Measuring progress in global tuberculosis control:

WHO Policy and Recommendations

DRAFT

IMPORTANT NOTE:

The main policies and recommendations are summarized on page 6 (synthesis of recommendations for measuring TB incidence, prevalence and mortality), page 10 (TB incidence), page 20 (TB prevalence), page 28 (TB mortality) and page 34 (evaluation of the impact of TB control on TB incidence, prevalence and mortality).

Policies and recommendations on the measurement of TB prevalence (pages 19 to 26) are taken directly from the meeting report of the December 2007 meeting of the Task Force, except for text highlighted in yellow. Text highlighted in yellow is text that has been added or modified following the comments, advice and recommendations provided by WHO's Strategic and Technical Advisory Group on TB (STAG-TB) in June 2008. The issues covered by the text highlighted in yellow will be given particular attention during the presentation of this policy paper on Day 1 of the September 2008 Task Force meeting.

Policies and recommendations on the measurement of TB mortality (pages 27 to 30) are based on the Lancet Infectious Diseases paper (Dye et al) published in January 2008, which was endorsed at the December 2007 meeting of the Task Force.

Policies and recommendations on TB incidence are based on the material presented in the Lancet Infectious Diseases paper, and on discussions about ARI surveys during the June 2006 and December 2007 meetings of the Task Force. However, some of the ideas/recommendations have been developed further in this Policy Paper. For this reason, it is the section on measuring TB incidence (pages 7 to 18) that Task Force members should review most carefully in advance of the meeting, and which will be given particular attention during the September 2008 meeting itself.

Please also note that the reference numbers are correct but are not always in the correct order in this draft. This will be addressed in time for the Task Force meeting.

Text highlighted in green indicates text that requires further work/discussion.

Draft version 2. Date: 9 September 2008
TABLE OF CONTENTS

1. INTRODUCTION ........................................................................................................ 1

2. WHAT ARE THE GLOBAL TARGETS FOR TB CONTROL? .......................... 3

3. THE WHO GLOBAL TASK FORCE ON TB IMPACT MEASUREMENT .... 4

4. EVIDENCE BASE FOR WHO POLICY AND RECOMMENDATIONS ON MEASURING TB INCIDENCE, PREVALENCE AND MORTALITY ................. 4

5. WHO POLICY PACKAGE FOR MEASURING TB INCIDENCE, PREVALENCE AND MORTALITY ......................................................... 6

6. WHO POLICY AND RECOMMENDATIONS FOR MEASURING TB INCIDENCE ........................................................................................................ 7

7. WHO POLICY AND RECOMMENDATIONS FOR MEASURING TB PREVALENCE ........................................................................................................... 19

8. WHO POLICY AND RECOMMENDATIONS FOR MEASURING TB MORTALITY ............................................................................................................. 27

9. PRODUCING A SERIES OF ESTIMATES OF TB INCIDENCE, PREVALENCE AND MORTALITY FROM 1990 TO 2015 ........................................... 30

10. METHODS TO EVALUATE THE CONTRIBUTION OF TB CONTROL TO CHANGES IN TB INCIDENCE, PREVALENCE AND MORTALITY .............. 31

11. CONCLUSIONS ..................................................................................................... 36

12. REFERENCES ..................................................................................................... 37

13. ANNEXES .......................................................................................................... 41
1. Introduction

Tuberculosis (TB) was declared a global public health emergency by the World Health Organization (WHO) in 1993. At that time, there were an estimated 7 million new TB cases per year and 1.6 million deaths from TB, compared with an estimated 9.2 million cases and 1.7 million deaths from TB in 2006. With the introduction of the DOTS strategy in the mid-1990s, WHO began to systematically monitor progress in TB control using the two global indicators and related targets for TB control established by the World Health Assembly (WHA) in 1991.¹ The indicators were (i) the percentage of estimated new (incident) cases of smear-positive TB detected in DOTS programmes (case detection rate) and (ii) the percentage of detected cases successfully treated (successful treatment rate). The targets were to reach a 70% case detection rate and an 85% treatment success rate by 2000, a target year that was later reset to 2005. In 2005, the global case detection rate under DOTS reached 58% and the treatment success rate was 84.7%.²

The WHA targets have proved very useful for stimulating greater efforts to control TB in endemic countries, and case detection and treatment success rates are well-established and widely-used indicators of national TB programme (NTP) performance at global, regional and country level. Their limitation is that they do not directly measure whether or not the epidemiological burden of TB - measured in terms of cases and deaths - is being reduced. In other words, they measure the "outcomes" of TB control programmes, but not their impact.

Starting from the year 2000, global targets for TB control have been extended to include targets for reducing cases and deaths. These newer "impact" targets have been set within the Millennium Development Goals (MDGs) and by the Stop TB Partnership, with target years of 2015 and 2050.² The principal targets are to halt and reverse incidence by 2015, and to halve prevalence and death rates by 2015 compared to their level in 1990. Both the MDG and Stop TB Partnership targets are part of the WHO's Stop TB Strategy, and they were recognized in a WHA resolution on TB passed in 2007 (WHA 60.19).

Achieving the impact targets is now the focus of international and national efforts to control TB, and demonstrating whether or not they are achieved is of major importance for individual countries, the UN, WHO and the Stop TB Partnership, and a variety of technical, financial and development agencies. However, measuring changes in TB incidence, prevalence and deaths is challenging, and countries with a high burden of TB as well as their technical, financial and development partners are seeking clear guidance about how it should be done. This policy paper, which is the second in a series being produced by the WHO's Stop TB Department, is designed to provide such guidance. It is based primarily on the recommendations of the WHO Global Task Force on TB Impact Measurement, while also drawing on the wider literature on monitoring and evaluation.

The document is structured in nine major sections, which are:

- **What are the global targets for TB control (2)?** This section defines the global targets that have been set for 2015 and 2050 within the MDG framework and by the Stop TB Partnership;
- **The WHO Global Task Force on TB Impact Measurement (3).** This section describes the mandate, membership and organization of the Task Force;

Draft version 2. Date: 9 September 2008
• Evidence base for WHO policy and recommendations for measuring TB incidence, prevalence and mortality (4). This section explains how the work of the Task Force from 2006 to 2008 underpins the policy and recommendations included in this policy paper;

• WHO policy package for measuring TB incidence, prevalence and mortality to 2015 and beyond (5). This section summarizes the major WHO policies and recommendations for measuring TB incidence, prevalence and mortality from 2009 onwards. More explanation and details are then provided in sections 6 to 9;

• WHO policy and recommendations for measuring TB incidence (6). This section describes the six methods that are available to measure TB incidence, and the number of countries for which these methods were used to produce the most recent WHO estimates of TB incidence. It then provides WHO policy and recommendations for how to measure TB incidence (both its absolute value and trends over time) from 2009 onwards. Guidance material and tools, and sources of technical and financial support, are identified;

• WHO policy and recommendations for measuring TB prevalence (7). This section describes the two major methods that are available to measure TB prevalence, and the number of countries for which these methods were used to produce the most recent WHO estimates of TB prevalence. It then provides WHO policy and recommendations for how to measure TB prevalence (both its absolute value and trends over time) from 2009 onwards. Guidance material and tools, and sources of technical and financial support, are identified;

• WHO policy and recommendations for measuring TB mortality (8). This section summarizes the three major methods that are available to measure TB mortality, and explains the number of countries for which these methods were used to produce the most recent WHO estimates of TB mortality. It then provides WHO policy and recommendations for how to measure TB mortality (both its absolute value and trends over time) from 2009 onwards. Guidance material and sources of technical and financial support are identified;

• Estimates of TB incidence, prevalence and mortality from 1990 to 2015 (9). The recommendations in sections 6 to 9 focus on strengthening of surveillance and implementation of surveys of the prevalence of TB disease from 2008 onwards. To assess whether or not the global targets of halving TB prevalence and death rates by 2015 compared with 1990 are achieved, estimates are needed from 1990 to 2015. This section explains how WHO, together with other members of the Task Force, propose to produce such estimates;

• WHO policy and recommended methods for measuring the contribution of TB control to changes in TB incidence, prevalence and mortality (10). Besides measuring changes in TB incidence, prevalence and mortality, it is important to evaluate the extent to which these changes are due to changes in TB control. This section provides guidance on such evaluations.

The document also includes a list of references and six annexes. Annex 1 lists the membership of the WHO Global Task Force on TB Impact Measurement. Annex 2 provides extracts from the WHA resolution on TB passed in 2007. Annexes 3 and 4 provide further details on surveys of the prevalence of TB infection and disease, respectively. Annex 5 lists countries that met Task Force criteria for implementing surveys of the prevalence of TB disease, but which are not among the 21 global focus countries selected by the Task Force. Annex 6 lists technical and financial partners for ongoing or planned surveys of the prevalence of TB disease.
2. What are the global targets for TB control?

Global targets and indicators for TB control have been developed within the framework of the MDGs as well as by the Stop TB Partnership and the WHA. While the targets are global, they are commonly used within WHO regions and by individual countries. The principal and most widely-quoted targets and indicators are summarized in Box 1.

The impact targets are to halt and reverse the incidence of TB by 2015, and to halve TB prevalence and death rates by 2015 compared with a baseline of 1990. The incidence target is MDG Target 6.C, and it has been adopted by the Stop TB Partnership. The MDG framework includes indicators, but not targets, for TB prevalence and death rates. The targets of halving TB prevalence and death rates by 2015 compared with 1990 were set by the Stop TB Partnership, based on a resolution of the year 2000 meeting of the Group of Eight (G8) industrialized countries (in Okinawa, Japan). Halving TB death and prevalence rates by 2015 are much more difficult targets to achieve than is the target of ensuring that incidence is falling by 2015.

As explained in the introduction, outcome targets were established by the WHA in 1991. When first set, the targets were to achieve a case detection rate of at least 70% under DOTS and to reach a treatment success rate of at least 85% in DOTS cohorts by 2000. The target year was later reset to 2005. While 2005 has now passed, the indicators of case detection and treatment success remain part of the MDG and Stop TB Partnership framework for measuring progress in TB control.

Box 1. Goals, targets and indicators for TB control

**Millennium Development Goals**

Goal 6: Combat HIV/AIDS, malaria and other diseases
Target 6.C: Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases

Indicator 6.8: Incidence, prevalence and death rates associated with tuberculosis
Indicator 6.9: Proportion of tuberculosis cases detected and cured under DOTS (the internationally recommended strategy for TB control)

**Stop TB Partnership targets**

By 2005: At least 70% of people with sputum smear-positive TB will be diagnosed (i.e. under the DOTS strategy), and at least 85% cured. These targets were first set by the WHA of WHO in 1991

By 2015: The global burden of TB (per capita prevalence and death rates) will be reduced by 50% compared to 1990 levels. This means reducing prevalence to \( \approx 150 \) cases per 100 000 population or lower and deaths to \( \approx 15 \) deaths per 100 000 population per year or lower by 2015 (including TB cases co-infected with HIV). The number of people dying from TB in 2015 should be less than 1 million, including those co-infected with HIV

By 2050: The global incidence of active TB will be less than 1 case per million population per year (the criterion for elimination adopted within the USA)

---

1 To avoid double-counting of deaths from AIDS and TB, WHO world health statistics on TB deaths exclude TB deaths among people co-infected with HIV. However, the aim of TB control is to eliminate the disease from whole populations. In the annual report on global TB control, WHO therefore publishes TB statistics both overall and for the HIV-positive and negative sub-populations separately.
The Stop TB Partnership has also set the goal of eliminating TB by 2050. The target for elimination is to have less than 1 case per million population.²

Besides outcome and impact targets, it is also possible to set other kinds of targets for TB control (input, process and output targets).⁶ For example, the Stop TB Partnership's Global Plan to Stop TB for the period 2006 to 2015 includes financial (input), process (e.g. percentage of the population for whom community-based care is available) and output (e.g. number of cases treated in DOTS programmes, or number of HIV-positive TB patients enrolled on antiretroviral treatment) targets. Since these targets are not the subject of this policy paper, they are not discussed further here.

3. The WHO Global Task Force on TB Impact Measurement

To respond to the need to measure progress towards the 2015 targets for reductions in TB incidence, prevalence and mortality, WHO established a Global Task Force on TB Impact Measurement (hereafter the Task Force) in June 2006. The Task Force includes experts in TB epidemiology, representatives from major technical and financial agencies, and representatives from countries with a high burden of TB (Annex 1). Its mandate is to produce a robust, rigorous and widely-endorsed assessment of whether the 2015 targets for reductions in TB incidence, prevalence and mortality are achieved at global level, for each WHO region and in individual countries; to regularly report on progress towards these targets in the years leading up to 2015; and to strengthen national capacity in monitoring and evaluation of TB control. The importance of this work is reflected in a WHA resolution passed in 2007 (WHA 60.19), in which WHO is required by its member states to report on whether the 2015 global targets for TB control are achieved, to report on progress in the interim, and to help strengthen health information systems (Annex 2).

4. Evidence base for WHO policy and recommendations on measuring TB incidence, prevalence and mortality

The WHO policies and recommendations included in this document are based on the following four foundations:

- a systematic review of the methods that are available to measure TB incidence, prevalence and mortality, which was published in a peer-reviewed journal in January 2008;
- discussions and related recommendations of the WHO Task Force on TB Impact Measurement, from three meetings held between June 2006 and September 2008;
- recommendations and advice provided by the WHO's Strategic and Technical Advisory Group on TB (STAG-TB), including review of a draft version of this document in June 2008; and
- the wider literature on monitoring and evaluation.

² The criterion for elimination presented in Box 1 is that defined by the Centers for Disease Control, USA. It differs from a European recommendation that elimination be defined as less than one smear-positive case per million population.
All policies and recommendations included in this document have been endorsed by the full Task Force.

The systematic review of the methods that are available to measure TB incidence, prevalence and mortality was co-authored by members of the Task Force and published in *Lancet Infectious Diseases* (LID) in January 2008. The LID paper was endorsed at the December 2007 meeting of the Task Force, and was itself based on the discussions and recommendations of the first Task Force meeting held in June 2006. The methods used for the systematic review of the evidence on which the LID paper was based are summarized in Box 2. A few of the 137 references cited in the LID paper are quoted in this document, where this is thought to be particularly useful. Otherwise, for the full set of references used, readers should consult the original paper.

