Measuring Progress Towards the Millennium Development Goals for TB Control

Task Force convened under the auspices of the
Strategic and Technical Advisory Group (STAG),
Stop Tuberculosis Department,
World Health Organization

SCOPE AND PURPOSE

The objective of the Task Force, commissioned by WHO's STAG, is to:

Develop a costed plan to measure progress towards the Millennium Development Goals for tuberculosis control

To do this, the Task Force will identify the methods that could and should be used to measure the implementation and epidemiological impact (or effect) of TB control. TB control over the coming decade will be guided by the 2nd Global Plan to Stop TB, and following the expanded Stop TB Strategy based on DOTS. The Task Force will consider, among other things:

- methods for strengthening routine TB surveillance;
- ways to improve and standardize the analysis of surveillance data;
- where surveys of infection and active disease have already been done, and should be done in future;
- the comparative merits of surveys of infection and active disease (with reference to surveillance data);
- how to set standards for carrying out prevalence surveys (of both infection and disease);
- how new technology for diagnosis and treatment will affect measurement procedures;
- the funding required to provide technical assistance to key countries, which countries these are, and possible sources of this funding;
- ways to review, update and disseminate existing WHO guidelines and recommendations for assessing TB surveillance systems in countries;
- methods for translating the measurements into action on TB control (e.g. dissemination, resource mobilization)

PROCESS

The plan will be developed in consultation with technical and other experts representing WHO, NGOs, donor governments and other funding agencies, and health ministries of TB endemic countries. This paper has been prepared to facilitate the first round of consultation, which will address 10 key questions about measurement (below). The plan
will be drafted by a core writing committee (to be assembled), based on the answers to these questions, and on the results of parallel research on various aspects of the methodology (e.g. inventory of surveys, review of costs, critical appraisal of tuberculin surveys).

The Task Force will present its plan for measurement to STAG in June 2006. It will recommend which surveillance and survey methods and systems should be established and used in which countries, their strengths and weaknesses, how they could be implemented, what measurement will cost and how it will be paid for. STAG will be asked to endorse the technical aspects of the plan (epidemiology and financing).

As part of the process of developing the plan, the Task Force will present interim findings (from answers to the questions below) to the Stop TB Coordinating Board (10-11 November 2005, Assisi). It will seek the Board's advice on the scope and future of the process, and on the development and implementation of recommendations, especially with regard to sources of funding.

TECHNICAL BACKGROUND

Targets for TB control

The establishment of targets within the framework of the Millennium Development Goals (MDGs, Box 1) is providing greater impetus in the evaluation of health programmes. National TB control programmes are now thinking more actively about how to measure the epidemiological impact of control, in addition to monitoring DOTS implementation. The measurement of DOTS implementation has focused on assessing progress in case detection and treatment success. The evaluation of impact (due to any control method, not just DOTS) requires the measurement of TB incidence, prevalence and deaths between 2005 and MDG target year 2015. The success, or failure, of TB control needs to be evaluated for the principal endemic countries and regions, and for the world as a whole.

The targets in Box 1 have been derived from, and are intended to be as consistent as possible with, a series of resolutions and directives made by various bodies since 1991 (G8, UN, WHO, Stop TB Partnership, IUATLD, CDC). However, they are not fully, internally consistent from an epidemiological standpoint. For example, to halve the TB death rate by 2015 is much more demanding than the target of ensuring that incidence is falling by that year. There is an informal agreement within WHO that TB prevalence and deaths should exclude patients coinfected with HIV, mainly to avoid double counting deaths from AIDS and TB. However, the aim of TB control is to eliminate the disease from whole populations. WHO therefore reports TB statistics and estimates for both the HIV-positive and negative sub-populations.

The criterion for elimination in Box 1 is that defined by the US CDC, and differs from a European recommendation that elimination be defined as less than one smear-positive case per million in 2050. We propose to adopt the CDC recommendation for two reasons:
(i) the intention is to eliminate all forms of TB; (ii) by 2050, diagnosis by smear microscopy should be obsolete.

