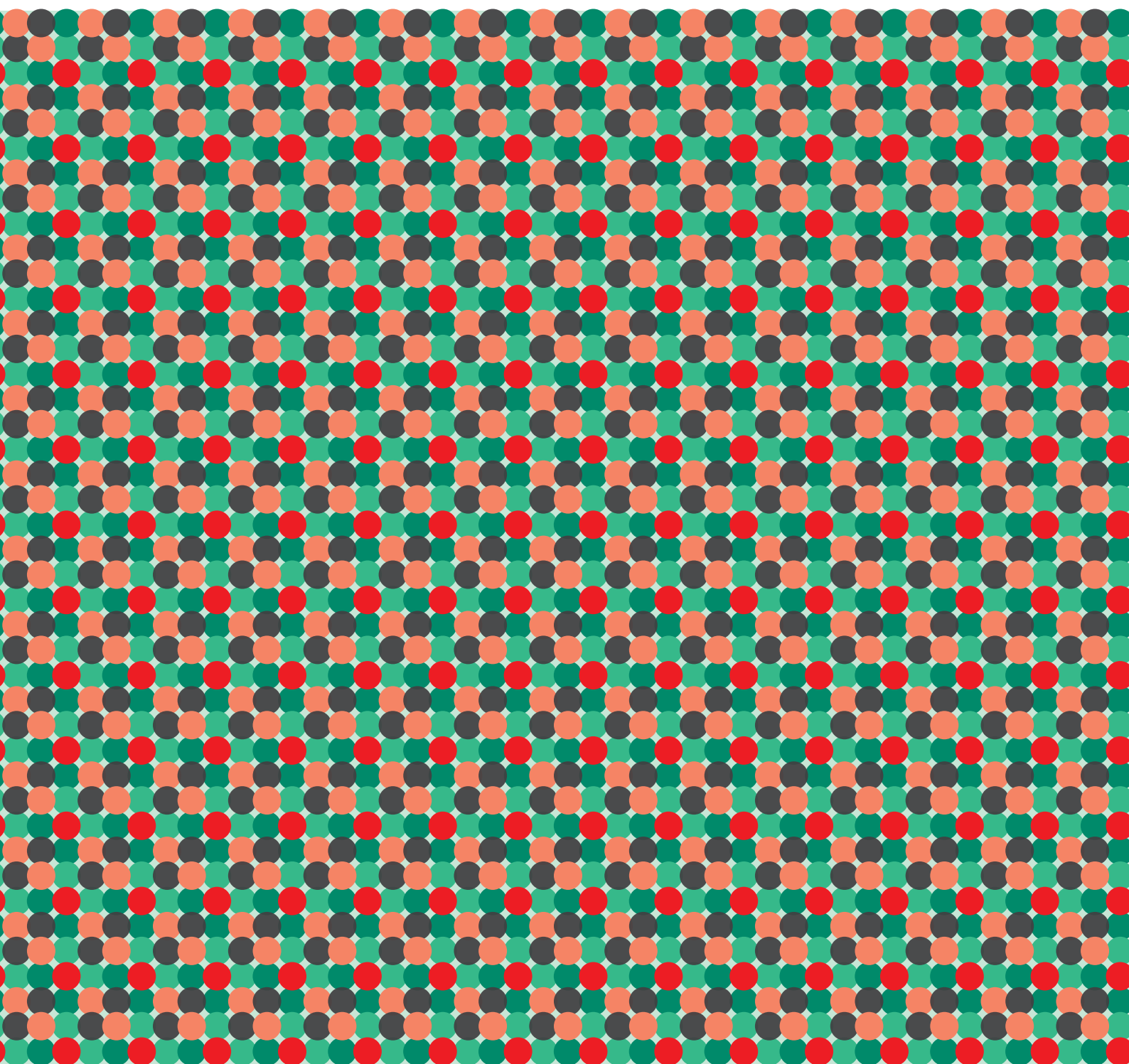


STANDARDS AND BENCHMARKS FOR TUBERCULOSIS
SURVEILLANCE AND VITAL REGISTRATION SYSTEMS



CHECKLIST AND USER GUIDE

The WHO Global Task Force on
TB Impact Measurement



WHO Library Cataloguing-in-Publication Data

Standards and benchmarks for tuberculosis surveillance and vital registration systems:
checklist and user guide.

