ANNEX 1

Methods used to estimate the burden of disease caused by TB

This annex explains the methods that were used to produce estimates of the global burden of disease caused by TB (measured in terms of incidence, prevalence and mortality). It has nine major sections:

- **General approach.** This section provides some background information about the methods used to produce estimates of disease burden.
- **Definitions.** This section defines TB incidence, prevalence and mortality, the case fatality rate (CFR) and the case notification rate. It also explains the regions for which estimates of disease burden are produced and sources of information on population estimates.
- **Estimates of TB incidence, 1990–2010.** This section explains the main methods used to estimate TB incidence, and the countries for which they have been applied. Specific attention is given to estimates for China and India.
- **Estimates of HIV prevalence among incident TB cases, 1990–2010.** This section explains the methods used to estimate the prevalence of HIV among incident cases of TB.
- **Estimates of TB prevalence, 1990–2010.** This section explains the methods used to estimate TB prevalence. These are national surveys of the prevalence of TB disease and indirect estimates based on combining estimates of incidence with estimates of the duration of TB disease.
- **Estimates of the number of cases of multidrug-resistant TB (MDR-TB).** This section explains how estimates of the proportion of notified cases of TB that had MDR-TB in 2010 were produced and used to assess the number of prevalent cases of MDR-TB in 2010. Methods to analyse trends in the proportion of new cases of TB with MDR-TB among notified cases 1994–2010 are also explained.
- **Estimates of TB mortality, 1990–2010.** This section explains the two methods used to estimate TB mortality. These are direct measurements from vital registration (VR) or survey data and indirect estimates based on combining estimates of TB incidence with estimates of the CFR. The countries for which these methods have been used are explained. Methods for estimating TB mortality in HIV-infected individuals and TB mortality by age and sex are also described.
- **Projections of TB incidence, prevalence and mortality.** This section explains how projections up to 2015 were produced.
- **Uncertainty framework.** This section explains the general approach to including uncertainty in all estimates.

1. General approach

Estimates of the burden of disease caused by TB (measured in terms of incidence, prevalence and mortality) are produced annually by WHO using information gathered through surveillance systems (case notifications and death registrations), special studies (including surveys of the prevalence of disease and in-depth analyses of surveillance data), expert opinion and consultations with countries. Two recent publications provide up-to-date guidance about how TB incidence, prevalence and mortality should be measured,[1] based on the work of the WHO Global Task Force on TB Impact Measurement.[2] The methods used to estimate the burden of disease were updated in 2009 following 18 months of work by an expert group convened by the Task Force. Improvements to methods included systematic documentation of expert opinion and how this has been used to produce estimates of disease burden, simplification of models,[3] updates to parameter values based on the results of systematic reviews, much greater use of mortality data from VR systems and systematic documentation of uncertainty (hence the uncertainty intervals shown on all of the estimates of disease burden in this report).

2. Definitions

2.1 Incidence, prevalence, mortality, the case fatality rate and the case notification rate

Incidence is defined as the number of new and relapse cases of TB (all forms) occurring in a given year. Relapse cases are defined as a new episode of TB in people who have had TB in the past and for whom there was bacteriological confirmation of cure and/or documentation that treatment was completed (**Box 3.1**, **Chapter 3**). In the remainder of this Annex, relapse cases are referred to as *recurrent* cases, in line with expected changes in terminology that will be introduced by WHO in the near future and because the term is more useful when explaining the estimation of TB incidence. Recurrent cases may be true relapses or a new episode of TB caused by reinfection. In current case definitions, both relapse cases and patients who require a change in treatment are called "retreatment cases". However, people with a continuing episode

[1] *TB impact measurement: policy and recommendations for how to assess the epidemiological burden of TB and the impact of TB control*. Geneva, World Health Organization, 2009 (Stop TB policy paper no. 2; WHO/HTM/TB/2009.416). The policy paper is available on the Task Force's website **www.who.int/tb/advisory_bodies/impact_measurement_taskforce**

[2] For further details, see the Task Force web site at: **www.who.int/tb/advisory_bodies/impact_measurement_taskforce**. The review is also the basis for the TB component of the update to the Global Burden of Disease, due for publication in 2011 (**www.who.int/topics/global_burden_of_disease**).

[3] For example, some parameter values are now estimated only at global level or for regions, rather than for each country individually.

of TB that requires a treatment change are prevalent cases, not incident cases.

Prevalence is defined as the number of TB cases (all forms) at a given point in time.

Mortality. According to the latest revision of the international classification of diseases (ICD-10), TB mortality is the number of deaths caused by TB in HIV-negative people. TB deaths among HIV-positive people are classified as HIV deaths in ICD-10. For this reason, estimates of deaths caused by TB in HIV-positive people are presented separately from those in HIV-negative people.

The **case fatality rate** is the risk of death from TB among people with active TB disease.[1]

The **case notification rate** refers to new and recurrent episodes of TB notified to WHO for a given year, expressed per 100 000 population. The case notification rate for new and recurrent TB is important in the estimation of TB incidence. It is important to highlight, however, that in some countries information on treatment history may be missing for some cases. When data on treatment history are not available, recurrent cases cannot be distinguished from cases whose treatment was changed, since both are registered and reported in the category "retreatment". An assessment of data for patients reported in the "unknown history" category is conducted with national TB control programmes (NTPs) to determine the proportion of such patients that is included in the category of recurrent cases.