The MDGs provide the framework for evaluating the progress in TB control due to implementation of the second Global Plan to Stop TB, 2006-2015 (hereafter GP2), which will be launched at the next World Economic Forum in Davos (January 2006).

**Box 1. Goals, Target and Indicators for TB Control**

**Millennium Development Goal 6: Combat HIV/AIDS, malaria and other diseases**

**Target 8:** Have halted by 2015 and begun to reverse the incidence of malaria and other major diseases

**Indicator 23:** Prevalence and death rates associated with tuberculosis

**Indicator 24:** Proportion of tuberculosis cases detected and cured under DOTS (internationally recommended TB control strategy)

**Stop TB Partnership Targets**

**By 2005:** At least 70% of people with infectious TB will be diagnosed (i.e. under the DOTS strategy), and at least 85% cured.

**By 2015:** The global burden of TB (prevalence and death rates) will be reduced by 50% relative to 1990 levels. This means reducing prevalence to $\approx 150$ per 100,000 or lower and deaths to $\approx 15$ per 100,000 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 TB disease will be less than 1 case per million population per year (the criterion for TB "elimination" adopted within the USA).

Although many indicators are used to monitor and evaluate TB control (see the recent "Compendium"), we propose that the scope of this discussion be limited to those directly relevant to the MDGs. That is, the Task Force would not be concerned with, for example, defining and costing systems for monitoring drug supplies.

**What is known about TB epidemiology and control**

WHO uses routine surveillance\(^1\) reports and survey data to evaluate the scale and direction of the TB epidemic, and to evaluate the implementation and impact of TB control. Annex 1 contains an overview of epidemiology and control. Clearly, the

\(^1\) In this paper "surveillance" means the routine tracking of the TB epidemic by counting cases and deaths using the same data collection system over time; "monitoring" is the routine tracking of information about the control programme and its outcomes.
conclusions given in this summary are subject to various uncertainties. However, nothing can be measured with perfect precision - and in this sense every statistic is an "estimate". The key question is whether the data, and the conclusions drawn from them, are accurate enough to make key decisions about TB control. If not, what are the most important deficiencies, and how can they be rectified?

**Measurement of key indicators: pros and cons**

Table 1 summarizes the characteristics of key epidemiological indicators, and of the methods for their measurement. Annex 2 contains a fuller account of the attributes of these indicators.

Treatment outcomes are routinely reported in DOTS cohorts. The cohort system is well-established in most countries and should be used as the basis for modifications that will allow for the monitoring of, for example, patients carrying drug-resistant bacilli or who are co-infected with HIV.

The case detection "rate", as defined by WHO, is the ratio of smear-positive cases reported in any one year to the estimated incidence of smear-positive cases in that year. This is the most controversial of the indicators used to track DOTS implementation, mainly because the denominator - incidence - is hard to measure accurately.

Incidence is the central measure of progress towards elimination, which is the principal, long-term goal of TB control. The best method of evaluating incidence annually on a large scale (i.e. nationally and globally) is through routine surveillance, provided the data from surveillance is of proven accuracy, i.e. assuming that all (or nearly all) cases are detected. When most countries have good surveillance systems, case detection will become redundant as an indicator of implementation. Meanwhile, incidence estimates derived indirectly from tuberculin surveys (giving risk of infection, Styblo method), prevalence of disease (divided by duration of illness), or from TB deaths (divided by case fatality) can be useful to cross-check the analysis of routine surveillance data, but may not give accurate estimates of absolute incidence or trends in incidence.

As a measure of the short-term (<5 years) effect of TB control, incidence is less suitable than prevalence because it changes more slowly (so changes are statistically harder to detect), and there are usually fewer incident cases in any one year than prevalent cases at any one time. Consequently, in highly-endemic countries, especially those with weak surveillance systems, the effect of TB control is better measured by consecutive disease prevalence (cross-sectional) surveys. Cost and logistic complexity are the main drawbacks of prevalence surveys. Some high-burden countries will probably be able to do two or more prevalence surveys between now and 2015, but perhaps not all. In countries that can do surveys, they are unlikely to be done at less than 5-yearly intervals.

