*, 2003; 17:1053–1061.
16. Mwangulu FBD et al. Cotrimoxazole prophylaxis reduces mortality in human immundeficiency virus-positive tuberculosis patients in Karonga District, Malawi. *Bulletin of the World Health Organization*, 2004; 82:354–363.
17. Chimzizi RB et al. Voluntary counselling, HIV testing and adjunctive cotrimoxazole are associated with improved TB treatment outcomes under routine conditions in Thyolo District, Malawi. *International Journal of Tuberculosis and Lung Disease,* 2004; 8:579–585.
18. Chimzizi RB et al. Counselling, HIV testing and adjunctive cotrimoxazole for TB patients in Malawi: from research to routine implementation. *International Journal of Tuberculosis and Lung Disease,* 2004; 8:938–944.
19. Harries AD, Salaniponi FM, Nunn PP, Raviglione M. Performance-related allowances within the Malawi National Tuberculosis Control Programme. *International Journal of Tuberculosis and Lung Disease,* 2005; 9:138–144.

20. Kritski AL et al. Two decades of research on tuberculosis in Brazil: state of the art of scientific publications. *Revista de Saúde Pública*, 2007;41 (Suppl):9–14.

21. Motta MCS et al. Access to tuberculosis diagnosis in Itaboraí Municipality – Rio de Janeiro State – Brazil: the patients' point of view. *International Journal Tuberculosis and Lung Disease*, 2009;13(9):1137–41.

22. Arcêncio RA et al. City tuberculosis control coordinators' perspectives of patient adherence to DOT in São Paulo State, Brazil, 2005. *International Journal Tuberculosis and Lung Disease*, 2008;12(5):527–31.

23. Figueiredo TM et al. Performance of primary health care services in tuberculosis control. *Revista de Saúde Pública,* 2009;43(5):825–31.

24. Bates I et al. Evaluating health research capacity building: an evidence-based tool. *PLoS Medicine*, 2006; 3(8):e299.

25. Bates I et al. Indicators of sustainable capacity building for health research: analysis of four African case studies. *Health Research Policy and Systems,* 2011;28;9(1):14: doi:10.1186/1478-4505-9-14.