2.2 Regions

Regional analyses are generally undertaken for the six WHO regions (that is, the African Region, the Region of the Americas, the Eastern Mediterranean Region, the European Region, the South-East Asia Region and the Western Pacific Region). For analyses related to MDR-TB, nine epidemiological regions were defined. These were African countries with high HIV prevalence, African countries with low HIV prevalence, Central Europe, Eastern Europe, high-income countries,[2] Latin America, the Eastern Mediterranean Region (excluding high-income countries), the South-East Asia Region (excluding high-income countries) and the Western Pacific Region (excluding high-income countries). The list of countries in the first six of these nine regions is provided in **Appendix 1**; the other countries are listed under the WHO regions of which they are a part in **Annex 3**.

2.3 Population estimates

Where population sizes are needed to calculate TB indicators, the 2010 revision of estimates provided by the United Nations Population Division (UNPD) was used.[3] The UNPD estimates sometimes differ from those made by countries.

3. Estimates of TB incidence, 1990–2010

No country has ever undertaken a nationwide survey of TB incidence because of the large sample sizes required and associated major logistic and financial challenges. As a result, there are no direct measurements of the incidence of TB. Theoretically, data from TB surveillance systems that are linked to health systems of high coverage and performance may capture all (or almost all) incident cases of TB. However, as yet no standard and widely-endorsed criteria and benchmarks for classifying TB surveillance systems are available. The WHO Global Task Force on TB Impact Measurement is working on the development of such standards (**Chapter 2**).

In the absence of direct measurements, estimates of TB incidence for almost all countries rely on methods described in **sections 3.1–3.4**. The methods used to estimate TB incidence in China and India are explained separately, in **section 3.5** and **section 3.6** respectively, following national workshops held in China (in June 2011) and India (in July 2011).

It should be emphasized that incidence estimates are no longer derived from surveys of the prevalence of tuberculous infection as measured in tuberculin surveys. The WHO Global Task Force on TB Impact Measurement has agreed that methods for deriving incidence from the prevalence of infection are unreliable. The Task Force has also stated that it is doubtful whether repeat tuberculin surveys provide a reliable estimate of the trend in TB incidence.[4]

3.1 Estimating TB incidence from estimates of the proportion of cases detected

Notification data for new and recurrent cases have been analysed in combination with evidence about the coverage of the TB surveillance system and expert opinion in six regional workshops and country missions held during the period 2009–2011, according to a framework developed by the WHO Global Task Force on TB Impact Measurement (**Figure 2.2**, **Chapter 2**). By mid-2011, these workshops and country missions had covered 96 countries (**Figure 2.1**, **Chapter 2**).

For the 96 countries covered by these regional workshops and country missions, incidence was estimated according to the following equation:

---

[1] Straetemans M et al. Assessing tuberculosis case fatality ratio: a meta-analysis. *PLoS One*. 2011, 6(6):e20755.

[2] High-income countries are defined by the World Bank as countries with a per capita gross national income (GNI) of US$ 12 276 or more in 2010.

[3] **http://esa.un.org/unpd/wpp/unpp/panel_population.htm**; accessed August 2011.

[4] *TB impact measurement: policy and recommendations for how to assess the epidemiological burden of TB and the impact of TB control*. Geneva, World Health Organization, 2009 (Stop TB policy paper; no. 2 (WHO/HTM/TB/2009.416).

$$\text{incidence} = \frac{\text{case notifications}}{1 - \text{underreporting}}$$

Expert opinion about the proportion of TB cases[1] that were not reported was elicited for three reference years (1997, 2003 and, depending on when the workshop was held, either 2008 or 2009). This was done following in-depth analysis of notification data (including data from sub-national administrative levels), programmatic data reflecting efforts in TB control (for example, data on infrastructure, staffing, the performance of services and funding) and (where available) data from inventory studies.[2] In addition, data on access to health care from Demographic and Health Surveys and the overall performance of health systems (using indicators such as the infant mortality rate) were used to substantiate opinion on the proportion of cases with no or very limited access to health care (**Table A1.1**).

A full description of the methods used in these workshops is available in a report of the workshop held for countries in the African Region (in Harare, Zimbabwe, December 2010).[3]

TABLE A1.1

Sources of information and data on TB incidence used in regional workshops and country missions

POSSIBLE CATEGORIES OF INCIDENT CASES	SOURCES OF DATA	
Do not have physical or financial access to health care	Demographic and health surveys, KABP[a] surveys	Capture–recapture modelling
Seek care, but TB not diagnosed	Survey	
TB diagnosed, but not reported	"Inventory" survey	
Reported cases	TB surveillance	

[a] KABP = knowledge, attitudes, behaviour and practices.