Besides incidence and prevalence, it is vital to have accurate counts of TB deaths. One reason (among others) is that most of the burden of TB, as measured in terms of DALYs (the common currency of morbidity and mortality), is due to premature deaths of young
adults. Moreover, TB deaths can usually be reduced more quickly than new cases in good chemotherapy programmes. The measurement problem is that a TB death is a rare event even in the so-called high-burden countries. The long-term goal is that TB deaths should be reported among all death registrations (coded as in ICD-10), in systems that give data of proven accuracy. In countries where vital registration is currently poor or non-existent (i.e. most high-burden countries), the best near-term evaluation of TB deaths is likely to come from general mortality surveys where death by cause is assessed by verbal autopsy (VA).

Tuberculin surveys have long been used as an instrument for measuring the prevalence of TB infection, and for estimating the annual risk of infection (ARI) from prevalence data. ARI is not an indicator used in the MDG framework, but accurate measures of the risk of infection, and the changing risk, would be valuable in support of measurements of incidence, prevalence and deaths. KNCV has been contracted by WHO to carry out a review of the utility of tuberculin surveys, especially with respect to the measurement of transmission, risk of infection and incidence at population level, and is due to report by the end of September 2005.

In summary, accurate routine surveillance and vital registration - to measure incidence, treatment success, and mortality - is the goal towards which all TB control programmes should be moving. When most countries have complete surveillance reports, case detection will become obsolete as an indicator of programme implementation. While routine recording methods are being strengthened in all countries over the next decade, surveys of disease prevalence and mortality can be carried out in selected countries, perhaps representing different regions, to monitor progress due towards the 2015 MDG targets. The performance of tuberculin surveys is unpredictable, and this technique may, or may not, provide supporting evidence on transmission and its trend. Surveys of infection, disease and death are unlikely to give, through indirect calculations, accurate estimates of incidence, though current WHO estimates for many countries rely on these indirect calculations.

**Sources of data**

WHO has records of more than 80 million TB patients from routine surveillance, 1980-2003, and more than 17 million patients that have been registered in DOTS programmes. Almost all of the 211 countries and territories report annually to WHO, but with data of variable quality. WHO asks for nationally aggregated data, which are usually compiled from data aggregated at district level. In 2005, few countries have case-based reporting systems that record and track individual patients, though the number is growing (especially among middle-income countries).

Population-based surveys of the prevalence of infection and disease, and occasionally of incidence, have been carried out since the 1950s. Annex 3 lists all the surveys of infection and disease that are known to WHO, including those which are in progress or planned.
Table 1. Advantages and disadvantages of various epidemiological measurements for TB control

Text in orange refers to attributes of the indicator; regular text refers to attributes of the measurement technique.