**Box 2. Methods used for a systematic review of the evidence about how TB incidence, prevalence and mortality can be measured**

<table>
<thead>
<tr>
<th>Search strategy and selection criteria for Lancet Infectious Diseases paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data for the review were identified by searches of Medline, ISI Web of Science, plus the libraries of all authors. Search terms used in the online databases were, in various combinations, “tuberculosis”, “environmental mycobacteria”, “incidence”, “prevalence”, “mortality”, “infection”, “annual risk of infection”, “tuberculin skin test”, “mixture method”, “mirror-image method”, “survey”, “surveillance”. Among many thousands of papers on this topic, only the key papers in English that are relevant to the methodology were selected.</td>
</tr>
</tbody>
</table>

Task Force meetings were held in June 2006, December 2007 and September 2008. The 2006 meeting covered general recommendations about how TB incidence, prevalence and mortality should be measured. The 2007 meeting discussed the three major strategic tracks of work to be pursued by the Task Force (assessment and use of routine surveillance data; implementation of surveys of the prevalence of TB disease and infection; and methods used to produce estimates of TB incidence, prevalence and mortality from 1990 to 2015 using surveillance and survey data) and how these would be organized. Particular attention was given to the assessment of which countries should implement surveys of the prevalence of TB disease and the methods to be used in such surveys. The 2008 meeting focused on the direct and indirect measurement of TB incidence (its absolute value and trend) using routine surveillance data supplemented by data from sources such as operational research studies and surveys.

Details of Task Force meeting discussions and recommendations can be found in meeting reports, a series of background papers, and peer-reviewed papers.

A draft version of this policy document was reviewed by STAG-TB at its June 2008 meeting. The advice and recommendations that were provided, which focused on surveys of the prevalence of TB disease, were addressed during finalization of the document.

The recommendations in section 10 in particular draw on the wider literature on monitoring and evaluation, as well as publications that have specifically evaluated the impact of TB control on the epidemiological burden of TB.

---

*Please note that it is likely that some of the recommendations for measuring TB incidence in this draft will be modified or refined during the September 2008 Task Force meeting.*
5. WHO policy package for measuring TB incidence, prevalence and mortality

The WHO policy package for measuring TB incidence, prevalence and mortality is shown in Table 1. This package is explained in more detail in sections 6 to 10.

Table 1: The WHO policy package for measuring TB incidence, prevalence and mortality from 2008 to 2015 and beyond

<table>
<thead>
<tr>
<th>General</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Improve routine surveillance until notified TB cases reliably record all incident TB cases and vital registration data reliably record all TB deaths</td>
</tr>
<tr>
<td>2. Strengthen capacity in monitoring and evaluation</td>
</tr>
<tr>
<td>3. Periodic review and updating of data, assumptions and analytical methods used to produce WHO estimates of TB incidence, prevalence and mortality</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement of TB incidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Periodic analysis of the reliability and coverage of TB notification data to estimate the total number of incident TB cases and trends in TB incidence</td>
</tr>
<tr>
<td>5. Standard framework and related tool for systematic analysis of the reliability and coverage of TB notification data, to be developed and promoted by the Task Force</td>
</tr>
<tr>
<td>6. Certification of TB notification data for countries where analyses using the standard framework/tool show TB notification data are a close proxy (direct measure) of TB incidence</td>
</tr>
<tr>
<td>7. Cross-validation of estimates of TB incidence using TB mortality data from vital registration systems</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement of TB prevalence</th>
</tr>
</thead>
<tbody>
<tr>
<td>8. Surveys of the prevalence of TB disease in 21 global focus countries, designed and implemented according to WHO guidelines and Task Force recommendations</td>
</tr>
<tr>
<td>9. Indirect estimates of TB prevalence based on estimates of TB incidence and the duration of TB disease for countries that do not implement surveys of the prevalence of TB disease</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Measurement of TB mortality</th>
</tr>
</thead>
<tbody>
<tr>
<td>10. Development/strengthening of vital registration systems so that all TB deaths are reliably recorded</td>
</tr>
<tr>
<td>11. Sample vital registration used as an interim solution where national vital registration systems are not yet available</td>
</tr>
<tr>
<td>12. Indirect estimates of TB mortality based on estimates of TB incidence and the case fatality rate for countries without reliable national or sample vital registration systems</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Evaluation of the impact of TB control</th>
</tr>
</thead>
<tbody>
<tr>
<td>13. Periodic studies to evaluate the impact of TB control on TB incidence, prevalence and mortality</td>
</tr>
</tbody>
</table>
6. WHO policy and recommendations for measuring TB incidence

6.1 Definition of TB incidence

The incidence of TB is the number of cases of TB that occur each year. It is usually reported as the total number of cases, or as the number of cases for a given unit of population (e.g. number of incident cases per 100 000 population, or number of incident cases per million population). TB incidence changes relatively slowly. Even high-quality control programmes are expected to achieve a decline in the incidence rate of only 5–10% per year (in the absence of HIV co-infection).

6.2 Methods for measuring TB incidence

Six methods are available to measure TB incidence in a given year.12 These are:

1. Direct measurement from TB notification data i.e. TB notifications are assumed to equal (or be a very close proxy for) TB incidence. This method is only possible if there is strong evidence that all TB cases are diagnosed and that all diagnosed TB cases are notified;

2. Direct measurement from prospective cohort studies, in which the number of incident TB cases is measured by following a large population cohort (around 400 000 people need to be followed for one year if incidence is thought to be around the global average of 100 per 100 000 population);

3. Indirect estimation based on surveys of the annual risk of infection (ARI) with TB. Incidence is estimated as the ARI multiplied by 50 (based on the Styblo assumption or "rule" that an ARI of 1% is equivalent to around 50 smear-positive cases per 100 000 population);

4. Indirect estimation from studies of the prevalence of TB disease. Incidence is estimated as the prevalence of TB disease divided by the estimated average duration of disease;

5. Indirect estimation from vital registration data that include TB mortality data. TB incidence is estimated as the number of TB deaths divided by the estimated case fatality rate; and

6. Indirect estimation from assessment of the completeness of TB notification data. For example, if TB notifications are estimated to include 60% of incident cases (60% case detection rate), then TB incidence is estimated as TB notifications/0.6.

To estimate trends in TB incidence, these six methods can be used every year, or at regular intervals:

- Method 1 (direct measurement from TB notification data) can be used every year if the necessary data exist;
- Method 2 (prospective cohort study) could be used at regular intervals, but in practice the challenges of costs and logistics mean that such studies have hardly ever been undertaken (existing examples are limited to a specific population group in the Republic of Korea and to a limited geographical area in south India)13, 14;
- Methods 3 and 4 can be used at periodic intervals, provided that surveys use consistent methods;
Table 2. Number of countries for which the major methods available to measure TB incidence (in absolute terms) and trends in TB incidence were used to produce WHO estimates for 2006

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Direct measurement from TB notification data (mostly in 1997)*</td>
<td>20</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>20</td>
</tr>
<tr>
<td>2. Prospective cohort study (direct method)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>3. ARI (tuberculin) survey (indirect method)</td>
<td>-</td>
<td>-</td>
<td>India, Somalia</td>
<td>-</td>
<td>-</td>
<td>15</td>
<td>Myanmar, DPRK</td>
<td>19</td>
<td></td>
</tr>
<tr>
<td>4. Disease prevalence survey (indirect method)</td>
<td>-</td>
<td>-</td>
<td>Bangladesh, Indonesia, China</td>
<td>-</td>
<td>-</td>
<td>9</td>
<td>Pakistan</td>
<td>13</td>
<td></td>
</tr>
<tr>
<td>5. Estimates or counts of TB deaths from vital registration data (indirect method)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>Brazil, S. Africa, Mexico</td>
<td>-</td>
<td>-</td>
<td>3</td>
<td></td>
</tr>
<tr>
<td>6. Assessment of completeness of TB notification data in a specific year (mostly in 1997) (indirect method)</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>151</td>
<td>6***</td>
<td>157</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>20</td>
<td>-</td>
<td>4</td>
<td>1</td>
<td>3</td>
<td>175</td>
<td>9</td>
<td>212</td>
<td></td>
</tr>
</tbody>
</table>

*considered here as direct measurement from TB notification data because the case detection rate was estimated to be 90% or more.
**country's own notification data for 97 countries, regional trend for 98 countries. For almost all countries, trends in TB notifications (all forms of case) assumed to be the same as trends in TB incidence.
***Belize, Iraq, PNG, Sri Lanka, Thailand, Timor-Leste.

Example of how to read the table: There are two countries for which (i) TB incidence was estimated in absolute terms in 1997 from ARI survey data and (ii) trends in TB incidence have been estimated from a series of ARI surveys (India, Somalia). There are 151 countries for which (i) TB incidence in absolute terms was estimated from an assessment of the completeness of TB notification data in 1997 and (ii) trends in TB incidence in years before and after 1997 have been estimated from TB notification data, mostly using the assumption that trends in TB notifications (all forms of case) are the same as trends in TB incidence.
Method 5 can be used every year provided that both (i) mortality data are timely, high quality and complete, or are timely and of consistent quality and completeness and (ii) the case fatality rate is constant.

Method 6 can be used to measure trends in TB incidence in one of two major ways: (i) it can be used at regular intervals or (ii) a one-time estimate can be combined with an estimate of the trend in TB incidence. Trends in TB incidence can be estimated from TB notification data, from a series of ARI surveys, from a series of surveys of the prevalence of TB disease, from a series of TB mortality data, or (in theory but never in practice) a series of prospective cohort studies. For TB notifications to be considered a direct proxy of trends in TB incidence, programmatic/health system efforts to find and treat TB cases must remain constant during the years for which TB notification data are used as a direct proxy for trends in TB incidence (and likely for 1–2 years before the start of the series of notification data). If programmatic/health system efforts are not constant - for example, there are changes (either positive or negative) in the quality and coverage of TB diagnostic and treatment services, and/or there are changes (either positive or negative) in the range of interventions being used to find and treat cases - then changes in TB notifications may not correspond well to changes in TB incidence.

The number of countries for which each of these methods was used to produce WHO estimates of TB incidence in 2006 is shown in Table 2. As this shows, in 2006 the WHO estimates of TB incidence for most countries were based on indirect estimates of the completeness of TB notification data (method 6), usually for 1997, combined with estimates of trends in TB incidence using trends in TB notifications for years before and after 1997.

6.3 WHO policies and recommendations

The main WHO policies and related recommendations about how TB incidence should be measured (both its absolute value and trends) are summarized in Box 3. These are explained in more detail in the following subsections (sections 6.3.1 to 6.3.5).

6.3.1 Strengthening routine TB surveillance

The best method for measuring TB incidence is through a routine surveillance system that captures reliable and comprehensive data about new cases of TB (defined above as method 1). All countries should strengthen their surveillance systems (TB-specific recording and reporting systems, and/or general health information systems) until TB notifications can be considered a direct measure (close proxy) of TB incidence.

Strong foundations for effective TB surveillance already exist. Standard WHO recording and reporting forms exist and are widely used. For most NTPs, data collection, management and reporting are a core part of their activities. The UN Statistical Division as well as agencies such as the Global Fund have developed, and continue to develop, methods for assessing the quality of data. The Health Metrics Network (HMN), a partnership housed by WHO, is aiming "to increase the availability and use of timely and accurate health information by catalysing the joint funding and development of core country health information systems". This offers an opportunity to improve TB surveillance in the context of more general strengthening of health

---

4 This is because case-finding efforts can affect incident cases from previous years (that are now part of a backlog of prevalent cases).
information systems. Training courses on TB surveillance and epidemiology have been developed and are regularly organized by WHO and its collaborating centres, as well as by the International Union against TB and Lung Disease (IUATLD), USAID, Centers for Disease Control (CDC), Measure Evaluation and the TB research institute in Japan (RIT).

**Box 3. WHO policies and recommendations for measuring TB incidence**

**Policy 1.** WHO will promote the strengthening of routine TB surveillance (of cases and deaths) in all countries as the major mechanism for measuring progress in TB control.

**Policy 2.** WHO, together with other members of the Task Force, will develop a standardized framework and related tool for producing and documenting estimates of TB incidence (both its absolute value and trends). The framework will focus on analysis of the reliability and coverage of TB notification data, and will include definition of benchmarks that would need to be met for TB notification data to be considered a direct measure (close proxy) of TB incidence.

**Policy 3.** WHO, together with other members of the Task Force, will support countries to conduct systematic assessments of the reliability and coverage of TB notification data, and will use the results to produce updated estimates of TB incidence (its absolute value and trend) and the case detection rate. If appropriate, countries will be invited to request the Task Force to certify their data as providing a direct measure of TB incidence.

**Policy 4.** WHO will promote the development of vital registration (VR) systems and will support countries to use the mortality data from these systems to help cross-validate estimates of TB incidence, particularly in countries where (i) VR data are considered more reliable and complete than TB notification data and (ii) the necessary linkages between TB mortality and notification data can be made.

**Policy 5.** WHO will not encourage the use of ARI surveys to measure TB incidence.

**Recommendation 1.** All countries should strengthen their surveillance systems until TB notifications are a direct measure (close proxy) of TB incidence.

**Recommendation 2.** Countries should periodically conduct a systematic assessment of TB incidence (its absolute value and trends), using a standard framework and tool for analysing and documenting the reliability and coverage of TB notification data. This framework and tool will be developed by the Task Force. Findings should be used to produce better estimates of TB incidence (its absolute value and trends) and better estimates of the case detection rate. It will also help to identify where and how TB surveillance needs to be strengthened and where TB control needs to be improved so that the case detection rate can be increased.

**Recommendation 3.** Countries that would like their TB notification data to be certified by the Task Force as providing a direct measure of TB incidence should approach the Task Force so that a formal evaluation can be undertaken.

**Recommendation 4.** Countries with reliable and complete mortality data from vital registration systems should consider using these data to cross-validate estimates of TB incidence, especially where these data are likely to be more reliable and complete than TB notification data.

**Recommendation 5.** ARI surveys should not be used to measure TB incidence in most countries. Their use should be limited to a few countries where there is good reason to believe that ARI data can be interpreted reliably and where there is no feasible alternative approach to estimating TB incidence.

Case-based national TB information systems allow much more precise analysis of TB data than is possible with aggregated data that are compiled on a quarterly basis. For this reason, countries should develop case-based recording and reporting systems using flexible computer solutions. Ideally, these systems should be web-based to allow remote data entry and real-time data management and reporting. They should also be designed with flexibility and safety in mind, so that systems can be readily adapted when information needs change. If general health information systems provide an
adequate platform for case-based reporting of TB cases (reports are timely, reliable and complete), then for improved efficiency countries are encouraged to integrate case-based TB reporting into general health information systems.