Distributions of the proportion of cases that were not reported in the three reference years were assumed to follow a Beta distribution. Reasons for using Beta distributions include the following:

- They are continuous and defined on the interval (0, 1). Since the variance of the proportions of cases that were not reported tend to be large as a result of high uncertainty, random draws of numbers from a normal distribution would yield numbers outside the interval (0, 1). The use of truncated normal distributions may result in excess density towards one of the bounds.
- They are not necessarily symmetrical.
- They are defined with two parameters that can be estimated from available data using the method of moments.[4]

The shape and scale parameters necessary to define the Beta distribution were computed using the method of moments, as follows:

First, the variance for the distribution was taken as:

$$V = \left(\frac{u - l}{4} \right)^2$$

where l and u are the lower and upper bounds of the plausible range for the proportion of incident cases that were reported (also referred to as the case detection rate in **Chapter 3**).

Shape 1 (noted α) and 2 (noted β) follow from:

$$s = \frac{E(1 - E)}{V} - 1$$

$$\alpha = sE$$

$$\beta = s(1 - E)$$

where E is the expected value of the istribution (**Table A1.2**).

Time series for the period 1990–2010 were built according to the characteristics of the levels of underreporting that were estimated for the three reference years. A cubic spline extrapolation of V and E, with knots set at the reference years, was used for countries with low-level or concentrated HIV epidemics. In countries with a generalized HIV epidemic, the trajectory of incidence from 1990 to the first reference year (usually 1997) was based on the annual rate of change in HIV prevalence. Incidence trajectories were derived from the series of notified TB cases using Monte Carlo simulations from which expected values, 2.5th and 97.5th centiles were extracted. All computations were conducted in the R statistical environment.[5]

If there were insufficient data to determine the factors leading to time-changes in case notifications, incidence was assumed to follow a horizontal trend going through the most recent estimate of incidence.

[1] Defined as cases of all forms of TB, including sputum smear-positive pulmonary cases, sputum smear-negative pulmonary cases, and extrapulmonary cases.

[2] Measurements from "inventory" studies can be used to quantify the number of cases that are diagnosed but not reported to national surveillance systems. In some circumstances, data from these studies can be used to estimate the number of cases that were not diagnosed as well (using capture-recapture methods. A useful reference on capture-recapture methods is: Chao A et al. The applications of capture-recapture models to epidemiological data. *Statistics in Medicine*, 2001, 20(20):3123–3157.

[3] See **www.who.int/tb/advisory_bodies/impact_measurement_taskforce**. The tools (called TISAT and the Workbook) used in regional workshops and country missions are also available on the Task Force's web site.

[4] Rényi A. *Probability theory*. New York, Dover Publications Inc., 2007.

[5] R Development Core Team. *R: a language and environment for statistical computing*. Vienna, R Foundation for Statistical Computing, 2009 (**www.R-project.org**).

TABLE A1.2

Parameter estimates used to produce estimates of TB incidence, prevalence and mortality

MODEL PARAMETER	DISTRIBUTION	DISTRIBUTION PARAMETERS[b]
Incidence, high-income countries	Beta[a]	$\alpha = \bar{I} \cdot \left[\frac{\bar{I}(1-\bar{I})}{V} - 1 \right]$ $\beta = (1-\bar{I}) \cdot \left[\frac{\bar{I}(1-\bar{I})}{V} - 1 \right]$ *where $\bar{I}$ was set at 1.3 times the notification rate, noted N, and V is defined by:* $V = \left[\frac{0.3}{4} N \right]^2$
HIV prevalence among incident TB	Beta[a]	$\alpha = \bar{x} \cdot \left[\frac{\bar{x}(1-\bar{x})}{V} - 1 \right]$ $\beta = (1-\bar{x}) \cdot \left[\frac{\bar{x}(1-\bar{x})}{V} - 1 \right]$ *Where $\bar{x}$ is the expected value and V is given by:* $V = \left[\frac{u-l}{4} \right]^2$
Duration of disease, non-notified HIV-negative cases of TB	Uniform	l = 1, u = 4 (years)
Duration of disease, non-notified HIV-positive cases of TB	Uniform	l = 0.01, u = 0.2 (years)
Duration of disease, notified HIV-negative cases of TB	Uniform	l = 0.2, u = 2 (years)
Duration of disease, notified HIV-positive cases of TB	Uniform	l = 0.01, u = 1 (years)

[a] The probability density function of the Beta distribution is: $f(x; \alpha, \beta) = \frac{x^{\alpha-1}(1-x)^{\beta-1}}{\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\,dt}$

[b] u and l denote upper and lower bounds.

3.2 Estimating TB incidence from data on case notifications and expert opinion for high-income countries

For high-income countries, the level of TB incidence was assumed to be distributed between the notification rate for new and recurrent cases combined (lower uncertainty bound, noted *l*) and 1.3 times the notification rate (upper uncertainty bound, noted *u*), as informed by expert opinion. The distribution of incidence was assumed to follow a Beta distribution with shape and scale parameters computed using the method of moments, as described above.