<table>
<thead>
<tr>
<th>MEASURE</th>
<th>ADVANTAGES</th>
<th>DISADVANTAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td>Prevalence of infection</td>
<td>Risk of infection changes relatively quickly in response to control (but prevalence of infection, from which risk is calculated, changes slowly).</td>
<td>Measures infection, not disease burden; not an MDG indicator.</td>
</tr>
<tr>
<td>From tuberculin surveys</td>
<td>Relatively cheap and logistically straightforward.</td>
<td>Results often hard to interpret where infection rates are low and where BCG coverage is high or where exposure to environmental mycobacteria is high; measures average risk of infection over past 5-10 years; Styblo 1:50 rule for indirectly estimating disease incidence may not be applicable under chemotherapy, or where HIV infection rates are high.</td>
</tr>
<tr>
<td>Prevalence of disease</td>
<td>Component due to incidence changes slowly in response to control; response to control. MDG indicator.</td>
<td>Component due to duration of illness changes.</td>
</tr>
<tr>
<td>From population-based surveys</td>
<td>Accurate measure of bacteriologically confirmed disease; should change quickly in response to control; surveys useful where routine surveillance data are poor, and are a platform for related investigations e.g. of interactions between patients and health system.</td>
<td>Costly; logistically complex (especially with radiography), therefore cannot be measured annually; does not easily lead to an estimate of TB incidence (denominator of WHO case detection rate), because duration is hard to assess.</td>
</tr>
<tr>
<td>Incidence of disease</td>
<td>Direct measure of denominator of WHO case detection rate; MDG indicator.</td>
<td>Changes slowly following reductions in transmission.</td>
</tr>
<tr>
<td>From case notifications</td>
<td>Direct measure of incidence; absolute incidence can be assessed from routine case reports where case detection judged to be high; trends can be judged from series of routine case reports, if measured consistently; every country now has a surveillance system, reporting annually or sub-annually.</td>
<td>Case detection mostly low in high-burden countries (underestimates incidence), and may vary through time (inaccurate trends).</td>
</tr>
<tr>
<td>From longitudinal, population-based cohort studies</td>
<td>Direct measure of incidence.</td>
<td>Costly; logistically complex; requires carefully judged survey interval and follow-up of individuals.</td>
</tr>
<tr>
<td>TB mortality</td>
<td>Direct measure of TB burden accounting for a high proportion of DALYs; case fatality falls quickly in a new control programme; MDG indicator.</td>
<td>Component due to incidence changes slowly in response to control; hard to reduce case fatality further in low-burden countries.</td>
</tr>
<tr>
<td>From observations on patient cohorts</td>
<td>Direct observation of number of patients dying.</td>
<td>Deaths observed are those in cohort only, not in the population at large, and not beyond the period of cohort follow-up; deaths among defaulters and transfers usually unknown; TB not always the cause of death for patients on TB treatment.</td>
</tr>
<tr>
<td>From product of incidence and case-fatality rate</td>
<td>Simple and widely applicable.</td>
<td>Relies on accurate measures of incidence (above) case-fatality rate and case fatality; case fatality measurable in observed DOTS cohorts, but not among patients treated elsewhere or untreated. Approximate at best.</td>
</tr>
<tr>
<td>From vital (death) registrations (VR)</td>
<td>Direct measure of TB deaths and trends; can be reported annually or sub-annually.</td>
<td>VR does not yet exist in many high-burden countries (notably in Africa and Asia); typically underestimates TB deaths; sensitivity and specificity untested.</td>
</tr>
<tr>
<td>From verbal autopsy (VA)</td>
<td>Review of registered deaths can improve accuracy of cause of death statistics.</td>
<td>Sensitivity and specificity of VA not fully evaluated; where no death registration system exists, laborious to compile deaths from a rare disease, and requires large sample sizes.</td>
</tr>
</tbody>
</table>
With a view to monitoring progress towards the MDGs (2015 compared with 1990), and progress due to the implementation of GP2 (2015 compared with 2005), some countries have already carried out surveys that effectively set baselines. China has done prevalence surveys in 1990 and 2000. Indonesia's 2004 prevalence survey gave results that can be compared against surveys done in the early 1980s. Both sets of surveys show significant reductions in prevalence. Cambodia carried out a well-executed prevalence survey in 2002, with results that are still to be published. India successfully carried out a nationwide tuberculin survey between 2000 and 2003, and the Tuberculosis Research Centre (TRC, Chennai) is closely and regularly monitoring changes in TB infection, incidence, prevalence and deaths in the "Model DOTS Project" in Tamil Nadu. National or sub-national disease prevalence surveys are planned or under way in Armenia, Bangladesh, Eritrea, Myanmar (selected provinces) and Viet Nam. A tuberculin survey will be done in Nepal. Some of these surveys have been funded through GFATM grants.