1.Tuberculosis – epidemiology. 2.Vital statistics. 3.Cause of death. 4.Epidemiologic monitoring. 5.Benchmarking. 6.Checklist. I.World Health Organization.

ISBN 978 92 4 150672 4

(NLM classification: WF 205)

© World Health Organization 2014

All rights reserved. Publications of the World Health Organization are available on the WHO website (www.who.int) or can be purchased from WHO Press, World Health Organization, 20 Avenue Appia, 1211 Geneva 27, Switzerland (tel.: +41 22 791 3264; fax: +41 22 791 4857; e-mail: bookorders@who.int).

Requests for permission to reproduce or translate WHO publications –whether for sale or for non-commercial distribution– should be addressed to WHO Press through the WHO website (www.who.int/about/licensing/copyright_form/en/index.html).

The designations employed and the presentation of the material in this publication do not imply the expression of any opinion whatsoever on the part of the World Health Organization concerning the legal status of any country, territory, city or area or of its authorities, or concerning the delimitation of its frontiers or boundaries. Dotted lines on maps represent approximate border lines for which there may not yet be full agreement.

The mention of specific companies or of certain manufacturers' products does not imply that they are endorsed or recommended by the World Health Organization in preference to others of a similar nature that are not mentioned. Errors and omissions excepted, the names of proprietary products are distinguished by initial capital letters.

All reasonable precautions have been taken by the World Health Organization to verify the information contained in this publication. However, the published material is being distributed without warranty of any kind, either expressed or implied. The responsibility for the interpretation and use of the material lies with the reader. In no event shall the World Health Organization be liable for damages arising from its use.

layout and design: blossoming.it

WHO/HTM/TB/2014.02

Abbreviations

ART	antiretroviral therapy
BMU	basic management unit
CBS	Central Bureau of Statistics
CDC	Centers for Disease Control and Prevention
DST	drug susceptibility test
DRTB	drug-resistant tuberculosis
EPTB	extrapulmonary tuberculosis
EQA	external quality assurance
ETS	Enhanced Tuberculosis Surveillance
HIV	human immunodeficiency virus
HMIS	health management information system
ICD	International Classification of Diseases
ID	identification
LQAS	lot quality assurance sampling
M&E	monitoring and evaluation
MCH	maternal and child health
MDR-TB	multidrug-resistant tuberculosis
NAAT	nucleic acid amplification tests
NTP	national tuberculosis programme
NTSS	National Tuberculosis Surveillance System
ONS	Office of National Statistics
PHE	Public Health England
PMTCT	prevention of mother-to-child transmission
PTB	pulmonary tuberculosis
RIVM	National Institute for Public Health and the Environment
SOP	standard operating procedures
SRL	Supranational TB Reference Laboratory
TB	tuberculosis
TERG	Technical Evaluation Reference Group
UK	United Kingdom
USA	United States of America
USAID	United States Agency for International Development
VR	vital registration
WHO	World Health Organization
ZTLS	zonal TB and leprosy supervisor

Contents

Abbreviations	4
Acknowledgements	6
Introduction	8
Standards and benchmarks for TB surveillance and vital registration systems: Checklist	15
Data, materials and personnel requirements to undertake the checklist	31
Standards and benchmarks for TB surveillance and vital registration systems: User guide	41
Standard B1.1: <i>Case definitions are consistent with WHO guidelines</i>	42
Standard B1.2: <i>TB surveillance system is designed to capture a minimum set of variables for all reported TB cases</i>	45
Standard B1.3: <i>All scheduled periodic data submissions have been received and processed at the national level</i>	48
Standard B1.4: <i>Data in quarterly reports (or equivalent) are accurate, complete and internally consistent (For paper-based systems only)</i>	53
Standard B1.5: <i>Data in the national database are accurate, complete, internally consistent and free of duplicates (For electronic case-based or patient-based systems only)</i>	63
Standard B1.6: <i>TB surveillance data are externally consistent</i>	75
Standard B1.7: <i>TB surveillance data are internally consistent over time</i>	78
Standard B1.8: <i>All diagnosed cases of TB are reported</i>	95
Standard B1.9: <i>Population has good access to health care</i>	99
Standard B1.10: <i>Vital registration system has high national coverage and quality</i>	103
Standard B2.1: <i>Surveillance data provide a direct measure of drug-resistant TB in new cases</i>	106
Standard B2.2: <i>Surveillance data provide a direct measure of the prevalence of HIV infection in TB cases</i>	110
Standard B2.3: <i>Surveillance data for children reported with TB are reliable and accurate, and all diagnosed childhood TB cases are reported</i>	113
Glossary of terms	117
Appendix 1: <i>Data verification tools for standard 1.4</i>	126
Appendix 2: <i>Applying the lot quality-assessment sampling (LQAS) theory</i>	133
Appendix 3: <i>Examples of methods used in The Netherlands, UK and USA to ensure and improve internal consistency of TB surveillance data at the national level</i>	136