6.3.2 Standardized and systematic assessment of TB incidence via evaluation of the reliability and coverage of TB notification data

For most countries, estimates of TB incidence published up to 2008 were based on two things (Table 2):

- an estimate of the case detection rate in 1997 i.e. the absolute number of incident cases of TB in 1997 was estimated as the number of notified TB cases in 1997 divided by the estimated proportion of TB incident cases that were notified in 1997 (method 6);
- an assumption that trends in TB notifications (of all forms of case) since 1997 were a direct proxy for trends in TB incidence. In other words, trends in TB incidence after 1997 were assumed to be the same as trends in TB notifications.

From 2009 onwards, an updated approach to measuring TB incidence (both its absolute value and trends) is needed:

- it is over ten years since the original estimates of TB incidence were made;
- the assumption that trends in TB notifications reflect trends in TB incidence is increasingly problematic, given programmatic efforts to increase case-finding;
- there are now more data and more experience of analysing TB surveillance data that can be drawn upon to produce updated estimates.

From 2009, the Task Force will focus on two approaches to better measurement of TB incidence (both its absolute value and trends). These are:

- direct measurement of TB incidence (both its absolute value and trends) from TB notification data (defined above as method 1). This is likely to be feasible in a relatively small number of countries before 2015, but it will set the standard to which all countries should aspire;
- a more systematic and standardized approach to the indirect measurement of TB incidence, combining periodic assessment of the case detection rate (i.e. the proportion of incident TB cases accounted for in TB notification data, defined as method 6 above) with analysis of the extent to which trends in TB notifications are a proxy of trends in TB incidence.\(^5\)

For both of these approaches, a common analytical framework and tool will be developed by the Task Force. The framework will be based on three things:

- a standard set of analyses of available TB notification data, designed to assess the completeness and reliability of these data as well as the extent to which TB notification data account for incident TB cases;
- analysis of available TB notification data and a standard set of variables that can influence trends in TB notifications. Depending on the results, TB notification data could be judged to be a direct proxy of trends in TB incidence, or they could be used to measure trends in TB incidence after adjustment for the influence of factors besides changes in TB incidence, or they could be judged too unreliable to provide any measure of trends in TB incidence; and

---

\(^5\)It should be noted that attempting to estimate case detection rates at sub-national level is not recommended, since such assessments require sub-national (rather than national) estimates of TB incidence.
a standard approach to estimating the number of TB cases that are not accounted for in TB notification data.

A prototype version of this framework is shown in Table 3.

For a country's TB notification data to be considered a direct measure of TB incidence, it would need to be shown that available TB notification data are reliably recording all incident cases (or missing only a negligible number of such cases). For other countries, the three sets of analyses defined above should allow a better indirect assessment of the fraction of TB cases included in TB notification data for a given year (i.e. the case detection rate), a better assessment of trends in TB incidence, and identification of what needs to be done to improve the quality and coverage of TB notification data.

An example of what the recommendations mean in practice is provided in Box 4, based on a recent analysis of TB incidence (its absolute value and trend) in Kenya.

**Standardized analysis of available TB notification data**

The first step to assessing whether TB notification data are directly measuring TB incidence (both its absolute value and trends) is to analyse available TB notification data. This analysis should include assessment of the completeness of reporting by reporting units (for example, the number of expected reports can be compared with the number of reports actually received), assessment of whether there is duplication or misclassification of data, exploration of variability geographically and over time (to check for internal consistency), and comparisons with values that are expected given existing knowledge of TB epidemiology (e.g. the fraction of pulmonary cases that are sputum smear-positive). This assessment should be used as a basis for removing duplications and misclassifications, to impute for missing data, and to make an initial assessment of the extent to which TB notifications account for all incident TB cases.

To facilitate these analyses, TB notification data should be available in electronic format (either as aggregated data or, ideally, as data for individual patients) and disaggregated by (i) time period (e.g. quarters) (ii) geographical or administrative unit (iii) case type (iv) smear results and (v) age and sex.

**Standardized analysis of the extent to which TB notification data are a proxy for trends in TB incidence**

TB notification data can be used as a direct proxy for trends in TB incidence if TB notification data are both a) reliable and b) account for all (or almost all) incident cases over the time period being considered. They may also be a good proxy for trends in TB incidence when incident cases are missing from routine TB notification data, provided that the proportion of incident TB cases being notified (the case detection rate) remains constant (or relatively constant). For the proportion of incident TB cases being notified to remain the same, certain conditions must apply. These include consistency in the completeness of reporting (e.g. all districts have reported TB

---

6 TB cases develop from a large and widely-distributed reservoir of latent infection, such that the true incidence of TB does not usually vary greatly across small areas or short time-periods (<5 years). Large geographic variation as well as large changes in notifications over short time periods may indicate that cases are being over- or under-reported in some places and time-periods.
Table 3: Framework for assessment of TB incidence (its absolute value and trends)

<table>
<thead>
<tr>
<th>Part 1: Standardized analysis of available TB notification data</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step</td>
</tr>
<tr>
<td>1. Assess completeness of reporting from lowest administrative level to national level</td>
</tr>
<tr>
<td>2. Check for duplications and misclassifications</td>
</tr>
<tr>
<td>3. Compare standard indicators across geographical areas (sub-national analysis) and over time, and compare standard indicators with commonly-observed or expected values in TB epidemiology</td>
</tr>
</tbody>
</table>

Part 2: Standardized analysis of whether trends in TB notifications are a good proxy of trends in TB incidence

<table>
<thead>
<tr>
<th>Example of variables: HIV prevalence in general population, number of health units providing TB diagnostic and treatment services, number of NTP staff, NTP funding, number of smears examined per diagnosed TB case, % cases among immigrants, prevalence of risk factors for TB</th>
</tr>
</thead>
<tbody>
<tr>
<td>4. Compare trends in TB notifications with trends in variables that can influence TB incidence and TB notifications</td>
</tr>
</tbody>
</table>

Part 3: Standardized assessment of the fraction of cases missing from routine TB notification data

<table>
<thead>
<tr>
<th>Example of variables: % cases pulmonary/extrapulmonary, % pulmonary cases that are smear-positive, age/sex distribution, % change in notifications per year, notifications per 100,000 population</th>
</tr>
</thead>
<tbody>
<tr>
<td>5. Compile/assess evidence about whether cases are being diagnosed and treated by care providers linked to NTP but not recorded in TB notification data</td>
</tr>
<tr>
<td>6. Compile/assess evidence about whether TB cases are being treated (but not notified) by providers not linked to NTP</td>
</tr>
<tr>
<td>7. Compile/assess evidence about whether cases are presenting but not being diagnosed at health care facilities</td>
</tr>
<tr>
<td>8. Compile/assess evidence about whether cases are not going to health facilities despite having access to them</td>
</tr>
<tr>
<td>9. Compile/assess evidence about whether there are cases without access to health facilities</td>
</tr>
</tbody>
</table>
notifications for all quarters of all years for which trends are being estimated), and consistency in programmatic efforts to diagnose and treat TB during the years being considered (e.g. consistency in the number of diagnostic and treatment units, consistency in the level of NTP funding and staffing, consistency in the number of sputum smears being examined per TB case diagnosed, consistency in the knowledge/practices of staff). If there are changes in any of these indicators, then trends in TB notification data may not be a good proxy for trends in TB incidence (for an example to illustrate this point, see Box 4). It is therefore important to analyse whether changes in these indicators have occurred, and whether they are related to trends in TB notifications. Where a relationship exists, trends in TB notifications need to be adjusted to allow for the influence of these factors, before using them to estimate trends in TB incidence.

**Box 4. What do the recommendations mean in practice? An example from Kenya**

The incidence of TB in Kenya was indirectly estimated from TB notification data in 1997, as part of a global effort to estimate the global epidemiological burden of TB (see Dye et al, JAMA, 1998). The estimate was based on an expert assessment that the percentage of incident smear-positive cases being notified was 57% (i.e. 57% case detection rate). Until 2006, the trend in TB incidence before and after 1997 was assumed to be the same as the trend in TB notifications (of all forms of TB case).

Kenya has experienced a generalized HIV epidemic since the early 1980s and substantial efforts to improve the quality and coverage of TB diagnosis and treatment services were made from 2001 onwards. This made it difficult to disentangle the effect of HIV (which affects TB incidence) from the effect of programme performance on TB notifications, which in turn made it difficult to estimate the trend in TB incidence. Between September 2006 and December 2007, estimates of the absolute value of TB incidence and the trend in TB incidence were jointly reviewed by WHO and the NTP. This was done in the context of new evidence and new analysis. The major new sources of evidence were (i) data on trends in HIV-positive and HIV-negative TB notifications separately (ii) a direct measure of the prevalence of HIV among TB patients (iii) a recent survey of the prevalence of HIV in the general population and (iv) evidence about how programme performance had changed during the period 1996–2006. Both (i) and (ii) became available following the introduction of provider-initiated HIV testing for TB patients in 2005. Evidence about programme performance during the period 1996–2006 was compiled during 2007. The four principal indicators used were: the number of health units where TB diagnosis was available, the number of health units where TB treatment was available, the number of NTP staff at national, provincial and district level, and NTP funding. For all four of these indicators, there was a clear relationship with trends in TB notifications from 2001 to 2006, while HIV-related data suggested that the epidemic peaked around 2000 and had not caused any increase in TB incidence from 2001 to 2006. In combination, these new data provided strong evidence that the increase in TB notifications after 2001 was due to programmatic improvements (and not increases in TB incidence). This led to a downward revision in the estimate of TB incidence in 2006, an adjustment of the estimated trend in TB incidence, and an upward revision in the estimated case detection rate (to 70%). The original estimate of TB incidence (and case detection) in 1997 was left unchanged.

To allow reliable measurement of trends in TB incidence from 2007 onwards, maintaining high rates of HIV testing for TB patients is essential. This will allow trends in HIV-positive and HIV-negative TB notifications to be separated. Trends in HIV-negative TB notifications can be used to measure changes in case-finding. Comparison of trends in HIV-positive and HIV-negative TB notifications can be used to assess the impact of HIV on TB incidence.

The NTP has not yet undertaken a thorough investigation of the reliability and coverage of its TB notification data, but this could be used to further improve estimates of TB incidence (in absolute terms). Efforts to strengthen the routine surveillance system, including the introduction of new recording and reporting forms and expanded use of electronic recording and reporting systems, have begun.

For further details, see Mansoer J et al.63

As part of the standardized framework and tool that WHO will develop with other members of the Task Force, a minimum or "essential" set of analyses of patterns and
trends in TB notifications will be defined. This will include analysis of changes in reporting and their association with changes in TB notifications, and analysis of changes in a core set of indicators of programme performance (e.g. number of diagnostic and treatment units, number of staff, number of smears examined per diagnosed TB case, funding) and their association with changes in TB notifications. Results will be used to judge whether TB notifications can be used to reliably measure trends in TB incidence, whether they can be used to measure trends in TB incidence after adjustment for factors besides TB incidence that affect TB notifications, or whether they are too unreliable to assess trends in TB incidence.

**Standardized assessment of the fraction of cases being missed by routine TB notification data, based on the "Onion" model**

Analysis of available TB notification data is an essential component of any assessment of TB incidence and trends in TB incidence. However, on its own it is not sufficient to estimate TB incidence in absolute terms, because it will not identify how many TB cases exist but are not accounted for in TB notification data. A framework that can be used to understand where and why incident TB cases might not be accounted for in TB notification data, to investigate and quantify the proportion of incident TB cases that are captured in TB notification data, and to identify the kind of

**Figure 1. The “onion” model: a framework for assessing the fraction of TB cases accounted for in TB notification data, and how this fraction can be increased**
Table 4: Examples of methods that could be used to assess how many TB cases are missing from TB notification data

<table>
<thead>
<tr>
<th>Possible reason for cases to be missing from TB notification data</th>
<th>Examples of methods that could be used to directly measure how many TB cases are missing from TB notification data</th>
<th>Examples of published studies</th>
<th>Examples of supporting evidence that could be used</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cases diagnosed by NTP but not recorded in notification data (Ring 2)</td>
<td>Operational research can be used to study the number of cases that are missing from TB notification data. These studies typically involve prospectively collecting data from places where TB cases may be (i) diagnosed but not notified (ii) seeking care but not being diagnosed and (iii) experiencing symptoms but not seeking care.</td>
<td>Botha E et al (S. Africa)</td>
<td>Drugs sales in the private sector</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Health expenditures in private/NGO sectors, out-of-pocket expenditures</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Number of health facilities/private practitioners and proportion that are not collaborating with the NTP</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Prescriptions in pharmacies</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Regulations regarding prescribing and availability of drugs and their application in practice</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Knowledge and use of the international standards for TB care</td>
</tr>
<tr>
<td>Cases diagnosed by non-NTP providers that are not notified (Ring 3)</td>
<td>To assess the number of cases whose diagnosis is being missed at health care facilities and to assess the number of cases that are being correctly diagnosed and treated but not notified, a common approach is to introduce study registers at health facilities (including laboratories), in which TB suspects and TB cases are listed. These lists can then be compared with lists of notified cases. If at least 2 (but preferably 3 or more) lists can be generated, it may be possible to use capture-recapture methods to estimate total incident cases (i.e. to estimate not only cases that are missing from notifications, but also to estimate the number of cases that are missing from all lists i.e. cases that are not in contact with health facilities at all).</td>
<td>Miglioiri et al (Italy), Maung et al, (Myanmar), Lonnroth et al (Viet Nam), Ambe et al (India), Arora et al (India), Dewan et al (India)</td>
<td>Knowledge/attitudes/practices of health staff</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Suspect management practices</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Slides examined per TB suspect</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>% laboratories with satisfactory performance (based on EQA)</td>
</tr>
<tr>
<td>Cases presenting to health facilities that are not diagnosed (Ring 4)</td>
<td></td>
<td></td>
<td>Data on population knowledge, attitudes and practice (KAP) from TB-related KAP surveys</td>
</tr>
<tr>
<td>Cases that have access to health services but do not seek care (Ring 5)</td>
<td></td>
<td></td>
<td>Population access to health services e.g. % population living within a certain distance of a health facility</td>
</tr>
<tr>
<td>Cases that do not have access to health services (Ring 6)</td>
<td></td>
<td></td>
<td>Number of laboratories doing smear microscopy per 100 000 population</td>
</tr>
<tr>
<td>All reasons listed above</td>
<td></td>
<td></td>
<td>Number of nurses and doctors per 100 000 population compared with international norms of what is required</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Data from major household/demographic surveys</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Vital registration data showing what proportion of TB deaths never accessed TB diagnosis and treatment</td>
</tr>
</tbody>
</table>

Prevalence survey from Myanmar | Prevalence of TB disease survey in which questions about health-seeking behaviour and contact with health services are asked. See also section 7. |
programmatic or health system interventions that might be required to increase the fraction of incident TB cases being recorded in TB notification data, is shown in Figure 1. This framework was first presented to the international TB community in 2002, and has been termed the "onion" model.