A growing number of general mortality surveys based on verbal autopsy are providing more data about TB deaths. VA surveys linked to studies of tobacco smoking have been carried out in India and South Africa. Surveys of death by cause have also been done in Tanzania (TEHIP study), Brazil (Rio de Janeiro), and more will be done in India (Registrar General of India prospective study of one million deaths) and Indonesia.

**Studies of the impact of TB control**

Despite much circumstantial evidence, there are very few recent, persuasive, studies of the epidemiological impact of TB control. WHO and collaborators have carried out detailed studies of the impact of DOTS in Peru and China. But there has been no obvious, large-scale impact in some countries that have had good DOTS programmes for several years (India, Viet Nam). Doubts therefore remain, not just about DOTS impact in specific countries, but also about whether the principles of control by chemotherapy, derived largely from observations made in Europe and North America, apply in high-burden countries today. More studies of impact are urgently needed to dispel doubts on both levels.

**Current mechanisms for standardization of procedures and for improvement of surveillance systems**

Procedures for TB surveillance have been standardized through the joint work of the technical agencies represented on this Task Force, plus many other collaborators. Modifications to these procedures are made through consultations with all interested parties, and coordinated by WHO. The procedures are made operational through the WHO-recommended recording and reporting forms used at district level (now under revision), and through the mechanism of international reporting to WHO. Recommended procedures are reinforced through the use of software packages that require standard input data, such as WHO's STAR and EpiCentre and CDC's BOTUSA (or its extension, the Electronic TB Register).
As an aid to standardizing data collection procedures, the UN Statistical Division has begun to assess the quality of data used to describe progress towards the MDGs. The methods are crudely applicable to TB, and a refined process for certifying the quality of TB data is highly desirable.

Beyond the standardization of procedures, other mechanisms for strengthening surveillance include the newly-established Health Metrics Network (HMN), a partnership based in WHO, which aims "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 (HIS)". It seems possible, in principle, that TB monitoring could be improved in the context of HMN, provided this contributes to the process of strengthening health information overall.

Training courses also strengthen monitoring and evaluation. WHO and collaborating centres (e.g. Sondalo), IUATLD, JIT, USAID and CDC all run training courses to improve various aspects of surveillance and epidemiology, though these are not coordinated.

**Cost of measurement**

Funding for all the above activities has come from various sources. USAID provides core funding for WHO's work on monitoring and evaluation, and supports M&E courses run with Measure Evaluation. BMGF funds the Health Metrics Network, and population surveys are now being done with grants from GFATM. It is part of the remit of this Task Force to consider sources of funding to implement its recommendations on measurement. This may include discussion of how this money might complement funding that is available within countries, and funding that is available for broader initiatives in health measurement (beyond TB).

However, no assessment had yet been made of the cost of measuring progress towards the MDGs. The scale of investment needed clearly depends on what needs to be done, which is the main concern of this Task Force. However, the costs can be divided into funds required to carry out surveys and surveillance in countries, and the funds required for external technical assistance (TA). The members of this Task Force have, collectively, much experience of budgeting for both kinds of activities. To facilitate the budgeting exercise, it will be extremely useful to bring together all available data on costs, as the overarching plan for measurement is being developed.
ANNEX. CHARACTERISTICS OF KEY EPIDEMIOLOGICAL INDICATORS, AND OF THE METHODS USED FOR THEIR MEASUREMENT

Case detection rate and incidence

The case detection "rate", as defined by WHO (based on the early work of Styblo), is the ratio of cases reported in any one year to the estimated incidence in that year. The principal measure used by WHO refers to smear-positive cases only (though WHO also reports estimates of case detection for all forms of TB).

The numerator of this ratio is questionable in (a) its focus on smear-positive cases, to the exclusion of smear-negative in some settings, and (b) whether case reports are biased by the need to satisfy international targets, or by the provision of free treatment for smear-positive but not for smear-negative patients. Far more controversial is the denominator of the case detection rate, which requires accurate measures of incidence.