Acknowledgements

This document, The standards and benchmarks for tuberculosis surveillance and vital registration systems: Checklist and user guide, was developed as part of the work of the World Health Organization's (WHO) Global Task Force on TB Impact Measurement. One of the Task Force's three major strategic areas of work is strengthening TB surveillance; the development of the Checklist and accompanying user guide were top priorities of the Task Force from 2011 to 2013, and it is hoped that they will provide a foundation for strengthening measurement of the level of and trends in TB disease burden using routine surveillance data from notification and vital registration systems in many countries.

The Checklist and user guide were both developed by a core team of nine people: Laura Anderson (Public Health England, UK), Emily Bloss (Centers for Disease Control and Prevention, USA), Katherine Floyd (WHO headquarters), Philippe Glaziou (WHO headquarters), Irwin Law (WHO headquarters), Charalambos Sismanidis (WHO headquarters), Hazim Timimi (WHO headquarters), Deanna Tollefson (Centers for Disease Control and Prevention, USA) and Matteo Zignol (WHO headquarters). The group was led and coordinated by Emily Bloss and Irwin Law, with broad guidance from Katherine Floyd.

The development of the Checklist and user guide benefited from the inputs of many people, including during a meeting of the Task Force's subgroup on TB surveillance held in September 2011 and a meeting of the full Task Force in March 2012. The core team is particularly grateful to Ibrahim Abubakar (Public Health England and University College London, UK), Lori Armstrong (Centers for Disease Control and Prevention, USA), Amal Bassili (WHO Eastern Mediterranean Regional Office), Ana Bierrenbach (independent consultant, Brazil), Martien Borgdorff (University of Amsterdam, the Netherlands), Udo Buchholz (Koch Institute, Germany), Kevin Cain (Centers for Disease Control and Prevention, USA), Chen-Yuan Chiang (the Union, France), Ted Cohen (Harvard University, USA), Andrei Dadu (WHO Regional Office for Europe), Connie Erkens (KNCV Tuberculosis Foundation, the Netherlands), Dennis Falzon (WHO headquarters), Vahur Hollo (European Centre for Disease Control and Prevention, Sweden), Joseph Imoko (WHO Country Office, Uganda), Nico Kalisvaart (KNCV Tuberculosis Foundation, the Netherlands), Hillary Kipruto (WHO Country Office, Kenya), Eveline

Klinkenberg (KNCV Tuberculosis Foundation, Ethiopia), Eugene McCray (Centers for Disease Control and Prevention, USA), Kathy O'Neill (WHO headquarters), Sai Pothapregada (the Global Fund, Switzerland), Marina Takane (WHO headquarters), Marieke van der Werf (European Centre for Disease Control and Prevention, Sweden), Sara Whitehead (Centers for Disease Control and Prevention, USA) and Norio Yamada (Research Institute for Tuberculosis, Japan) for their contributions to the drafting or careful review of material and to discussions at these meetings. Special thanks are due to Jaap Broekmans for his excellent chairmanship of the two meetings at which the Checklist and user guide were discussed and for his guidance and input throughout the development process.

The authors would also like to thank the national TB programme staff in the 11 countries where the Checklist was piloted: Brazil, China, Egypt, Estonia, Japan, Kenya, the Netherlands, Thailand, Uganda, the UK and the USA.