In the onion model, only TB cases in the first ring are found in TB notification data. The relative size of rings 2 to 6 determines the proportion of TB incident cases being accounted for in TB notification data. Although conceptually simple, quantification of the fraction of TB cases that are missing from TB notification data (Rings 2 to 6) is challenging. Examples of the methods and supporting evidence that could be used to assess how many cases exist in rings 2 to 6 are provided in Table 4.

**Support to countries to use the framework and tool**

A tool to help translate the framework described above into practice will be developed. Countries will then be supported to use it via mechanisms such as workshops and country missions. Besides allowing better measurement of TB incidence and better understanding of where TB surveillance and TB control as a whole need to be strengthened, this process should help to strengthen national and local capacity in monitoring and evaluation (e.g. data management, data analysis, operational research and writing of reports and papers).

**6.3.3 Certification of TB notification data**

Countries with TB notification data that can be shown to be reliably accounting for all (or close to all) incident TB cases using the framework described in section 6.3.2 will be encouraged to approach the Task Force so that their data can be certified as a direct measure of TB incidence. Certification will be based on a formal assessment by Task Force members and other experts.

**6.3.4 Vital registration data**

In addition to analysis of TB notification data, another form of routine surveillance data - vital registration data that include data on deaths from TB - can help to cross-validate estimates of TB incidence and trends in TB incidence. Countries with vital registration systems that include mortality data of high reliability and coverage should consider using these data to cross-validate estimates of TB incidence (described as method 5 above). There are three examples of countries for which this has already been done: Brazil, Mexico and South Africa. For other countries considering the use of vital registration data for this purpose, however, two points should be highlighted: (i) vital registration data should be considered more timely, reliable and complete than TB notification data and (ii) it should be feasible to link mortality records with TB notification records.

**6.3.5 ARI surveys and prevalence of TB disease surveys**

ARI surveys to measure TB incidence are not recommended in most countries. This is because survey results are usually difficult or impossible to interpret. The use of ARI surveys should be limited to a few countries where there is good reason to believe that data can be interpreted reliably and where there is no feasible alternative approach to estimating TB incidence. Further details about ARI surveys are provided in Annex 3.
Surveys of the prevalence of TB disease are likely to provide an additional source of evidence about TB incidence. However, these surveys will probably be undertaken in a relatively small number of countries. Moreover, even when data on disease prevalence are available, there will be considerable uncertainty in any estimates of TB incidence, both because of the wide confidence intervals on estimates of TB prevalence and uncertainty about the duration of TB disease. Further details are provided in section 7.

6.4 Guidance material and tools

At the time of writing (mid-2008), the standardized framework and tool for evaluation of TB notification data, including definition of the benchmarks to be used to determine whether or not a country's TB notification data can be certified as directly measuring TB incidence, remains under development. A first draft of a set of standard methods and a prototype tool were developed for a workshop with the theme "Analysing progress towards the 2015 targets for TB control in Latin America: the role of routine surveillance data". The framework and tool will be further developed and piloted during 2009, including in a workshop with countries from the European region and in meetings of the Task Force's subgroup on routine surveillance (see Annex 1). They will become part of WHO guidance material after endorsement by the Task Force.

Studies in Kenya, Morocco and Viet Nam provide three recent examples of in-depth studies of TB incidence. In Kenya (see Box 4), data on HIV prevalence in the general population and among TB patients specifically, data on programmatic efforts to increase case-finding and a series of TB notification data were used to re-estimate TB incidence in absolute terms as well as its trend over time. In Morocco, an analysis of trends in TB notifications from 1996 to 2005 indicated that TB incidence rates had fallen more slowly in men than in women, and that TB incidence had fallen more slowly than expected. In Viet Nam, the lack of any decline in TB notification rates could not be explained by increasing efforts to diagnose TB, suggesting that stable TB notification rates are not concealing reductions in TB incidence.

Such reductions in TB incidence had been expected given that the WHA targets of a 70% case detection rate and an 85% treatment success rate had been assessed to have been achieved. There are two published studies that have measured TB incidence through a population cohort study; one from the Republic of Korea and one from southern India. Neither study provides evidence of TB incidence at national level.

Guidelines on ARI surveys have been produced by KNCV and the National Tuberculosis Institute in Bangalore, India, based on previous guidelines and reviews.

6.5 Technical and financial support

Technical support is available from a variety of agencies, including those that are represented on the Task Force. Sources of financial support include national or local budgets, the Global Fund (which recommends that 7% of a grant should be for monitoring and evaluation), and other donors that are committed to supporting monitoring and evaluation.

Subsequently, a survey of the prevalence of disease has suggested that TB incidence is higher than previously thought, and that the case detection rate is lower than previously thought.
7. WHO policy and recommendations for measuring TB prevalence

7.1 Definition of TB prevalence

The prevalence of TB is the number of TB cases at a given point in time. It is usually reported as the total number of prevalent cases, or as the number of prevalent cases for a given unit of population (e.g. number of cases per 100 000 population). Prevalence determines the risk of TB infection in a community i.e. how much transmission is occurring. The prevalence of TB is (approximately) the incidence of TB multiplied by the average duration of disease. Improved case-finding and treatment both shorten the duration of disease, so prevalence responds more rapidly than incidence to changes in TB control. Periodic assessment of the prevalence of TB disease can therefore be more useful for measuring the short-term impact of TB control (e.g. within five years) than are efforts to measure changes in TB incidence.

Box 5: Surveys of the prevalence of TB disease: important facts and limitations

1. **Extrapulmonary cases.** Typically, surveys do not look for or identify extrapulmonary cases. The diagnostic methods needed to diagnose such cases are invasive and it would be difficult to apply them in the context of a population-based survey. For this reason, investigations to diagnose extrapulmonary TB are not recommended in the WHO guidelines on surveys of the prevalence of TB disease.

2. **Children.** Typically, surveys do not try to identify TB cases among children. This is because the current technology used for screening is too invasive (e.g. venepuncture, radiation and gastric tubing) for use in healthy children. Alternative ways to measure the epidemiological burden of TB in children are to strengthen surveillance and to conduct systematic contact tracing. To measure the number of TB cases in children in population-based surveys, better diagnostic tools are needed.

3. **Pulmonary cases not confirmed by bacteriology.** Cases of TB that are not confirmed by smear or culture will not be identified by a survey of the prevalence of TB disease. Diagnosis of such cases requires patient follow-up that is not part of survey procedures.

4. **HIV-positive cases.** The recommended screening strategy includes culture examinations, so culture-positive (but smear-negative) TB among HIV-positive individuals will be diagnosed.

5. **Estimates of the total prevalence of TB disease in the community.** Uncertainties about the proportion of TB cases accounted for by children and extrapulmonary adult cases mean that even with an (imprecise) estimate of the prevalence of smear and/or culture-positive TB, it can be difficult to estimate the total prevalence of TB in the community.

6. **Adding TB prevalence surveys to existing surveys or survey platforms.** This is difficult if not impossible. Surveys for other diseases are usually designed to estimate the prevalence of diseases and conditions for which the prevalence is much higher than the prevalence of TB. This means that the size of the population studied (required sample size) is much smaller than the 100 000 to 200 000 people who need to be included in a survey of the prevalence of TB disease. Adding a TB-prevalence component to an existing Demographic and Health Survey (DHS), for example, would overwhelm the DHS. In addition, TB surveys depend on mobile X-rays, radiographers to read the X-rays, facilities for collecting and transporting sputa, and laboratories to process the samples; in contrast most surveys depend mainly (or only) on the results of questionnaires. It is therefore more likely that surveys of other diseases and health conditions would be added to surveys of the prevalence of TB disease, rather than vice versa.

7. **Using surveys to assess the coverage of routine TB notification data.** Surveys can be useful for identifying how many cases of active TB have been to health care facilities and how many are accounted for in TB notification data. This information will help to estimate the fraction of incident TB cases being missed from TB notification data, which in turn will help to produce better estimates of TB incidence and the case detection rate (see also section 6).
7.2 Methods for measuring TB prevalence

There are two methods for estimating the prevalence of TB (see Annex 4 for further details). The first is direct measurement using a cross-sectional population-based survey. Such surveys typically require sample sizes of around 200 000 people (if the estimate of TB prevalence prior to the survey is around the global average of 100 per 100 000 population), and implementation is expensive (typically US$ 1–2 million per survey) and logistically challenging. The second is indirect measurement, with TB prevalence estimated as TB incidence multiplied by the average duration of disease. In WHO estimates published up to 2008, estimates of TB prevalence in all but 13 countries were based on multiplying estimates of TB incidence by the estimated average duration of disease.

Changes in TB prevalence over time are best measured by implementing at least two (but preferably more) surveys at large enough intervals. Alternatively, changes in TB prevalence over time can be estimated from changes in TB incidence and the average duration of disease. In other words, TB prevalence in a given year can be estimated as incidence multiplied by the average duration of disease (in years).

Besides the challenges of costs and logistics, important basic facts and limitations about surveys of the prevalence of TB disease are highlighted in Box 5

7.3 WHO policies and recommendations

WHO policies and recommendations about the measurement of TB prevalence are summarized in Box 6. These are explained in the following two subsections. Further details are also available in the report of the December 2007 meeting of the Task Force.

Box 6. WHO policy and recommendations for measuring TB prevalence

<table>
<thead>
<tr>
<th>Policy 1.</th>
<th>WHO will promote the implementation of national surveys of the prevalence of TB disease in 21 global focus countries that have been identified by the Task Force, and will assist these countries to access the necessary technical and financial support.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendation 1.</td>
<td>There are 21 countries that should implement at least one survey of the prevalence of TB disease between 2008 and 2015. Some of the countries without prior survey data should carry out two surveys during the period 2008–2015.</td>
</tr>
<tr>
<td>Recommendation 2.</td>
<td>Prevalence of TB disease surveys should be implemented according to the WHO guidelines published in 2007.</td>
</tr>
<tr>
<td>Recommendation 3.</td>
<td>In a survey of the prevalence of TB disease, the sampled population should at a minimum be screened using both X-rays and a questionnaire about symptoms of TB. This is Strategy 3 in the WHO guidelines. All sampled individuals should be asked about their health-seeking behaviour.</td>
</tr>
<tr>
<td>Recommendation 4.</td>
<td>Any TB cases identified during a survey of the prevalence of TB disease should be tested for HIV according to national policy and standard practice. Questions about other diseases and risk factors can be included if they will not compromise the TB component of the survey.</td>
</tr>
<tr>
<td>Recommendation 5.</td>
<td>All countries that implement prevalence of TB disease surveys should also promote the strengthening of routine TB surveillance, and should use surveys as an opportunity to strengthen national and local capacity in monitoring and evaluation.</td>
</tr>
<tr>
<td>Recommendation 6.</td>
<td>For countries that do not implement prevalence of TB disease surveys, TB prevalence should be indirectly estimated using methods recommended by the Task Force.</td>
</tr>
</tbody>
</table>
7.3.1 Countries where national population-based surveys of the prevalence of TB disease are recommended

The Task Force has identified 21 countries (termed global focus countries) where surveys of the prevalence of TB disease are strongly recommended. These 21 countries are listed in Table 5. These countries will be given particular attention and support by the Task Force, including provision of training to survey principal and co-investigators, and the matching of at least one technical partner to each country (see also Annex 6). The criteria that were used to select these countries are shown in Table 6. To be selected, a country had to meet at least one of the four sets of criteria.

Table 5. The 21 countries in which the implementation of at least one survey of the prevalence of TB disease between 2008 and 2015 is strongly recommended

<table>
<thead>
<tr>
<th>Country</th>
<th>Criteria met</th>
<th>Region*</th>
<th>High-Burden?</th>
<th>Estimated TB SS** prevalence in 2005 (per 100,000 population)</th>
<th>Survey data from 2008 or earlier exist</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>2,4</td>
<td>AFR high HIV</td>
<td>Yes</td>
<td>154.2</td>
<td>No</td>
</tr>
<tr>
<td>Malawi</td>
<td>1,2,3,4</td>
<td>AFR high HIV</td>
<td>No</td>
<td>239.1</td>
<td>No</td>
</tr>
<tr>
<td>Mozambique</td>
<td>1,2,3</td>
<td>AFR high HIV</td>
<td>Yes</td>
<td>244.7</td>
<td>No</td>
</tr>
<tr>
<td>Nigeria</td>
<td>1,2,3,4</td>
<td>AFR high HIV</td>
<td>Yes</td>
<td>226.3</td>
<td>No</td>
</tr>
<tr>
<td>South Africa</td>
<td>2,3</td>
<td>AFR high HIV</td>
<td>Yes</td>
<td>395.8</td>
<td>No</td>
</tr>
<tr>
<td>Uganda</td>
<td>1,2,3,4</td>
<td>AFR high HIV</td>
<td>Yes</td>
<td>237.2</td>
<td>No</td>
</tr>
<tr>
<td>UR Tanzania</td>
<td>1,2,3,4</td>
<td>AFR high HIV</td>
<td>Yes</td>
<td>204.7</td>
<td>No</td>
</tr>
<tr>
<td>Zambia</td>
<td>2,3</td>
<td>AFR high HIV</td>
<td>No</td>
<td>291.4</td>
<td>No</td>
</tr>
<tr>
<td>Ghana</td>
<td>1,2</td>
<td>AFR low HIV</td>
<td>No</td>
<td>158.3</td>
<td>No</td>
</tr>
<tr>
<td>Mali</td>
<td>1,2,3,4</td>
<td>AFR low HIV</td>
<td>No</td>
<td>243.1</td>
<td>No</td>
</tr>
<tr>
<td>Rwanda</td>
<td>1,2,3</td>
<td>AFR low HIV</td>
<td>No</td>
<td>278.1</td>
<td>No</td>
</tr>
<tr>
<td>Sierra Leone</td>
<td>1,2,3</td>
<td>AFR low HIV</td>
<td>No</td>
<td>416.0</td>
<td>No</td>
</tr>
<tr>
<td>Pakistan</td>
<td>1,4</td>
<td>EMR</td>
<td>Yes</td>
<td>132.5</td>
<td>Yes</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>4</td>
<td>SEA</td>
<td>Yes</td>
<td>142.1</td>
<td>Yes</td>
</tr>
<tr>
<td>Indonesia</td>
<td>4</td>
<td>SEA</td>
<td>Yes</td>
<td>106.6</td>
<td>Yes</td>
</tr>
<tr>
<td>Myanmar</td>
<td>4</td>
<td>SEA</td>
<td>Yes</td>
<td>75.6</td>
<td>Yes</td>
</tr>
<tr>
<td>Thailand</td>
<td>2,4</td>
<td>SEA</td>
<td>Yes</td>
<td>84.4</td>
<td>Yes</td>
</tr>
<tr>
<td>Cambodia</td>
<td>2,3</td>
<td>WPR</td>
<td>Yes</td>
<td>267.3</td>
<td>Yes</td>
</tr>
<tr>
<td>China</td>
<td>4</td>
<td>WPR</td>
<td>Yes</td>
<td>89.4</td>
<td>Yes</td>
</tr>
<tr>
<td>Philippines</td>
<td>4</td>
<td>WPR</td>
<td>Yes</td>
<td>165.9</td>
<td>Yes</td>
</tr>
<tr>
<td>Viet Nam</td>
<td>4</td>
<td>WPR</td>
<td>Yes</td>
<td>89.5</td>
<td>Yes</td>
</tr>
</tbody>
</table>

*AFR high HIV = African countries with high HIV prevalence. AFR low HIV = African countries with low HIV prevalence. SEA = South-East Asia. WPR = Western Pacific.