At present, under-reporting is likely to be a more important deficiency than over-reporting; that is, case detection rates are probably well below "100%" in many high-burden countries. The following approaches can be used, in principle, to assess incidence and case detection rates:

(a) Direct measurement of case detection. This can be done through simple inventory methods (e.g. noting the number of reporting units that actually submit reports, or by comparing reports from different sources), by calculating standard reference indicators (e.g. ≈ 70% of pulmonary TB patients should be sm+), or by more formal techniques such as capture-recapture analysis. These methods will usually give at least an approximate estimate of case detection, and help to strengthen the surveillance system in the process, with long-term benefits. Capture-recapture methods have infrequently been used for studies of TB, but probably have a great deal more to offer.

(b) Incidence measured in longitudinal, population-based surveys. Surveys should follow individuals in a population to monitor the number that develop TB through time. The interval between surveys must be neither too short (costly, logistically demanding) nor too long (missing cases arising), and the study population must be sufficiently large to guarantee precision. Such surveys are rarely done because incidence rates are low (measured per 100,000 population, even in areas where TB is considered to be highly-endemic).

(c) Incidence measured from routine case reports. The ultimate, practical method for monitoring incidence, en route to elimination, is to ensure that all countries have surveillance systems with complete coverage; that is, detecting (nearly) 100% of TB patients as they become ill. All countries now have functioning TB surveillance systems, though they are currently highly variable in quality. Provided case notifications are

---

2 Because the case detection rate is actually a ratio, its value can exceed 100%, for reasons explained in Global Tuberculosis Control, WHO Report 2005.
reported consistently, they will at least give an accurate assessment of trends. In general, the assessment of trends is more reliable than the assessment of absolute incidence.

Three further methods of estimating incidence, in principle, are:

(d) Incidence of smear-positive TB estimated from annual risk of infection, derived from tuberculin surveys, using the Styblo rule (1% ARI = 50 sm+/100,000/yr).

(e) Incidence estimated from the ratio of prevalence ($P$, single cross-sectional survey) to the duration of illness, $d$ ($I = P/d$).

(f) Incidence derived from the ratio of the TB death rate, $D$, to the TB case fatality rate, $f$ ($I = D/f$).

All three require data from surveys and surveillance, to be obtained by methods discussed below.

**Treatment success**

Treatment outcomes are routinely reported in DOTS cohorts. The established system should clearly be used as the basis for modifications that will allow for the monitoring of patients carrying drug-resistant bacilli or who are co-infected with HIV. There are sources of error in the DOTS system, but these are typically small, and it is often clear how they can be corrected. Treatment success is defined as cure + completion, but completion is an unsatisfactory outcome in populations where drug resistance is frequent. The period of observation can be extended to follow longer courses of treatment for patients carrying drug-resistant mycobacteria. The fate of patients lost to follow-up (defaulted, transferred, diagnosed but not registered, registered but not evaluated) is unknown, so these outcomes must be minimized. Outside DOTS cohorts, the biggest problem is the failure to provide any evaluation of treatment outcomes. Biases are possible where there are strong incentives to meet targets, so reports need to be closely verified. Because treatment outcomes are likely to vary among different kinds of patients, it will be appropriate to maintain disaggregated data in some settings, e.g. for patients who are coinfected with HIV, or carrying drug resistant bacilli.

**Prevalence of active TB**

Prevalence is a relatively responsive indicator of the impact of chemotherapy, which typically changes faster than incidence. There are usually more prevalent cases at any one time than incidence cases in any one year. The prevalence of smear-positive and bacteriologically positive TB can be measured accurately in population-based surveys (usually cluster randomized), though precision is always compromised by the sample size needed for a rare disease like TB. An additional virtue of prevalence surveys is that they can be used to evaluate the quality of surveillance by identifying which patients with active disease have already been captured by the routine reporting system. Questionnaires can be used to explore the reasons why some patients are diagnosed and treated for TB
while others are not. In general, prevalence surveys provide a platform upon which to explore the interactions between patients and the health system, and the links between TB and other social and economic factors (potential determinants of health). Some questionnaire surveys have attempted to find out the duration of an episode of active TB (or the distribution of durations), with a view to estimating incidence from prevalence (see above). The results are of doubtful accuracy (patients typically underestimate how long they have been ill).