Funding support was provided from the Bill and Melinda Gates Foundation, the Global Fund and the United States Agency for International Development (USAID).

The document was edited by Tim France (Inis Communication), and its layout and cover were designed by Giacomo Frigerio (Blossom).

The authors declared no conflict of interest, financial or personal, in the preparation of this document.

Introduction

Background

A major goal of tuberculosis (TB) surveillance is to provide an accurate measure of the number of new TB cases and related deaths that occur each year, and to be able to assess these trends over time. In some countries, TB surveillance already meets the standards necessary to do this, but in others, there are important gaps in the TB surveillance system that make this impossible. For example, TB cases that are diagnosed in the private sector go unreported in many settings, and in many countries with a high burden of TB, people with TB may not access health care and therefore not be diagnosed at all. Furthermore, many countries lack vital registration systems with the geographical coverage and quality required to accurately measure deaths caused by TB. Therefore, the *Checklist of standards and benchmarks for TB surveillance and vital registration systems* was developed with the following objectives:

- To assess a national surveillance system's ability to accurately measure TB cases and deaths.
- To identify gaps in national surveillance systems that must be addressed in order to improve TB surveillance.

The results of national assessments using the checklist can be used to identify which countries have surveillance systems that already provide an accurate measure of the number of TB cases and deaths that occur each year, and to define the actions necessary to strengthen surveillance in countries in which gaps are identified. Following the 2012 recommendations of the Global Fund's Technical Evaluation Reference Group (TERG) and a collaborative agreement between the Global Fund and the World Health Organization (WHO), there was a new aim to integrate assessments of TB surveillance using the checklist within Global Fund grant mechanisms. As such, assessments should be timed to coincide with periodic reviews, programme reviews or Global Fund phase II grant renewals, with results used to develop monitoring and evaluation (M&E) investment plans that can be supported through subsequent Global Fund grants. This collaboration has great potential to help strengthen TB surveillance in more than a hundred countries receiving Global Fund grants for TB care and control worldwide.

The checklist was developed by a team of experts in disease surveillance in conjunction with expert advice from meetings organized by WHO in September 2011 and May 2012. The checklist underwent two rounds of

field-testing in eleven countries, including Brazil, China, Egypt, Estonia, Japan, Kenya, the Netherlands, Thailand, Uganda, the United Kingdom and the United States of America, and was revised accordingly.

What does the checklist specifically assess?

The checklist has two parts: Part A provides a general description of the TB surveillance system that is being assessed; Part B (Section 1) is a checklist for TB surveillance and vital registration systems that includes sections covering data quality, system coverage and TB mortality data from vital registration systems. Part B (Section 2) includes the supplementary standards for surveillance of TB/human immunodeficiency virus (HIV) coinfection, drug-resistant TB and childhood TB.

Part A comprises 18 questions that characterize the national TB surveillance system and sets the background for Part B, which consists of 13 standards and their associated benchmarks. The standards are general statements about the criteria for a high-performance TB surveillance system; nine standards are related to the measurement of TB cases and one is related to measurement of TB deaths. There are three supplementary standards that can be used to assess whether a national TB surveillance system can be certified as providing a direct measure of the number of drug-resistant TB cases, HIV-positive TB cases and/or childhood TB cases.

Benchmarks define (in quantitative terms wherever possible) the level of performance considered sufficient to meet respective standards. To ensure that the most complete data are available for review, the assessments are designed to use data for the most recent complete calendar year, unless otherwise stated in the user guide. Depending upon the timeliness of the reporting and finalization of data validation procedures in the system, the lag time may range from no delay to up to one year. In some instances, data from additional years are needed to assess trends over time, or data from only a single quarter are required to reduce the burden of data collection. It is anticipated that an assessment of a TB surveillance system using the checklist would take place every three to five years.

For Parts A and B of the checklist, key actions should be recorded that: 1) address the identified gaps in the surveillance and vital registration systems that prevent them from accurately measuring TB cases and deaths, and 2) help the system improve the quality of TB surveillance based on well-established best practices. An estimated budget to support activities that could bridge these gaps will assist in developing an M&E investment plan.