A further 32 countries that met at least one of the four sets of criteria shown in Table 6 are listed in Annex 5. These countries may also decide to implement surveys for the purposes of producing better national estimates of the epidemiological burden of TB, but will not be the focus of Task Force efforts to support survey implementation. The rationale for selection of the 21 countries shown in Table 5 was as follows:

- surveys of the prevalence of TB disease are expensive and logistically difficult to implement, and providing the necessary technical and financial support to all
of the 53 countries\(^8\) that met at least one of the four sets of criteria for undertaking a survey would be challenging;

- taken together, the 53 countries represent about two-thirds of total TB cases globally. For certain regions, the countries that met the criteria accounted for a very high percentage of total TB cases: 98% of total cases in the African Region, 97% of total cases in the Western Pacific Region, and 75% of total cases in the Eastern Mediterranean Region. Undertaking surveys in a smaller number of countries is sufficient for the purposes of measuring progress at global and regional level, and represents a much more efficient use of available technical and financial resources;
- all of the 22 high-burden countries that met at least one of the four sets of criteria in Table 6 were included;
- in combination, the 21 countries account for a substantial share of the regional number of TB cases in the four WHO regions where routine surveillance systems are weakest (i.e. excluding the European Region and the Region of the Americas);
- in Africa, the selected countries represent central, west, east and southern Africa, and include Anglophone, Francophone and Lusophone countries.

Table 6. Criteria used to select countries that are candidates for implementing surveys of the prevalence of TB disease during the period up to 2015

<table>
<thead>
<tr>
<th>Group 1 →</th>
<th>Criteria</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimated smear-positive TB prevalence rate in 2006 ≥ 100 per 100,000 population and 2. Accounts for ≥ 1% of the estimated total number of smear-positive TB cases globally for 2006 and 3. Case detection rate (CDR) in 2005 ≤ 50% or &gt;100%</td>
<td>Major contribution to global burden of TB Sample size small enough to make surveys feasible in terms of cost and logistics Excludes countries whose contribution to the global burden of TB is insignificant for the purposes of global and regional assessments of burden and impact CDR ≤ 50% or &gt; 100% indicate weak reporting systems and problematic TB estimates, respectively</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 2 →</th>
<th>Criteria</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimated smear-positive TB prevalence rate in 2006 ≥ 70 per 100,000 population and 2. Accounts for ≥ 1% of the estimated total number of smear-positive TB cases globally in 2006 and 3. Estimated HIV prevalence rate in the adult population (15 to 49 years) in 2005 ≥ 1%</td>
<td>Less stringent criteria for the TB prevalence rate, but incorporates countries with high HIV prevalence and therefore where there is potential for a rapid increase in TB incidence and prevalence rates</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 3 →</th>
<th>Criteria</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Estimated smear-positive TB prevalence rate in 2006 ≥ 200 per 100,000 population and 2. Accounts for ≥ 0.5% of the estimated total number of smear-positive TB cases globally in 2006</td>
<td>Less stringent criteria for country's contribution to global burden of disease, but incorporates countries with particularly high TB prevalence rates</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Group 4 →</th>
<th>Criteria</th>
<th>Explanation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Survey implemented between 2000 and 2007 or 2. Survey planned before 2010</td>
<td>Prior survey data allow monitoring of trends High motivation of NTP to conduct a survey</td>
<td></td>
</tr>
</tbody>
</table>

Sources of data used to apply the criteria shown in Table 1 were 1) Global Tuberculosis Control: Surveillance, Planning, Financing. WHO 2007; 2) WHO TB data collection form for 2007; 3) Report on the global AIDS epidemic, UNAIDS/WHO, 2006.

---

\(^8\) Of the 53 countries, 33 were in the African Region, 7 in the Western Pacific Region, 5 in the South-East Asia Region, 4 in the Eastern Mediterranean Region, 3 in the European Region and 1 in the Region of the Americas.
A single prevalence of TB disease survey in countries with no prior survey data will be useful for helping to improve current estimates of the epidemiological burden of TB as well as providing a baseline for any surveys conducted after 2015. However, assessing changes in the prevalence of TB requires surveys to be carried out on (at least) two separate occasions, several years apart. For this reason, some of the African countries that are among the list of 21 priority countries will need to conduct two prevalence of disease surveys between 2008 and 2015 (Table 5).

For countries that implement prevalence of TB disease surveys, it is important that these surveys do not distract from the need to simultaneously establish mechanisms to strengthen routine surveillance of TB cases and deaths.

### 7.3.2 Recommended methods to be used in surveys of the prevalence of TB disease

Countries that implement surveys of the prevalence of TB disease should follow the guidelines published by WHO in 2007, which were universally endorsed by the Task Force.\(^9\)

In some cases, the guidelines present options rather than a strong recommendation about which of these options should be implemented. The Task Force discussed these options, as well as several other important issues, during its December 2007 meeting. The following specific recommendations and advice were agreed upon:

1. **Screening strategy.** The minimum screening strategy that should be used in surveys of the prevalence of TB disease is Strategy 3 in the WHO guidelines. In this strategy, the sampled population is screened using X-rays and a questionnaire about symptoms. All TB suspects are then asked to provide two sputum samples for smear microscopy and culture examination. While screening all of the sampled population with an X-ray considerably increases the cost of a survey, it reduces the workload of microscopy laboratories and is thought to provide the best possible estimates.

2. **Sub-national surveys and estimates.** In general, countries should not attempt to estimate TB prevalence at sub-national as well as national level. The production of estimates at sub-national level greatly increases the sample size that is required, which in turn increases costs. Estimates obtained at national level are satisfactory for the purposes of measuring progress towards the 2015 global targets. If countries have a particular interest in comparing specific groups of the population or geographical areas (e.g. urban vs. rural, coast vs. inland), then the strata should be clearly pre-defined and the sample size calculated accordingly. Stratification may also be useful for improving the precision of the national estimate of TB prevalence produced by a survey.

3. **HIV testing.** All TB cases identified in a survey should be offered HIV testing according to national policy and standard practice. In general, however, HIV testing should not be undertaken for all of the sampled population. Reasons include: (i) providing HIV screening for all of the sampled population is logistically difficult; (ii) obtaining informed consent from all participants may reduce the percentage of people willing to participate in the survey; and (iii) it may be difficult to ensure that results are provided to all those who are tested.
4. **Combining surveys of the prevalence of TB disease with surveys for other diseases.** It is difficult if not impossible to add surveys of the prevalence of TB disease to other surveys or survey platforms (see Box 5). It is more feasible to add surveys of other diseases to surveys of the prevalence of TB disease. However, while adding other surveys to surveys of the prevalence of TB disease allows collection of additional data, the disadvantage is that it will increase the time required for interviews as well as the overall complexity of the survey (e.g. training needs, logistics and analysis). The diseases and conditions that it is most suitable to survey in combination with TB have the following characteristics: (i) they occur mostly in the adult population (≥15 years); (ii) the prevalence is higher than the prevalence of TB; (iii) screening methods do not diverge, or diverge only marginally, from those used for TB; and (iv) they help with the differential diagnosis of TB e.g. chronic respiratory diseases. Collection of data beyond that needed for a survey of TB prevalence should only be attempted if it will not compromise the quality of the basic TB survey data.

5. **Collection of data on health-seeking behaviour.** Collection of data about health-seeking behaviour and the extent to which identified cases had already had contact with health services is strongly recommended. Results can be used to assess how many cases have not had contact with health services, the number that had not been diagnosed despite visiting health services, and the number of cases that had not been notified due to health-care providers not being linked to the NTP. Such findings will help to identify the fraction of cases likely to be included in TB notification data, reasons for lack of access to TB care and the absence of notification, and to develop interventions that will accelerate progress in TB control (see also section 6.3.3). Recent examples of surveys that have included the collection of such data are those undertaken in the Philippines and Myanmar.

6. **Collection of data on socio-economic status and risk factors for TB.** The collection of data on socio-economic status and risk factors for TB should be carefully considered. It is essential that the time and effort required to collect such data do not compromise the quality of the basic survey data. Recent experience from a survey of the prevalence of TB disease in the Philippines will help to inform decisions about whether or not similar information should be collected in future surveys.

7. **Drug susceptibility testing (DST).** Positive cultures from a survey of the prevalence of TB disease should be tested for drug susceptibility, provided that patients who are found to have drug resistance will be able to access appropriate treatment. Including DST in a survey of the prevalence of TB disease is especially relevant for countries that have not yet implemented a representative drug resistance survey. However, it should not be a substitute for conducting such a survey. The overall number of drug-resistant cases found in a survey of the prevalence of TB disease (which typically identifies a total of around 100–300 TB cases) is likely to be too small to provide precise estimates of the prevalence of drug-resistant TB (although the data will help to calculate the sample size needed for a drug resistance survey and may provide an initial estimate of the prevalence of drug resistance). Further discussion of DST in the context of surveys of the prevalence of TB disease is provided in Annex 11 of the WHO guidelines.

---

9 This means that not many questions would need to be added to the questionnaire and the main diagnostic tool required would be an X-ray machine.
8. **Combining prevalence of TB disease surveys with ARI surveys.** In the "Model DOTS Project" in south India, surveys of disease and infection conducted at the same time have allowed calibration of results and estimation of the national prevalence of TB disease by extrapolation from sub-national data. However, since the results of ARI surveys are usually difficult to interpret, combining ARI surveys with surveys of the prevalence of TB disease is not encouraged (see also section 6).

Important lessons have also been learned from recently implemented surveys in Asia. Some key "Dos" and "Don'ts" related to surveys of the prevalence of TB disease are highlighted in Box 7.

**Box 7. What to do and what to avoid when designing and implementing a disease prevalence survey** *(NB this Box is due to be updated by Task Force members who have supported recent surveys)*

<table>
<thead>
<tr>
<th>Do</th>
<th>Don'ts</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Follow the WHO guidelines on disease prevalence surveys published in 2007.</td>
<td>1. Don't underestimate the sample size required.</td>
</tr>
<tr>
<td>2. Make sure that the sample size calculations are done correctly, allowing for the design effect associated with cluster sampling as well as the likely survey participation rate.</td>
<td>2. Don't underestimate the likely budget requirement.</td>
</tr>
<tr>
<td>3. When making comparisons between surveys, strike a balance between the need for similarity in the methods used and the value of making use of newer and better screening methods and diagnostic tools.</td>
<td>3. Don't neglect the importance of strengthening the routine TB surveillance system just because a prevalence survey is being implemented.</td>
</tr>
<tr>
<td>4. Remember to adjust survey design and sample size calculations if one of the survey objectives is to compare prevalence among different geographic areas and/or socioeconomic groups.</td>
<td>4. Don't make the mistake of thinking that a prevalence survey will allow a straightforward re-estimation of the case detection rate. The denominator of the case detection rate is incidence, not prevalence. Converting prevalence estimates to estimates of incidence requires assumptions about the duration of disease; there will also be uncertainty about the true incidence rate because of uncertainty in the estimate of prevalence.</td>
</tr>
<tr>
<td>5. Consult recent survey protocols known to be of high-quality as well as technical experts to ensure the survey is designed correctly.</td>
<td>5. Don't make the mistake of thinking that a prevalence survey will allow a straightforward re-estimation of the case detection rate. The denominator of the case detection rate is incidence, not prevalence. Converting prevalence estimates to estimates of incidence requires assumptions about the duration of disease; there will also be uncertainty about the true incidence rate because of uncertainty in the estimate of prevalence.</td>
</tr>
<tr>
<td>6. Pay attention to data management and analysis at the beginning, rather than leaving it till the survey has been completed.</td>
<td>6. Don't underestimate the likely budget requirement.</td>
</tr>
<tr>
<td>7. Use available budget templates to estimate the funding required.</td>
<td>7. Don't neglect the importance of strengthening the routine TB surveillance system just because a prevalence survey is being implemented.</td>
</tr>
<tr>
<td>8. Use the most recent population count or estimate available in the country to calculate sampling weights and prevalence rates.</td>
<td>8. Don't make the mistake of thinking that a prevalence survey will allow a straightforward re-estimation of the case detection rate. The denominator of the case detection rate is incidence, not prevalence. Converting prevalence estimates to estimates of incidence requires assumptions about the duration of disease; there will also be uncertainty about the true incidence rate because of uncertainty in the estimate of prevalence.</td>
</tr>
<tr>
<td>9. Use appropriate statistical methods to analyse the data, accounting for sampling design and missing data, and seek appropriate statistical expertise from the outset.</td>
<td>9. Don't use inappropriate statistical methods to analyse the data, accounting for sampling design and missing data, and seek inappropriate statistical expertise from the outset.</td>
</tr>
</tbody>
</table>

---

**7.4 Guidance material and tools**

Guidelines on the design, implementation and analysis of prevalence of TB disease surveys were published by WHO in 2007. They were developed over two years (2006 and 2007) as a collaborative effort by several technical agencies, with coordination provided by the WHO Western Pacific Regional Office (WPRO). Several of the individuals who contributed chapters are members of the Task Force.