In reality, because of cost and logistics (due to the complexity of current methods), it seems unlikely that all 22 high-burden countries will be able to do two or more prevalence surveys to measure progress between 2005 and 2015. In some countries, it is possible to estimate prevalence nationally by calibrating the relation between risk of infection and prevalence, where both have been measured together locally (e.g. in the "Model DOTS Project" in south India).

Other population-based health survey platforms in existence (e.g. Demographic and Health Surveys, DHS) have not been used to measure prevalence because the definitive diagnosis of TB requires bacteriology (sputum smear microscopy or culture). In this context, a new, easy-to-use, sensitive and specific diagnostic tool that relies on a biomarker in e.g. in saliva or urine (preferably not blood) would allow a huge expansion of TB survey work within e.g. the DHS framework.

**Prevalence and risk of *M. tuberculosis* infection**

Tuberculin surveys are cheaper than disease prevalence surveys and less complicated logistically. However, it is often hard to determine the risk of infection from tuberculin data where infection rates are low and BCG coverage is high because the data on skin test indurations do not clearly distinguish between positive and negative children (i.e. the proportion of positives in a population level; it is still harder to determine whether individual children are infected). With the past five years, tuberculin surveys succeeded in India but failed in Cambodia, China and, arguably, Afghanistan. The use of standard cut-off points or antimodes gives results of doubtful validity because of the variable performance of tuberculin between surveys (due to variations in batch and field methods). Tuberculin surveys may be more generally useful in evaluating changes in the risk of infection through time, rather than the absolute risk. In addition, the Styblo rule may not apply where drug treatment is extensive (notwithstanding a recent example to the contrary, from India), and where many TB patients are co-infected with HIV.

**Death from TB in the whole population**

The accurate measurement of TB mortality is vital because, among other things, most of the burden of TB, as measured in terms of DALYs (the common currency of morbidity and mortality), is due to premature deaths of young adults. TB deaths can usually be reduced more quickly than incidence in good chemotherapy programmes.
As with case notifications, the ultimate goal is to develop an accurate system for reporting TB deaths, as part of vital registration (VR). Approximately 80 countries report TB deaths more or less accurately through VR, and WHO uses these death reports as estimates of TB mortality. With the exception of South Africa, no countries in Africa or Asia routinely report TB deaths (or do not have data that we are aware of). In countries elsewhere that do report TB deaths from VR, the accuracy of these reports is commonly unknown.

Some countries that do not have VR have begun to carry out mortality surveys, with death certifications further evaluated by verbal autopsy (VA, e.g. China, India). In general, surveys of the cause of death require large samples to get accurate results for TB, because a TB death is a rare event. VA as a technique for counting TB deaths has not yet been thoroughly validated, though validation studies are under way, e.g. in India.

In countries where TB deaths are not counted, or not reliably counted, they are calculated approximately from the product of estimated incidence and case fatality rates. The method is simple in principle, but requires accurate estimates of incidence and case fatality. Typically, the case fatality rate is poorly known outside treatment cohorts under observation.

The measurement of deaths within observed cohorts is important in assessing the performance of DOTS programmes. Ultimately these cohorts should include nearly all TB cases arising in any country, so that cohort outcomes converge more closely with national death registrations. However, convergence will never be complete unless the number of patients lost to follow-up (default, transfer) can be reduced to zero, and unless the period of cohort follow-up is extended well beyond six months. A smaller difficulty is that deaths during treatment are not always due to TB, but are nonetheless attributed to TB in the DOTS cohort system.