In the absence of country-specific data on the quality and coverage of TB surveillance systems, it was assumed that TB surveillance systems from countries in the high-income group performed similarly well, although the model does allow for stochastic fluctuations. The exception was the United Kingdom of Great Britain and Northern Ireland, where the underreporting of TB cases has been recently measured using inventory studies and capture-recapture modelling.[1] The results were used to measure TB incidence directly.

3.3 Estimating TB incidence from empirical measurements of disease prevalence

Incidence can be estimated using measurements from national surveys of the prevalence of TB disease combined with estimates of the duration of disease. Incidence is estimated as the prevalence of TB divided by the average duration of disease.

In practice, the duration of disease cannot be directly measured. For example, measurements of the duration of symptoms in prevalent TB cases that are detected during a prevalence survey are systematically biased towards lower values, since active case-finding truncates the natural history of undiagnosed disease. Measurements of the duration of disease in notified cases ignore the duration of disease among non-notified and untreated cases.

Literature reviews commissioned by the WHO Global Task Force on TB Impact Measurement have provided estimates of the duration of disease in untreated TB cases

[1] *Tuberculosis in the UK: annual report on tuberculosis surveillance in the UK 2010*. London, Health Protection Agency Centre for Infections, 2010 (**www.hpa.org.uk/web/HPAweb&HPAwebStandard/HPAweb_C/1287143581697**; accessed 15 July 2011.

from the pre-chemotherapy era (before the 1950s). The best estimate of the mean duration of disease (for smear-positive cases and smear-negative cases combined) in HIV-negative individuals is about three years. However, the proportion of incident cases that remain untreated is unknown. There are few data on the duration of disease in HIV-positive individuals.

When measurements from two prevalence surveys were available, trends in TB prevalence were derived by fitting a log-linear model to available measurements. When three or more prevalence measurements were available, the prevalence trajectory was built using cubic spline interpolation. If only one prevalence survey measurement was available, time-trends were assessed using in-depth analysis of surveillance data, as described above.

In this report, the prevalence to incidence method was used for only one country (Viet Nam), following a meeting in early 2009 in which consensus was reached among national experts and experts from WHO and the KNCV Tuberculosis Foundation.

3.4 Estimating TB incidence from previously published time-series of incidence

In all remaining countries (n=57), previously published time-series of TB incidence were extended by fitting a log-linear model to the estimates for 2006–2009, to predict a value for 2010.

3.5 TB incidence in China

As noted at the beginning of this section, if TB surveillance performs to high standards then the best source of information on TB incidence comes from routine notification data. In China, there is a web-based and case-based mandatory TB reporting system that has been fully operational since 2005. It covers very close to 100% of all detected TB cases.

During a national workshop held in Beijing in June 2011, incidence was estimated in two stages. First, the plausible interval for TB incidence in 2009 was set at 1–1.3 times the level of notifications, which is comparable to most high-income countries. This plausible interval was justified based on the observations that (i) the ratio of TB mortality to TB notifications in China was close to that observed in high-income countries; and (ii) the performance of the TB surveillance system is high. Second, trends in incidence were computed backwards in time from 2009 to 1990 and forwards in time to 2010 based on measured trends in rates of TB mortality and TB prevalence in adults, adjusted for the rapidly aging population. The workshop estimated that incidence in adults declined by 3–5% per year on average, with an age-adjusted decline of 3.4% per year (standard deviation, 0.58%). This estimate can be considered conservative given that the decline in TB prevalence is under-estimated (see **Box 2.6** in **Chapter 2** for further explanation).

3.6 TB incidence in India

Incidence for 2010 was estimated according to the methods described in **section 3.1**, including use of results from two subnational inventory studies. The level of underreporting for 2010 was estimated at 41% (range, 35–47%). National inventory studies will be needed to fully understand the extent to which TB cases are diagnosed in the private sector but not reflected in the national surveillance system.

In the absence of any clear trend in case notifications and no survey measurements taken before 2001, the trend in incidence was estimated to be flat between 1990 and 2001. This was also justified on the basis that implementation of the Revised TB Control Programme in India only began in parts of the country in 1999, with no evidence of improvements in TB control in the previous decade.

For the trend between 2001 and 2010, data from tuberculin surveys and notification data were used. Two national tuberculin surveys were conducted around 2000 and 2010. Despite difficulties in interpreting the second survey as a result of unfavourable distributions of reaction sizes as well as systematic differences between the two surveys (such as use of different tuberculins), the estimated decline in the annual risk of infection was estimated at 3.7% per year (95% confidence interval, 2.4–5.1% per year). This rate of decline was the basis for setting a prior beta distribution for the decline in incidence.

In districts with early implementation (1999–2003) of the Revised TB Control Programme, the average annual rate of decline in case notification rates varied between 0.6% and 3.6% per year. Combining the previous estimate from the tuberculin survey data with the observed notification data led to a posterior distribution of the annual rate of decline in TB incidence that had an expected value of 1.5% per year (standard deviation, 0.071).

3.7 Disaggregations of TB incidence

In this report, TB incidence is only disaggregated by HIV-infection status (see following section). The estimation of smear-positive TB incidence was discontinued in 2010, for reasons explained in detail in the global report published in 2010.