The guidelines have 18 chapters. In brief:

- Chapter 1 explains the purpose of surveys of the prevalence of TB disease, places their importance in the context of the available methods for estimating the epidemiological burden of TB and progress towards the MDG and Stop TB
Partnership targets, and summarizes the criteria that should be used to decide whether or not a survey should be implemented;

- Chapter 2 defines the objectives of a survey of the prevalence of TB disease;
- Chapters 3–7 describe the methodology to be used, covering sampling methods, screening strategies, diagnostic methods and case definitions;
- Chapter 8 discusses what measures of socioeconomic status and exposure to risk factors could be included in a survey of the prevalence of TB disease. Specific advice on how to measure selected risk factors is provided in an annex;
- Chapter 9 identifies the ethical issues that need to be considered when carrying out a survey;
- Chapters 10–14 cover the organizational aspects of survey implementation, and highlight how quality assurance and safety practices need to be applied to every component of measurement;
- Chapters 15–17 explain the critical role of data management, data analysis and reporting of results, and how their quality can be ensured;
- Chapter 18 gives a brief introduction to budgeting for a survey of the prevalence of TB disease.

The guidelines on disease prevalence surveys are also available in the form of a series of papers, which were published in a special issue of the International Journal of TB and Lung Disease in 2008. These papers include some updates to the guidelines, based on feedback and experience since the guidelines were published.

Survey protocols from recent high-quality surveys also provide useful guidance material. A particularly good example is the protocol developed for the survey that was implemented in Cambodia in 2002.

There are several published studies that document the results of surveys of the prevalence of disease. These include:

- national surveys implemented at five-year intervals from 1965 to 1995 in the Republic of Korea, which showed substantial reductions in parallel with improved treatment results and better case-management practices;
- national surveys conducted in China and Indonesia (China 1990 and 2000, Indonesia early 1980s and 2004), both of which were able to demonstrate large reductions in TB prevalence;
- studies undertaken in southern India over a study period of about 30 years;
- surveys in the Philippines conducted in 1983 and 1997, which documented a negligible decline in bacillary disease that was attributed, at least in part, to an inconsistent supply of anti-TB drugs and low-quality of treatment during this period.

### 7.5 Technical and financial support

A list of technical agencies that have supported recent surveys and that will support future surveys is provided in Annex 6. At the time of writing, the major source of financial support for the 21 global focus countries was the Global Fund. In 2008, USAID's TB Control Assistance Program (TBCAPz) committed funding for the survey planned in Pakistan.

Reference will be added once the publication date is known.
8. **WHO policy and recommendations for measuring TB mortality**

8.1 **Definition of TB mortality**

TB mortality is the number of deaths from TB that occur in a given year. It is usually reported as the total number of deaths from TB (e.g. millions of deaths), or as the number of deaths for a given unit of population (e.g. number of deaths per 100 000 population). TB mortality can usually be reduced more quickly than TB incidence because drug treatment cuts not only transmission but also the case fatality rate.

8.2 **Methods for measuring TB mortality**

There are three ways to measure TB mortality:

- **direct measurement using vital registration (VR) data.** This is possible if death registration data collected in VR systems are coded according to the International Statistical Classification of Diseases\(^\text{11}\) and that data are of proven completeness and accuracy;

- **direct measurement using verbal autopsy (VA) studies.** In such studies, a structured set of questions are asked of care-givers or family members of people who have died, with the aim of determining the cause of death.\(^\text{23, 24}\) Such studies can form part of a sample vital registration (SVR) system, or may be undertaken in conjunction with other studies. In an SVR system, births and deaths are registered from a nationally representative sample of sites throughout the country. All deaths are followed up at household level, where a VA interview is conducted with the aim of determining the underlying cause of death. The results for the sample are then extrapolated to the national population;

- **indirect measurement using estimates of case fatality rates and TB incidence.** Here, TB mortality is estimated as TB incidence multiplied by the estimated case fatality rate and TB incidence. The method is only as reliable as the underlying estimates of incidence and case fatality.

From a TB perspective, a general problem with death certification is that TB may be listed as one of the associated causes of deaths in VR systems, but not recorded as the underlying cause of death (as opposed to HIV/AIDS, which is always recorded as the underlying cause of death). Since it is usually only the underlying cause of death that is considered when national statistics are computed, these statistics may underestimate the number of deaths in which TB is a contributing factor.

The case fatality rate is easiest to assess for patients on treatment, especially in DOTS programmes. Case fatality rates outside DOTS programmes are much harder to estimate. Outcomes are not recorded and while the risk of dying is likely to be higher it is difficult to quantify. This creates considerable uncertainty about the overall case fatality rate among TB cases.

A further problem is that based on current guidelines, deaths recorded in DOTS cohorts may or may not be due to TB.

\(^{11}\) Preferably, ICD-10. It may be possible to use ICD-9 but this is not ideal.
To estimate case fatality rates accurately from the TB notification system, it is essential that:

- a high proportion of cases are detected and treatment outcomes are known for all (or almost all) patients;
- the outcomes of re-treatment cases are known;
- the validity of outcome data is routinely checked;
- recorded deaths are attributed to either TB or to other causes.

As DOTS cohorts account for an increasing share of total cases, cohort outcomes should converge more closely with national death registrations.

### 8.3 WHO policy and recommendations for measuring TB mortality

The main policies and related recommendations about how TB mortality should be measured are summarized in Box 8. The subsequent text explains these policies and recommendations and how they can be implemented. A Task Force meeting in 2009 will focus on the measurement of TB mortality, and this may result in more specific recommendations than those included here.

**Box 8. WHO policy and recommendations for measuring TB mortality**

<table>
<thead>
<tr>
<th>Policy 1</th>
<th>WHO will promote the strengthening of vital registration (VR) systems in all countries, as part of more general efforts to improve health information systems. Where reliable VR systems are not yet in place, the use of sample vital registration will be encouraged.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recommendation 1</td>
<td>The best way to measure deaths from TB is through a national vital registration system in which causes of death are coded using the ICD-10 system. All countries should strengthen their vital registration systems so that TB deaths, as well as other causes of death, can be reliably measured.</td>
</tr>
<tr>
<td>Recommendation 2</td>
<td>Where vital registration systems are weak or not yet developed, sample vital registration should be used as an interim solution for the reliable measurement of deaths, including deaths from TB.</td>
</tr>
</tbody>
</table>

The best way to measure the number of deaths from TB is via a national vital registration (VR) system. In the long run, all countries should be able to report TB deaths among routine death registrations (coded as in the International Statistical Classification of Diseases, ICD-10), in systems that give data of proven completeness and accuracy.

Currently, few countries with a high burden of TB have vital registration systems that can be used to directly measure TB deaths. In a review published in 2003, only about one-third of the 56 million deaths that occur each year (all causes) were reported by VR systems, and among 115 countries that report deaths and their causes only 23 were assessed to have high-quality data (≥ 90% complete, ill-defined codes < 10%). None of the countries that rank first to twenty-second in terms of total number of TB cases were among this list of 23 countries, although three of the 22 high TB-burden countries (Brazil, the Philippines and the Russian Federation) were among a list of 55 countries assessed to have data of medium quality (70–90% completeness). Of these three countries, the accuracy of the TB records, among all other causes of death, has been investigated only in Brazil. Two other high-burden countries (South Africa and Thailand) were assessed to provide death registrations of low quality. The majority of countries in the African and South-East Asia regions did not have national systems for death registration. Although the classification of countries has been questioned, the
review gives an indication of the scale of the challenge involved in developing or strengthening VR systems.

Better systems for reporting and recording TB deaths within VR systems need to be developed as part of wider efforts to increase death registration and improve health information systems. The resources that are needed to establish and maintain a functional VR system are considerable, requiring maintenance of the flow of information at national and sub-national levels (from data entry to analysis and dissemination of findings) as well as provision of training to doctors to ensure that death certificates are completed correctly.

Where both VR and TB notification systems already exist, systematic comparison (cross-referencing) should be used to:

- reduce the proportion of TB cases for whom no outcome data is available;
- check the validity of the mortality data held in the TB notification system;
- check the diagnosis of the cause of death that is stated in the VR system;
- identify how many TB deaths occurred without medical care being accessed (diagnosis of TB done close to or after death);
- identify how many TB cases died without being notified.

This cross-referencing can be done at national level using sophisticated record-linkage methodology. It can also be done at district level, at regular intervals, by manual comparison of records. Some countries may have additional sources of data on deaths from TB that could be used for cross-referencing with VR or TB notification data.

Besides cross-referencing, VR data should be used to better understand the epidemiology of TB and in turn to identify ways in which deaths from TB could be reduced. Analysis of variables listed on death certifications (e.g. age, sex, residence, occupation, place of death) can be used to better understand risk factors for TB deaths. In addition, while the ICD-10 system states that there can only be one single underlying cause of death, analysis of the multiple proximate or associated cause-of-death codes listed on the death certificates can be used to assess how many of the deaths attributed to other causes (notably HIV/AIDS) have TB as an associate cause, how many deaths are attributed to TB sequelae, and the extent to which TB deaths attributed to different clinical types and the availability of microbiological confirmation vary spatially and over time.

Where VR systems are weak or not yet developed, sample vital registration (SVR) represents the most promising interim solution for the reliable measurement of deaths (including deaths from TB). When applied together with validated VA procedures and implemented in a nationally representative sample of population clusters, it represents an affordable, cost-effective, and sustainable short- to medium-term alternative. As such, its use should be expanded. The major challenge for SVR studies based on verbal autopsy is to prove the validity of the diagnosis of cause of death.

**8.4 Guidance material**

Published studies that have analysed TB deaths using data from vital registration systems are scarce. Examples include:

- a descriptive study in Brazil. Using data from 2004, this found that if the number of deaths with TB as an underlying cause was added to the number of
deaths with TB as associate cause and with TB sequelae as an underlying cause, the total number of TB deaths as computed in the national statistics would be increased by 50%;

and

- a capture-recapture study in England. Using three sources of TB mortality data, this study found that the number of deaths among TB cases is underestimated in national TB surveillance data.\textsuperscript{43}

The use of verbal autopsy to estimate TB deaths has been described for India, China and Tanzania:

- in Chennai, India, verbal autopsy has been used to substantially reduce the number of deaths attributed to unspecified causes on death certificates. Some of these deaths were reclassified as TB deaths, although the accuracy of this reclassification is unknown;\textsuperscript{67, 68}

- in China, verbal autopsy was found to frequently misclassify leading causes of death among adults, although for a given cause of death the number of false-positives and false-negatives tended to cancel out;\textsuperscript{69}

- a study in Tanzania found that verbal autopsy for TB deaths met validation criteria (e.g. there was similarity between verbal autopsy and medical records).\textsuperscript{70}


The recommendations in sections 6, 7 and 8 focus on strengthening of TB surveillance and implementation of surveys of the prevalence of TB disease from 2008 onwards. To assess whether or not the targets of halving TB prevalence and death rates by 2015 compared to a baseline of 1990 are achieved or not, estimates are needed for the period 1990 to 2015. Moreover, few countries will have routine surveillance systems that directly measure TB incidence and mortality by 2015; in-depth assessments of the quality and coverage of routine surveillance data cannot be conducted every year; and only a relatively small number of countries will implement surveys of the prevalence of TB disease. This means that a set of methods is needed to:

- produce country-specific estimates back to 1990;

- produce estimates of the prevalence of TB disease for countries that do not implement surveys of the prevalence of TB disease, and to produce estimates of TB prevalence for all countries for the years in which a prevalence survey was not undertaken; and

- produce estimates of TB mortality for countries without vital registration or sample vital registration systems.

The range of available methods for doing each of these things is explained in sections 6 to 8. In each case there is more than one choice, and within each of the available methods there is room for judgement and interpretation of the available data. To respond to the challenges that this presents, and with a view to ensuring that estimates of TB incidence, prevalence and mortality are both credible and widely-endorsed (a key component of the Task Force's mandate), one of the three subgroups of the Task Force will conduct a periodic review of the data, assumptions and analytical methods that are used, with the first review conducted from 2008 to 2009. Such reviews will contribute not only to the data, assumptions and analytical methods used by WHO to produce the estimates included in the annual series of reports on global TB control, but
also to the estimates for TB being produced for an update of the Global Burden of Disease.\(^\text{12}\)

10. Methods to evaluate the contribution of TB control to changes in TB incidence, prevalence and mortality

The previous four sections (sections 6 to 9) have explained the variety of methods that can be used to measure TB incidence, prevalence and mortality, and the methods that WHO recommends should be used up to 2015. Such measurement is essential for answering the question of whether or not the 2015 global targets for TB control are achieved. It also prompts a further and more difficult question: To what extent are changes in TB incidence, prevalence and mortality attributable to interventions specifically designed to control TB, and to what extent are they due to other factors? In other words, what is the impact of TB control? In terms of the nature of the two questions being asked, this is a shift from "monitoring" to "evaluation".

The number of recent studies that have evaluated the epidemiological impact of TB control is relatively small. Examples include evaluations of the impact of DOTS on the prevalence of TB in China and the impact of DOTS on TB incidence and mortality in Peru. Most recently, the observation that there has been no obvious reduction in TB incidence in countries that have implemented DOTS programmes with high reported treatment success rates and estimated case detection rates for several years (e.g. Viet Nam, India)\(^\text{15}\) has raised concerns about the impact of DOTS, and whether the impact of chemotherapy on TB incidence that was observed in the mid-twentieth century in Western Europe and North America can be reproduced.\(^\text{46}\) More evidence about the impact of TB control on TB incidence, prevalence and mortality is needed.

This section starts by providing a definition of "impact evaluation", followed by a brief discussion of the methods that can be used to assess the impact of any intervention or programme. It then discusses how these methods can be applied to the evaluation of the epidemiological impact of TB control, using recent examples.

10.1 Definition of impact evaluation

There is no single and universally-used definition of impact evaluation. However, there is agreement that impact evaluations assess the broad, long-term impacts or results of interventions, and that they assess these long-term impacts or results by comparing an intervention with at least one relevant alternative. The relevant alternative could include no intervention, or an alternative intervention (or set of interventions). This emphasis on comparison with at least one alternative scenario allows analysis of the extent to which the intervention under evaluation caused changes in impact indicators or to what extent changes were due to other factors.

10.2 Methods for evaluating the impact of public health interventions

The gold-standard for evaluating an intervention is a randomized controlled trial (RCT). In an RCT, the intervention being evaluated is randomly applied to otherwise

\(^{12}\) The update of the Global Burden of Disease will produce estimates for 2005, due for publication in 2010.
similar individuals or populations. Provided that the study design is robust enough (e.g. sample size large enough, period of follow-up long enough) to detect real differences between the populations under investigation, any observed differences between the intervention and control group in the outcomes of interest can be attributed to the intervention. Three well-known examples of RCTs for public health interventions are: STI treatment for the prevention of HIV in Mwanza, Tanzania; mass STI treatment for HIV prevention in Rakai, Uganda; and STI treatment combined with a behavioural intervention in Masaka, Uganda.\(^{47-49}\)

Evaluating public health interventions through RCTs is often not possible.\(^{50}\) Reasons include the fact that interventions may already be widely implemented, and removing them is neither feasible or ethical; that randomization based on geographic areas can break down if people seek prevention or care services outside their own place of residence; and that there is often international, national or local pressure to implement interventions universally rather than in selected areas. Moreover, even when an intervention has been shown to have an impact in the context of a trial, its impact under routine programmatic conditions still needs to be shown.\(^{51}\) In theory, step-wise implementation can overcome these problems for new interventions, provided that the individuals or geographic areas that receive the intervention in each phase are truly comparable. "Natural experiments" when interventions are withdrawn for short periods for reasons beyond the control of a programme can also help to evaluate the effect of an intervention; the caveat is that in such situations it is rarely only the intervention that changes and the role of other factors has to be considered too.\(^{20}\)

Besides RCTs, impact assessments can be based on observation of what happened when an intervention or programme was implemented, and comparisons with (i) otherwise similar areas during the same time period and/or (ii) with the same area before a new intervention or programme was implemented and/or (iii) with a hypothetical scenario for the same area during the same time period. In the literature on evaluation, such evidence is said to allow "plausibility statements" to be made, as opposed to the "probability statements" that can be made based on RCTs.\(^{52}\)

A third approach to impact evaluation is to analyse trends in the impact indicators of interest in combination with trends in indicators that measure the implementation of the intervention or programme being evaluated. If changes in impact indicators can be explained by changes in indicators of programme/intervention implementation, this can be used as evidence of programmatic or intervention impact. Such evidence is said to allow "adequacy statements" to be made.

A checklist for the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND)\(^{13}\) has been developed and is useful when designing and reporting an impact evaluation.\(^{53}\) Whichever one of the three main approaches to impact evaluation is used, it is essential to define a causal framework. This should show how a programme or intervention is expected to influence long-term results or impacts, in the context of other external and internal factors. It should also explain any underlying assumptions.

\(^{13}\)This is the equivalent of the CONSORT statement developed for RCTs.
10.3 Impact evaluation in the context of TB control

Evaluation of the impact of TB control on TB incidence, prevalence and mortality at the population level using an RCT has never been attempted. To date, RCTs related to TB interventions have been used principally to assess the efficacy of alternative drug regimens and more recently the efficacy of alternative approaches to the supervision of treatment,\textsuperscript{14} in cohorts of individual patients. In both cases the cure rate was the principal outcome indicator.

Using observational data combined with (i) before-and-after comparisons and/or (ii) comparisons between similar areas with different approaches to TB control is more feasible. Even so, it has been done relatively rarely. Three examples are:

- **Evaluation of the impact of DOTS on the prevalence of TB disease in China.**\textsuperscript{20} This was possible because during the 1990s, the DOTS strategy was implemented in 50% of the country while the other half of the country did not implement DOTS. The impact of DOTS on the prevalence of TB disease was then assessed by comparing reductions in the prevalence of TB prevalence in the two areas, with prevalence measured in 1990 and 2000. The prevalence of TB disease was similar in the two areas in 1990, but in 2000 the prevalence of TB in DOTS areas was 32% lower than in non-DOTS areas;

- **Evaluation of the impact of DOTS on TB mortality in India.**\textsuperscript{55, 56} Here, a before-and-after comparison was used to evaluate how the introduction of the revised national TB control programme (RNTCP) had reduced the number of deaths from TB;

- **Evaluation of the impact of DOTS on TB incidence and mortality in Peru.**\textsuperscript{57} Notification data for the years 1958 to 2000 (32 years before the implementation of DOTS and ten years after its introduction), data on registered deaths from TB for 1948–1998, treatment outcome data and assumptions about the fraction of cases that were untreated, were used. TB incidence and mortality under DOTS (1991 to 2000) were estimated from these observed data, and compared with hypothetical projections of TB incidence and mortality in the absence of DOTS. Evidence that was used to support the case that reductions in TB incidence and deaths were due to DOTS included trends in the number of health units providing TB diagnosis and treatment services and the number of sputum smears examined.

While the number of impact evaluations based on observational data for TB is small, many countries have reported falling TB notifications over many years, as well as reduced TB mortality rates in cohorts of patients treated under DOTS. With such data, the impact of TB control programmes can be estimated as [total number of TB cases that would have occurred without DOTS × case fatality rate prior to DOTS] - [total number of TB cases that occurred with DOTS × case fatality rate under DOTS]. The evaluation of DOTS in Peru used this method. A more conservative approach would be to estimate the number of deaths averted by TB control as [number of patients treated] × [reduction in case fatality rate under DOTS]. While such indirect estimates of impact are inferior to direct measures of the number of deaths averted, they can be produced for most countries.

\textsuperscript{14} Direct observation of treatment (DOT) by different types of supervisor has been compared with unsupervised treatment.
Box 9. WHO policy and recommendations on evaluation of the impact of TB control on incidence, prevalence and mortality

| Recommendation 1. The impact of TB control on TB incidence should be evaluated using trends in TB incidence combined with evidence about the extent to which trends can be explained by changes in TB control. |
| Recommendation 2. The impact of TB control on TB prevalence should be evaluated using data from at least two population-based surveys of the prevalence of disease that used identical or similar methods, combined with evidence about the extent to which changes can be explained by changes in TB control. |
| Recommendation 3. The impact of TB control on TB mortality should be evaluated using trends in TB deaths combined with evidence about the extent to which trends can be explained by changes in TB control, and/or estimation of averted TB deaths using data on trends in TB incidence, reductions in mortality rates observed in patient cohorts, and the share of TB incident cases being treated in these patient cohorts. |
| Recommendation 4. Indicators that could be used to understand the extent to which changes in TB control explain changes in TB incidence, prevalence and mortality include: treatment success and case detection rates (outcome indicators); the numbers of people or patients receiving particular interventions (output indicators); the number of people trained and the number of sputum smears examined (process indicators); and the number of health units where TB diagnostic and treatment services are available, the number of staff working for the NTP, and funding (input indicators). |

Evaluation of the impact of TB control through analysis of trends in variables that are thought to influence TB incidence, mortality and prevalence is feasible, though to date it has not been systematically done. A recent example is an ecological analysis of the relationship between a variety of indicators and trends in TB incidence (with trends in notifications used as a proxy for incidence). Another recent study based on analysis of trends in HIV prevalence, GDP per capita, DOTS population coverage and treatment success rates also illustrates the use of trend data to explore the impact of TB control, although in this case the analysis focused on case detection and treatment success rates (outcome indicators) rather than incidence, prevalence or mortality.

Figure 2. An example of a causal framework

As noted in section 10.2, a causal framework should be defined when any of the above methods (RCT, observational data that allow comparison with what would have occurred in the absence of TB control or a specific strategy for TB control, analysis of trends) is used. A simple example is shown in Figure 2, based on the classification of indicators used in the TB compendium of indicators. In this causal framework, inputs (e.g. funding, staff, buildings) are transformed into processes (e.g. training, supervision, provision of drugs to TB patients, laboratory tests), which are then transformed into outputs (e.g. numbers of TB patients treated). Depending on the quantity and quality of treatment provided, these outputs are transformed into outcomes - the two major outcome indicators for TB control being successful treatment rates and case detection rates. Increases in case detection and successful treatment rates should cut TB incidence, prevalence and mortality because (i) the average duration of infectiousness should be reduced and (ii) the proportion of cases that dies should fall. The greater the effect on the duration of infectiousness and the greater the reduction in the risk of dying on treatment, the bigger will be the combined effect on TB mortality. However, it is possible for TB incidence to remain stable or even increase when case detection...
and successful treatment rates are rising, if the risk of developing TB at the population level is increasing (e.g. if the prevalence of risk factors for TB is rising).

This type of causal framework places emphasis on measures of programme performance besides the traditional indicators of case detection and successful treatment rates, as well as on factors that influence the epidemiology of TB besides TB control programmes. In particular, it shows the relevance of collecting data on a standard set of programmatic inputs, processes, and outputs to track and evaluate progress in TB control over time. For example, linked to the measurement of trends in incidence discussed in section 6, data on trends in programmatic inputs, processes and outputs can help to separate out the effect of programmatic changes on case notifications from the effect of underlying changes in incidence (Box 9). Where notification data are the main mechanism for measuring trends in incidence, this is particularly important. Data on programmatic inputs, processes and outputs can also be used to evaluate the extent to which changes in TB prevalence and mortality can be explained by changes in TB control (Box 9).

Compiling data on inputs, process and outputs in addition to the standard monitoring of case notifications and treatment success rates can be difficult and time-consuming. However, if done routinely or at regular intervals, it will provide a much stronger basis for evaluating both progress in TB control and the impact of TB control than currently exists. The Kenyan study described in Box 4 (section 6) and the study in Peru described in this section show that the collection and analysis of such data is possible, and that it can be used to draw important conclusions about the absolute level of TB incidence and mortality, and their trends over time. The TB data collection form used by WHO to collect data from NTPs has been revised with the aim of compiling and analysing such data more systematically and for more countries.
11. Conclusions

The 2015 impact targets for global TB control, set within the MDGs and by the Stop TB Partnership, are to ensure that the TB incidence rate is falling, and that TB prevalence and death rates are halved compared to their level in 1990. Achieving these targets is the focus of international and national efforts in TB control, and demonstrating whether or not they are achieved is of major importance for individual countries, the UN and Stop TB Partnership, and a variety of technical, financial and development agencies.

This policy paper has set out WHO recommendations about how progress in reducing TB incidence, prevalence and mortality should be measured in the years up to 2015, many of which will apply in subsequent years. The main recommendations can be summarized as follows:

- The ultimate aim for all countries should be to count, accurately and comprehensively, TB cases (incidence) and deaths (mortality) through routine surveillance and vital registration systems;
- To measure TB incidence, all countries need to improve the reliability and coverage of their routine TB notification data. Until these data provide a direct measure of TB incidence, TB incidence will need to be measured indirectly. This should be done via periodic in-depth assessments of the reliability and coverage of routine TB surveillance data using a standard framework and related tool. Such measurements may be supported or cross-validated by data on TB mortality available from vital registration systems, where such systems exist and are able to provide timely, complete and reliable data;
- To measure TB prevalence, there are 21 global focus countries that should implement at least one national population-based survey of the prevalence of TB between 2008 and 2015. These surveys should follow WHO guidelines published in 2007 and related Task Force recommendations. For other countries, TB prevalence will need to be measured indirectly from estimates of TB incidence and the average duration of disease;
- To measure TB mortality, all countries should strengthen their vital registration systems, and use sample vital registration as an interim solution. When vital registration or sample registration systems do not exist or do not provide reliable data, it will be necessary to estimate TB mortality indirectly from estimates of TB incidence and the case fatality rate;
- The Task Force will periodically review the data, assumptions and analytical methods used to produce estimates of TB incidence, prevalence and mortality, with particular focus on the data, assumptions and analytical methods used to produce indirect estimates. This work will include periodic review of the data, assumptions and analytical methods used to estimate TB incidence, prevalence and mortality back to 1990, as well as those used from 2008 onwards;
- Impact evaluations will need to be conducted for more countries, to improve the current evidence about the contribution of TB control to changes in TB incidence, prevalence and mortality.
12. References

53. Khatri GR, Frieden TR. Controlling tuberculosis in India. NEJM 2002;347:1420-5.
68. Gajalakshmi V, Peto R. Verbal autopsy of 80,000 adult deaths in Tamilnadu, South India. BMC Public Health 2004;4:47.
13. Annexes

Annex 1. Membership of the WHO Global Task Force on TB Impact Measurement (institutions and/or individuals)

**TB endemic countries:** Representatives from countries with a high burden of TB. Task Force meetings to date have included representatives from India, Indonesia, Malawi, Nigeria, the Philippines, South Africa, and Tanzania.

**International technical agencies with expertise in TB epidemiology:** CDC (Centers for Disease Control, Atlanta, USA); ECDC (European Centre for Disease Control, Stockholm, Sweden); KNCV Tuberculosis Foundation (The Hague, the Netherlands); RIT (Research Institute for Tuberculosis, Tokyo, Japan); the Union (International Union against TB and Lung Disease, Paris, France); PATH (Washington, USA), MEASURE Evaluation (Arlington, USA); WHO (HQ and Regional Offices).

**Financial agencies:** The Global Fund; USAID (United States Agency for International Development); the World Bank.

**Task Force Chair:** Jaap Broekmans, former Executive Director of KNCV Tuberculosis Foundation in the Netherlands and former Chair of the WHO Strategic and Technical Advisory Group on TB (STAG-TB).

Three subgroups have been established for each of three major strategic tracks of work that the Task Force will pursue. Membership is based on the areas of interest and expertise of Task Force members. Experts from outside the Task Force who have been nominated to each group are also listed.

**Area 1 (Routine surveillance data)**
Ana Bierrenbach (leader/coordinator, WHO-HQ), Chen-Yuan Chiang (the Union), Peter Gondrie (KNCV), Nico Kalisvaart (KNCV), Mehran Hosseini (WHO-HQ), Eugene McCray (CDC), Andrei Dadu (WHO/EURO), Ryuichi Komatsu (Global Fund), Amal Bassili (WHO/EMRO), Davide Manissero (ECDC, Stockholm), Ibrahim Abubakar (Health Protection Agency, UK), Charlotte Colvin (PATH), Daniel Boone (MEASURE Evaluation), Ted Cohen (Harvard University).