4. Estimates of HIV prevalence among incident TB cases, 1990–2010

The prevalence of HIV among incident cases of TB was directly estimated from country-specific and empirical data wherever possible. For the estimates published in this report, suitable data (as defined in **Table A1.3**) were available for a total of 544 country-year data points, up from 440 country-year data points in the previous year.

For the 3905 country-year data points for which surveillance data were either not available or for which the

TABLE A1.3

Source of data on HIV prevalence among incident TB cases

DIRECT MEASUREMENT OF THE PREVALENCE OF HIV IN TB PATIENTS	NUMBER OF COUNTRY-YEARS
National surveys	31
HIV sentinel surveillance	30
Provider-initiated testing and counselling with at least 50% coverage of testing	483
Total	**544**

percentage of TB patients tested for HIV was below 50%, the prevalence of HIV was estimated indirectly according to the following equation:

$$t = \frac{h\rho}{1 + h(\rho - 1)}$$

In this equation, t is HIV prevalence among incident TB cases, h is HIV prevalence among the general population (from the latest time-series provided by UNAIDS) and ρ is the incidence rate ratio (IRR) (defined as the incidence rate of TB in HIV-positive people divided by the incidence rate of TB in HIV-negative people).[1] We then let logit(t) be log(t/(1-t)) and logit(h) be log(h/(1-h)). Using data from countries where HIV prevalence has been estimated by UNAIDS as an independent variable, a linear model of logit-transformed t was fitted using logit-transformed h according to the following equation, written in matrix notation:

$$\hat{T} = X\beta$$

where $\hat{T}$ is a vector of predicted logit(t), X is an n x 2 matrix in which the first column holds 1s, and the second column holds logit(h). The vector β holds estimated model parameters.

Models were run using Monte Carlo simulations in which h was drawn randomly from a Beta distribution with shape parameters computed as described in **section 3.1**, (low and high uncertainty bounds are provided by UNAIDS – also see **Table A1.2**). The model was run 50 000 times using country-specific distributions for H and T (noted in capital letters to denote vectors or matrices) based on their uncertainty intervals. The uncertainty bounds for β were chosen as the 2.5th and 97.5th centiles.

The source of data used for each country is available upon request from tbdata@who.int.

5. Estimates of TB prevalence, 1990–2010

The best way to measure the prevalence of TB is through national population-based surveys of TB disease.[2,3] Data from such surveys are available for an increasing number of countries (**Chapter 2**). It should be noted, however, that measurements of prevalence are typically confined to the adult population. Furthermore, prevalence surveys exclude extrapulmonary cases and do not allow the diagnosis of cases of culture-negative pulmonary TB.

When there is no direct measurement from a national survey of the prevalence of TB disease, prevalence is the most uncertain of the three TB indicators used to measure disease burden. This is because prevalence is the product of two uncertain quantities: (i) incidence and (ii) disease duration. The duration of disease is very difficult to quantify because it cannot be measured during surveys of the prevalence of TB disease (surveys truncate the natural history of disease). Duration can be assessed in self-presenting patients, but there is no practical way to measure the duration of disease in patients who are not notified to NTPs.

Indirect estimates of prevalence were calculated according to the following equation:

$$P = \sum I_{i,j} d_{i,j},\ i \in \{1,2\},\ j \in \{1,2\}$$

where the index variable i denotes HIV+ and HIV–, the index variable j denotes notified and non-notified cases, d denotes the duration of disease in notified cases and I is total incidence. In the absence of measurements, we did not allow duration in notified cases to vary among countries. Given their underlying uncertainty, prevalence estimates should be used with great caution in the absence of direct measurements from a prevalence survey. Unless measurements were available from national programmes (for example, Turkey), assumptions of the duration of disease were used as shown in the last four rows of **Table A1.2**.

6. Estimates of the number of cases of MDR-TB

6.1 Proportion of notified cases of TB that have MDR-TB, 2010

Global and regional estimates of the proportion of new and retreatment cases of TB that had MDR-TB in 2010 were calculated using country-level information. If countries had reported data on the proportion of new and retreatment cases of TB that have MDR-TB from routine surveillance or a survey of drug resistance the latest available information was used. For countries that have not reported such data, estimates of the proportion of new and retreatment cases of TB that have MDR-TB were produced using modelling (including multiple imputation) that was based on data from countries for which data do exist. Estimates for countries without data were based on countries that were considered to be

[1] **www.unaids.org/en/dataanalysis/epidemiology/**, accessed 15 July 2011.

[2] Glaziou P et al. Tuberculosis prevalence surveys: rationale and cost. *International Journal of Tuberculosis and Lung Disease*, 2008, 12(9):1003–1008.

[3] *TB prevalence surveys: a handbook*. Geneva, World Health Organization, 2011 (WHO/HTM/TB/2010.17).

similar in terms of TB epidemiology. The observed and imputed estimates of the proportion of new and retreatment cases of TB that have MDR-TB were then pooled to give a global estimate, with countries weighted according to their share of global notifications of new and retreatment cases.

6.2 Trends in the proportion of new TB cases with MDR-TB, 1994–2010

Analysis of trends in the proportion of TB cases that have MDR-TB was restricted to new cases. Data were too patchy to allow analysis of trends in retreatment cases.

Countries or territories for which there were at least two measurements of the proportion of new TB cases that had MDR-TB between 1994 and 2010 were identified. A linear regression model of the log-transformed proportion of cases that have MDR-TB was fitted for every country, with only year as an independent variable. The restricting assumption of a linear association between changes in the proportion of cases that have MDR-TB and time was made because of the small number of measurements per country. The slope of this regression model represents the annual change in the proportion of new cases with MDR-TB. Modelling including multiple imputation was then used to produce estimates of the annual change in the proportion of cases that have MDR-TB for countries that have not reported data. Estimates for countries without data were based on countries to which they were considered to be similar in terms of TB epidemiology (see **Appendix 1** and **section 2.2**). Finally, the observed and imputed estimates were pooled to give global and regional estimates, with countries weighted according to their share of global notifications of new cases.

6.3 Numbers of prevalent cases of MDR-TB, 2010

The global estimate of the number of prevalent cases of MDR-TB in 2010 was derived in two steps. First, the weighted average of the proportion of new and retreatment notified cases that had MDR-TB was computed, to give an estimate of the proportion of all notified cases that had MDR-TB. This combined proportion was then multiplied by the estimated global prevalence of TB in the general population, under the assumption that the proportion of all cases that have MDR-TB was the same as the proportion of notified cases that have MDR-TB.

Country-specific estimates of the number of prevalent cases of MDR-TB in 2010 were not computed because only a few countries have directly measured the prevalence of TB in a population-based survey, and even among these countries data on the proportion of culture-positive pulmonary cases that had MDR-TB are not always available. To date, direct measurements of the number of prevalent cases of MDR-TB are available only for China, although several upcoming surveys will include assessments of drug resistance. In the absence of direct measurements at country level, country-specific estimates of the prevalence of MDR-TB suffer from much greater uncertainty compared with the uncertainty that surrounds global averages.

7. Estimates of TB mortality, 1990–2010

The best sources of data about deaths from TB (excluding those among HIV-positive people) are VR systems in which causes of death are coded according to ICD-10 (although the older ICD-9 and ICD-8 classification are still in use in several countries). Deaths from TB in HIV-positive people are coded under HIV-associated codes.

Estimates of TB mortality were produced directly from VR data or mortality surveys, or indirectly from estimates of TB incidence and case-fatality rates (CFRs). The source of data used in each country is available from tbdata@who.int upon request.

7.1 Estimating TB mortality from vital registration data and mortality surveys

Data from VR systems are reported to WHO by Member States and territories every year. In countries with functioning VR systems in which causes of death are coded according to the two latest revisions of the international classification of diseases (underlying cause of death: ICD-10 A15-A19, equivalent to ICD-9: 010-018), VR data are the best source of information about deaths from TB among people not infected with HIV. When people with AIDS die from TB, HIV is registered as the underlying cause of death and TB is recorded as a contributory cause. Since one third of countries with VR systems report to WHO only the underlying causes of death and not contributory causes, VR data usually cannot be used to estimate the number of TB deaths in HIV-positive people.

In 2010, 92 countries had well-functioning VR systems according to the following definition: (i) coverage of at least 70% of the population, and (ii) ill-defined causes of death (ICD-9 code B46, ICD-10 codes R00-R99) of <20% of all registered deaths.[1] In addition, mortality survey data from two countries were used (China and India), of which one (India) did not have VR data. Countries with mortality measurements included 6 of the 22 HBCs (Brazil, China, India, the Philippines, the Russian Federation and South Africa). However, we could not use the VR data on TB deaths from South Africa because large numbers of HIV deaths were miscoded as TB deaths.

Among the remaining 91 countries, there was a median of 9 years (interquartile range, 6–11) of VR data on TB mortality between 1991 and 2010 that met the above criteria, equivalent to 720 country-years. We assumed that

[1] Mathers CD et al. Counting the dead and what they died from: an assessment of the global status of cause of death data. *Bulletin of the World Health Organization*, 2005, 83:171–177.

the proportion of TB deaths among deaths not recorded by the VR system was the same as the proportion of TB deaths in VR-recorded deaths. For VR-recorded deaths with ill-defined causes, we assumed that the proportion of deaths attributable to TB was the same as the observed proportion in recorded deaths. We assumed errors in measurement (due to misclassifications) and assumptions (redistributions) to be binomially distributed.

7.2 Estimating TB mortality from indirect estimates of case-fatality rates and TB incidence

For the years in which VR or mortality survey data of sufficient quality and coverage were not available for the 91 countries defined above plus the 125 countries (as of 2010) without any direct measurement, mortality was estimated as the product of TB incidence and the CFR. CFRs were estimated separately for TB cases notified to NTPs and non-notified cases and, within these two groups, separate estimates were made for HIV-positive TB cases and HIV-negative TB cases (**Table A1.3**).