**Area 2 (Prevalence surveys)**
Ikushi Onozaki (leader/coordinator, WHO-HQ), Ana Bierrenbach (WHO-HQ), Marieke van der Werf (KNCV), Frank van Leth (KNCV), Philip Patrobás (WHO, Nigeria), Eliud Wandwalo (NTP, Tanzania), PG Gopi (formerly Tuberculosis Research Centre, Chennai, India), Norio Yamada (RIT/Japan anti-TB Association), Eugene McCray (CDC), Daniel Chemtob (NTP, Israel), VK Chadha (National Tuberculosis Institute, Bangalore, India), Amal Bassili (WHO/EMRO).

**Area 3 (Production of epidemiological estimates and evaluation of trends)**
Marieke van der Werf (co-leader/coordinator, KNCV), Philippe Glaziu (co-leader/coordinator, WHO-HQ), Ikushi Onozaki (WHO-HQ), Ana Bierrenbach (WHO-HQ), Eliud Wandwalo (NTP, Tanzania), Brian Williams (WHO-HQ), PG Gopi (formerly of Tuberculosis Research Centre, Chennai, India), VK Chadha (National Tuberculosis Institute, Bangalore, India), Norio Yamada (RIT/Japan anti-TB Association), Davide Manissero (ECDC, Stockholm), Daniel Chemtob (NTP, Israel), Andrei Dadu (WHO/EURO), Amal Bassili (WHO/EMRO).
Annex 2. Extracts from World Health Assembly resolution 60.19

The WHA requests the Director-General:

"… (5) to strengthen mechanisms to review and monitor estimates of impact of control activities on the tuberculosis burden, including incidence, prevalence and mortality with specific attention to vulnerable groups highly at risk, such as poor people, migrants and ethnic minorities; ...

...(8) to report to the Sixty-third World Health Assembly through the Executive Board on :

... (b) progress made in achieving the international targets for tuberculosis control by 2015, using the "proportion of tuberculosis cases detected and cured under DOTS" (Millennium Development Goal indicator 24) as a measure of the performance of national programmes, and tuberculosis incidence and "prevalence and death rates associated with tuberculosis" (Millennium Development Goal indicator 23) as a measure of the impact of control on the tuberculosis epidemic."
Annex 3: Further details on surveys of the annual risk of infection

Tuberculin skin test (TST) surveys of the annual risk of infection (ARI) are usually conducted with a sample of around 10,000 children aged from 5 to 15 years old. Children infected with *M. tuberculosis* are identified by the size of their skin-test reaction. The prevalence of infection combined with the mean age of sampled children (= average years of exposure) can then be used to estimate the annual risk of infection ARI. The incidence of smear-positive TB is then estimated using the rule-of-thumb proposed by Styblo, which is that an ARI of 1% is equivalent to 50–60 smear-positive cases per 100 000 population per year (equation 1). The value is usually assumed to be 50.

\[
\text{incidence (smear - positive)} = \text{annual risk of infection} \times \text{coefficient (usually 50)} \tag{1}
\]

Tuberculin skin test surveys have the advantage of being relatively cheap (around US$ 20,000 to US$ 50,000 each) and logistically straightforward to implement. However, the results can be difficult to interpret. Reasons for this include:

- positive responses to *M. tuberculosis* infection can be obscured by unpredictable cross-reactions from BCG vaccination or environmental mycobacteria. In general, the lower the ARI, the harder it is to distinguish the population of true positives from the population of cross-reactors;
- in recent surveys, the distributions of true and false positives often overlap to a large extent. This means that a single cut-off point cannot be used to identify the number of children that are infected;
- measuring the prevalence of infection requires rigorous application of standard methods;
- the Styblo rule-of-thumb, which is based on the assumption that each smear-positive case will infect 10 individuals in a year, that an untreated case remains smear-positive for two years, and that the reproduction number equals one (the epidemic is in a stable state), does not necessarily apply. Two recent reviews found that the number of contacts that became infected was much less than 10 in most settings.16, 17 This implies that the ratio of smear-positive incidence to ARI exceeds 50.

The value of TST surveys can be improved if the data can be calibrated with results for patients with active disease, and if it can be assumed that TST indurations have the same distribution among patients (usually adults) and the infected individuals studied in a TST survey (usually children). The most-commonly used tuberculin preparation appears to generate a plausible distribution of induration sizes for patients with a mode at around 18 mm, but may not always do so.

A second approach is to calibrate TST survey data with highly specific, interferon-\(\gamma\) release assays (IGRAs). Examples including the enzyme-linked immunospot (ELISpot) and the enzyme-linked immunosorbent assay (ELISA). The drawback of IGRAs is that they are costly (> US$10 per assay) and require blood taken by venepuncture (not merely a finger-prick). They are also less sensitive than TST for indicating whether a child or adult has ever been infected: in studies undertaken to date, a significant proportion of TB cases tested IGRA negative.
Annex 4. Further details surveys of the prevalence of TB disease

Population-based surveys of the prevalence of TB disease rely on cluster sampling (i.e. whole areas are selected, and then a suitable number of individuals is selected from each of these areas). Cluster sampling is used to reduce survey time and costs. Six national surveys were implemented between 1995 and 2007, in China, Cambodia, Eritrea, Indonesia, the Philippines, and the Republic of Korea. Further surveys are underway or planned (Annex 6).

Large sample sizes are needed to produce precise estimates of the national prevalence of TB (although not as large as the sample sizes needed to measure incidence). For example, if the prevalence of sputum smear-positive TB is 100 per 100 000 population, a random sample of 100 000 people can be expected to yield 100 cases with a 95% confidence interval of ± 20%. To allow for incomplete data as well as the need to increase sample size when cluster rather than simple random sampling is used (the so-called "design effect"), a good rule-of-thumb is to double the sample size compared to that needed for a simple random sample of individuals i.e. in this example, to 200 000 people. In the six national surveys listed above, the prevalence of smear-positive TB was measured with an accuracy that varied from ± 25% (China, the Republic of Korea) to ± 60% (Eritrea).

Surveys to measure TB prevalence sub-nationally (e.g. in states or provinces) can also be carried out. Recent examples include surveys in Bangladesh, Botswana, Cambodia, Ethiopia, India, and Uganda. However, for a given level of precision, a much larger overall sample size is needed if the objectives of a national prevalence survey include sub-national estimates of prevalence. If combined with ARI data, it may be possible to use sub-national data on the prevalence TB disease to derive a national estimate of the prevalence of TB disease. For example, calibration of ARI and prevalence of TB disease data from the same site in south India have been combined with ARI data from four different regions to construct a national estimate of disease prevalence for India.

One of the most critical methodological issues in any survey of the prevalence of TB disease is how to select which members of the survey sample will have sputum specimens taken for microscopy and culture examination. The major options are:

- Screening all sampled individuals with chest X-rays. Few cases of active TB will be missed when this approach is used;
- Screening all sampled individuals on the basis of the symptoms of chest illness (e.g. cough for 2 weeks). This will tend to miss mildly-symptomatic and asymptomatic cases;
- Screening using a combination of chest X-rays and symptoms. In recent surveys, this has been the most commonly used screening method;
- Obtaining sputum from all sampled individuals. This is low-cost and requires minimal technology. However, it requires consistent efforts in sputum collection and generates a large laboratory workload which must be accurately processed.

Most surveys of the prevalence of TB disease have been based on TB survey-specific sampling frames. However, other possibilities exist. The 2004 survey in Indonesia took advantage of a sampling frame established for the National Household Health Survey.
Changes in prevalence over time are best measured by implementing at least two (but preferably more) surveys at large enough intervals. The sample sizes required to measure changes in prevalence are larger than those needed for a one-time measurement of the total (absolute) number of cases. For example, if two surveys are carried out five years apart with the aim of detecting a 30% reduction in prevalence (from 100 cases per 100,000 population to 70 cases per 100,000 population), around 400,000 people would need to be examined in each survey. The lower the initial estimate of prevalence and the smaller the expected change in prevalence, the bigger the sample size that is needed to detect changes over time.

Besides the importance of large enough sample sizes, surveys to measure changes in TB prevalence over time need to be comparable in terms of study design, diagnostic protocols and case definitions. When new surveys are implemented, there can be a trade-off between maximizing the similarity of methods and making use of newer and better screening methods and diagnostic tools; if so, care is needed to ensure that the comparability of the surveys is not compromised.
Annex 5. Selection of countries in which surveys of the prevalence of TB disease should be undertaken during the period up to 2015

Table A6.1 Extended list of 32 countries that met one of the four sets of criteria for carrying out a survey of the prevalence of TB disease, but that were not selected as global focus countries by the Task Force

<table>
<thead>
<tr>
<th>Criteria met</th>
<th>Region</th>
<th>Country</th>
<th>High-burden</th>
</tr>
</thead>
<tbody>
<tr>
<td>2,3</td>
<td>AFR_high HIV</td>
<td>Botswana</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>AFR_high HIV</td>
<td>Burundi</td>
<td>No</td>
</tr>
<tr>
<td>1,2</td>
<td>AFR_high HIV</td>
<td>Central African Republic</td>
<td>No</td>
</tr>
<tr>
<td>2,3</td>
<td>AFR_high HIV</td>
<td>Congo</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>AFR_high HIV</td>
<td>Côte d'Ivoire</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>AFR_high HIV</td>
<td>DR Congo</td>
<td>Yes</td>
</tr>
<tr>
<td>1,3</td>
<td>AFR_high HIV</td>
<td>Ethiopia</td>
<td>Yes</td>
</tr>
<tr>
<td>2</td>
<td>AFR_high HIV</td>
<td>Lesotho</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>AFR_high HIV</td>
<td>Namibia</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>AFR_high HIV</td>
<td>Swaziland</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>AFR_high HIV</td>
<td>Zimbabwe</td>
<td>Yes</td>
</tr>
<tr>
<td>1,2</td>
<td>AFR_low HIV</td>
<td>Angola</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>AFR_low HIV</td>
<td>Burkina Faso</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>AFR_low HIV</td>
<td>Chad</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>AFR_low HIV</td>
<td>Eritrea</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>AFR_low HIV</td>
<td>Gambia</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>AFR_low HIV</td>
<td>Guinea</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>AFR_low HIV</td>
<td>Liberia</td>
<td>No</td>
</tr>
<tr>
<td>1,3</td>
<td>AFR_low HIV</td>
<td>Mauritania</td>
<td>No</td>
</tr>
<tr>
<td>1,2</td>
<td>AFR_low HIV</td>
<td>Niger</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>AFR_low HIV</td>
<td>Togo</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>EEUR</td>
<td>Armenia</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>EEUR</td>
<td>Russian Federation</td>
<td>Yes</td>
</tr>
<tr>
<td>1</td>
<td>EEUR</td>
<td>Tajikistan</td>
<td>No</td>
</tr>
<tr>
<td>1</td>
<td>EMR</td>
<td>Afghanistan</td>
<td>Yes</td>
</tr>
<tr>
<td>3,4</td>
<td>EMR</td>
<td>Djibouti</td>
<td>No</td>
</tr>
<tr>
<td>1,2</td>
<td>EMR</td>
<td>Sudan</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>LAC</td>
<td>Haiti</td>
<td>No</td>
</tr>
<tr>
<td>3</td>
<td>SEA</td>
<td>Timor-Leste</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>WPR</td>
<td>Lao PDR</td>
<td>No</td>
</tr>
<tr>
<td>4</td>
<td>WPR</td>
<td>Malaysia</td>
<td>No</td>
</tr>
<tr>
<td>1,2,3</td>
<td>WPR</td>
<td>Papua New Guinea</td>
<td>No</td>
</tr>
</tbody>
</table>
Annex 6. List of technical agencies providing support to recent or planned surveys of the prevalence of TB disease

*To be updated once all information reviewed and corrected as appropriate by endemic countries and technical agencies*

Ongoing or recently completed national prevalence surveys and supporting technical agencies

<table>
<thead>
<tr>
<th>Country</th>
<th>Technical Agencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bangladesh</td>
<td>ICDDR B(International Center for Diarrhoea Disease Research, Bangladesh), KNCV, (WHO)</td>
</tr>
<tr>
<td>Myanmar (subnational)</td>
<td>RIT/JATA WHO and JICA</td>
</tr>
<tr>
<td>Philippines</td>
<td>TDF (Tropical Disease Foundation), KIT (Korean Institute for TB), (US-CDC, RIT/JATA, WHO)</td>
</tr>
<tr>
<td>Indonesia</td>
<td>WHO</td>
</tr>
<tr>
<td>Vietnam</td>
<td>KNCV Tuberculosis Foundation, WHO, RIT/JATA</td>
</tr>
</tbody>
</table>

Upcoming national prevalence surveys that have already identified potential candidates supporting technical agencies (as of April 2008)

<table>
<thead>
<tr>
<th>Country</th>
<th>Technical Agencies</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kenya</td>
<td>KNCV Tuberculosis Foundation, KEMRI (Kenya Medical Research Institute), US-CDC</td>
</tr>
<tr>
<td>Malawi</td>
<td>Liverpool School of Tropical Medicine, National College of Medicine, REACH (Research on Equity and Community Health - Malawi), US-CDC</td>
</tr>
<tr>
<td>Mali</td>
<td>KNCV Tuberculosis Foundation</td>
</tr>
<tr>
<td>Nigeria</td>
<td>US-CDC</td>
</tr>
<tr>
<td>Rwanda</td>
<td>Damien Foundation, INS (Institut National des statistiques), Université Nationale du Rwanda - École de Sante Publique</td>
</tr>
<tr>
<td>Tanzania</td>
<td>NIMR (National Institute for Medical Research), MUCHS (Muhimbily University College of Health Sciences), KNCV Tuberculosis Foundation, Institute of Tropical Medicine, Antwerp, Belgium</td>
</tr>
<tr>
<td>Uganda</td>
<td>Kampala University</td>
</tr>
<tr>
<td>Zambia</td>
<td>ZAMBART (Zambia AIDS Related TB Research Team), RIT/JATA</td>
</tr>
<tr>
<td>Cambodia</td>
<td>RIT/JATA, WHO</td>
</tr>
<tr>
<td>Myanmar</td>
<td>WHO, RIT/JATA JICA, and PSI (Population Service International)</td>
</tr>
<tr>
<td>Thailand</td>
<td>RIT/JATA</td>
</tr>
<tr>
<td>Pakistan</td>
<td>WHO, the Union, KNCV Tuberculosis Foundation</td>
</tr>
</tbody>
</table>