For consistency with VR – or survey-based mortality estimates, CFRs were estimated such that they gave the best fit to the directly measured TB death rates (within their uncertainty ranges) across the 720 country-years of data from the 91 countries with functioning VR systems or survey data), in conjunction with WHO estimates of distributions of TB incidence in those countries. This statistical fitting used Bayesian linear models and was done separately for two groups of countries (high-income and all other countries), to account for differences in the ratio of reported TB mortality to TB notification rates among these two groups (data not shown).

The models used normal errors and Gibbs sampling:

$$y = (I - N)\beta_1 + N\beta_2 + \varepsilon,\ \varepsilon \sim N(0,\sigma^2)$$

where y is TB mortality from VR, I denotes TB incidence excluding people living with HIV, N denotes TB notifications excluding people living with HIV, and parameters β_1 and β_2 denote the CFR in non-notified and notified cases respectively. Semi-conjugate priors were set with an uninformative inverse Gamma prior on the conditional error variance:

$$b \sim N(b_i, B_i^{-2}),\ \sigma^2 \sim IG(5.10^{-4}, 5.10^{-4})$$

Priors b and their precision B were defined based on literature reviews,[1,2] and the country-year CFR parameters used by WHO for the years 1999–2008 (**Table A1.4**). Convergence of Markov Chains was assessed graphically and using two convergence diagnostic tests. Within each case category 1990–2010, mortality estimates were computed by taking the product of posterior distributions of the CFR, assumed to be time-independent (**Table A1.4**), and country-year specific distributions of estimated incidence.

TABLE A1.4

Estimates of TB case-fatality rates by case type and country

CASE TYPE AND COUNTRY GROUP	HIV-NEGATIVE	
	NORMAL PRIOR DISTRIBUTIONS[a] MEAN (STANDARD ERROR)	POSTERIOR DISTRIBUTIONS MEAN (STANDARD ERROR)
Non-notified: high-income countries	0.1 (0.01)	0.1 (0.0097)
Non-notified: other countries	0.4 (0.01)	0.32 (0.098)
Notified: high income countries	0.04 (0.01)	0.074 (0.0026)
Notified: other countries	0.05 (0.01)	0.058 (0.006)

[a] Priors and assumed distributions in HIV-negaitive cases were derived from (i) pooled estimates from random-effects modelling of literature review results and (ii) pooled estimates from the WHO global TB database of assumed country-specific CFRs (2008).

7.3 Estimates of TB mortality among HIV-positive people

A prior belief about the proportion of HIV deaths with TB as the contributory cause of death was set on the assumption of a beta distribution with parameters a and b. The prior proportion was set at 30% (standard deviation, 3%).[3] The likelihood for the estimated number of TB deaths among estimated HIV-positive incident TB cases was based on an assumed 50% CFR (standard deviation, 5%) in low and middle-income countries and a 20% CFR (standard deviation, 2%) in high-income countries, using the methods described above and from literature reviews.[4] Cases on antiretroviral therapy (ART) were assumed to benefit from the protective effect of ART, estimated at 48% (standard deviation, 0.45%) based on a recent literature review. The likelihood was defined as a beta density with parameters s+1 and f+1. By combining the beta prior with the likelihood function, the posterior is also of the beta form with parameters a+s and b+f. Posteriors were determined for each country-year data point.

7.4 Estimating TB mortality from disaggregated estimates of TB deaths by age and sex

For countries with VR data, it was possible to disaggregate estimated TB deaths by age (with age groups defined

[1] Straetemans M et al. Assessing tuberculosis case fatality ratio: a meta-analysis. *PLoS One*. 2011, 6(6):e20755.

[2] Tiemersma EW et al. Natural history of tuberculosis: duration and fatality of untreated pulmonary tuberculosis in HIV negative patients: a systematic review. *PLoS One*. 2011, 6(4):e17601.

[3] **www.unaids.org/en/dataanalysis/epidemiology/** accessed 15 July 2011.

[4] Straetemans M et al. The effect of tuberculosis on mortality in HIV positive people: a meta-analysis. *PLoS One*, 2010, 5(12):e15241.

as 0–4 years, 5–14 years, 15–24 years, 25–34 years, 35–44 years, 45–54 years, 55–64 years, ≥65 years) and sex, in line with the way in which deaths are reported. In countries with no functional VR system, the total number of estimated TB deaths was redistributed into the different age and sex strata according to the disaggregation of the combined population of countries with VR data (with standardization against the individual country's age and sex distribution). TB deaths in HIV-positive people were not disaggregated by age and sex due to limited data from countries with functional VR systems.

8. Projections of incidence, prevalence and mortality up to 2015

Projections of TB incidence, prevalence and mortality rates up to 2015 enable assessment of whether global targets set for 2015 are likely to be achieved at global, regional and country levels. Projections for the years 2011–2015 were made using log-linear regression models fitted to data from 2007–2010, with the assumption that recent trends would continue.

9. Estimation of uncertainty

There are many potential sources of uncertainty associated with estimates of TB incidence, prevalence and mortality, as well as estimates of the burden of HIV-associated TB and MDR-TB. These include uncertainties in input data, in parameter values, in extrapolations used to impute missing data, and in the models used.

We used fixed population values from the UNPD. We did not account for any uncertainty in these values.

Notification data are of uneven quality. Cases may be underreported (missing quarterly reports from remote administrative areas are not uncommon), misclassified (in particular, misclassification of recurrent cases in the category of new cases is common), or over-reported as a result of duplicated entries in TB information systems. The latter two issues can only be addressed efficiently in countries with case-based nationwide TB databases that include patient identifiers. Sudden changes in notifications over time are often the result of errors or inconsistencies in reporting, but may sometimes reflect abrupt changes in TB epidemiology (for example, resulting from a rapid influx of migrants from countries with a high burden of TB, or from rapid improvement in case-finding efforts).

Missing national aggregates of new and recurrent cases were imputed by cubic spline interpolation. Notification trajectories were smoothed using a penalized cubic splines function with parameters based on the data. Attempts to obtain corrections for historical data are made every year, but only rarely do countries provide appropriate data corrections.

Mortality estimates incorporated the following sources of uncertainty: sampling uncertainty in the underlying measurements of TB mortality rates from data sources, uncertainty in estimates of incidence rates and rates of HIV prevalence among both incident and notified TB cases, and parameter uncertainty in the Bayesian model. Time-series of TB mortality were generated for each country through Monte Carlo simulations.

Unless otherwise specified, uncertainty bounds and ranges were defined as the 2.5th and 97.5th centiles of outcome distributions. Throughout this report, ranges with upper and lower bounds defined by these centiles are provided for all estimates established with the use of simulations. When uncertainty was established with the use of observed or other empirical data, 95% confidence intervals are reported.

The model used the following sequence: (1) incidence estimation, (2) estimation of HIV-positive TB incidence, (3) estimation of mortality, (4) estimation of prevalence. By design, some steps were independent from each other (for example, step 4 may be done before or after step 3).

The general approach to uncertainty analyses was to draw values from specified distributions for every parameter (except for notifications and population values) in Monte Carlo simulations, with the number of simulation runs set so that they were sufficient to ensure stability in the outcome distributions. For each country, the same random generator seed was used for every year, and errors were assumed to be time-dependent within countries (thus generating autocorrelation in time-series). Regional parameters were used in some instances (for example, for CFRs). Summaries of quantities of interest were obtained by extracting the 2.5th, 50th and 97.5th centiles of posterior distributions.

Appendix 1. Epidemiological regions used for analyses related to MDR-TB

Africa – countries with high HIV prevalence: Botswana, Burundi, Cameroon, the Central African Republic, the Congo, Côte d'Ivoire, the Democratic Republic of the Congo, Ethiopia, Gabon, Kenya, Lesotho, Malawi, Mozambique, Namibia, Nigeria, Rwanda, South Africa, Swaziland, Uganda, the United Republic of Tanzania, Zambia, Zimbabwe.

Africa – countries with low HIV prevalence: Algeria, Angola, Benin, Burkina Faso, Cape Verde, Chad, the Comoros, Eritrea, the Gambia, Ghana, Guinea, Guinea-Bissau, Liberia, Madagascar, Mali, Mauritania, Mauritius, the Niger, Sao Tome and Principe, Senegal, Seychelles, Sierra Leone, Togo.

Central Europe: Albania, Bosnia and Herzegovina, Montenegro, Poland, Serbia, the former Yugoslav Republic of Macedonia, Turkey.

Eastern Europe: Armenia, Azerbaijan, Belarus, Bulgaria, Georgia, Kazakhstan, Kyrgyzstan, Latvia, Lithuania,

the Republic of Moldova, Romania, the Russian Federation, Tajikistan, Turkmenistan, Ukraine, Uzbekistan.

High-income countries: Andorra, Australia, Austria, the Bahamas, Bahrain, Barbados, Belgium, Bermuda, Brunei Darussalam, Canada, the Cayman Islands, Croatia, Cyprus, the Czech Republic, Denmark, Equatorial Guinea, Estonia, Finland, France, French Polynesia, Germany, Greece, Guam, Hungary, Iceland, Ireland, Israel, Italy, Japan, Kuwait, Luxembourg, Malta, Monaco, the Netherlands, New Caledonia, New Zealand, Northern Mariana Islands, Norway, Oman, Portugal, Puerto Rico, Qatar, the Republic of Korea, San Marino, Saudi Arabia, Singapore, Slovakia, Slovenia, Spain, Sweden, Switzerland, Trinidad and Tobago, the Turks and Caicos Islands, the United Arab Emirates, the United Kingdom, the United States, US Virgin Islands.

Latin America: Anguilla, Antigua and Barbuda, Argentina, Aruba, Belize, Bolivia (Plurinational State of), Brazil, British Virgin Islands, Chile, Colombia, Costa Rica, Cuba, Curaçao, Dominica, the Dominican Republic, Ecuador, El Salvador, Grenada, Guatemala, Guyana, Haiti, Honduras, Jamaica, Mexico, Montserrat, Nicaragua, Panama, Paraguay, Peru, Saint Kitts and Nevis, Saint Lucia, Saint Vincent and the Grenadines, Sint Maarten (Dutch part), Suriname, Uruguay, Venezuela. (Bolivarian Republic of).