The data, materials and personnel required to assess each standard and associated benchmark(s) listed below in section 2, followed by the by the user guide in section 3. The user guide was developed to provide instructions to implement the associated checklist of standards and benchmarks in an accurate and standardised way. The rationale for each standard and associated benchmark(s), and the methods that should be used to assess the benchmarks, are explained in the user guide. Specifically, the user guide provides a description of how and what data should be collected. For elements that require reviewing a sample of records, the user guide also explains how the sampling should be conducted. Examples are used to illustrate the methods described in the user guide, as well as recommended corrective actions to take if the benchmarks are not met. The user guide also defines key terms used in the checklist, and further lists the supporting appendices.

It is recognized that the standards and benchmarks related to health system coverage (Standard B1.9) and vital registration (Standard B1.10) are outside the purview of the TB programme. However, to assess the capacity of the surveillance system to accurately estimate TB burden, these two standards and associated benchmarks are deemed necessary.

In a few instances (e.g. Standards B1.4 and B1.8), where compilation of the necessary evidence may be difficult or impossible on a regular basis, it is acceptable to use evidence from the literature, reports of special studies or other related health surveys carried out in recent years to demonstrate that a standard is met, provided results from the assessment of other standards show that data quality within the system has not subsequently declined. This is explained in more detail in the user guide.

This checklist may also be used at the sub-national level, but this is not the primary purpose for which the tool was developed. It should also be noted that the checklist only assesses one part of system capacity and is not intended to assess the system's ability to fulfil other programmatic requirements, e.g. patient care, delivery of laboratory results, or drug stock management. Furthermore, the standards assess the outputs rather than the inputs or processes of the surveillance system, which will vary by country. Using this, along with information collected in the checklist's Table A, countries can identify areas where additional resources can be targeted to effectively strengthen their surveillance systems.

What is a certified TB surveillance system?

For a country's TB surveillance systems to be certified as providing a direct measurement of TB cases and TB deaths, all 10 standards and their associated benchmarks (Part B, Section 1) should be met. The three supplementary standards in Part B (Section 2) can be used to assess whether a TB surveillance system can be certified as providing direct and specific measures of the number of drug-resistant TB cases, HIV-positive cases of TB, and TB in children.

Certification provides an objective situation analysis of the current TB surveillance system. It is intended to provide a baseline and a framework that can be used to support improvements (if required) in the system. Subsequent assessments can be used to determine if targets are met based upon the initial assessments. Certification is based on the review of the system from the assessed time period. External peer review and endorsement of the findings by the WHO Global Task Force on TB Impact Measurement will be necessary for a country's system to be certified.

Who can undertake the checklist?

The checklist can be used by in-country national TB programme staff for self-assessment. All parts of the checklist should be undertaken by someone with an informed and current knowledge of the system that may include all or some of the following people:

- National tuberculosis programme (NTP) manager
- NTP programme officer
- NTP monitoring and evaluation office
- NTP statistician/epidemiologist
- NTP data manager
- WHO TB programme officer

What methods are required and how long does it take to complete the checklist?

The checklist requires an accurate and thorough collection of data from available sources. Therefore, a desktop review of all documents related to the checklist, including existing datasets and electronic surveillance systems, is necessary, and data audits at selected basic management units (BMU)¹ may be required as well. Interviews with the relevant stakeholders and partners may also be necessary to obtain the required information. Depending on how this information is stored, i.e. paper-based or electronic-based, it may take several weeks for the appropriate data to be extracted. Electronic-based data generally require less time to complete the checklist than paper-based systems. Time should also be allocated to summarize the findings of the checklist before dissemination.

1 Definition of a Basic Management Unit: A BMU is defined in terms of management, supervision and monitoring responsibility. A BMU for TB control may have several treatment facilities, one or more laboratories, and one or more hospitals. The defining aspect is the presence of a manager or coordinator who oversees TB control activities for the unit and who maintains a master register of all TB patients being treated, which is used to monitor the programme and report on indicators to higher levels. Typically, the units correspond to a government's second sub-national administrative division, which might be called, for example, a "district" or "county". It is internationally recommended that a BMU cover a population between 50 000 and 150 000 or up to 300 000 for large cities. (Source: Compendium of indicators for monitoring and evaluating national tuberculosis programmes. Geneva, World Health Organization, 2004 (http://www.who.int/tb/publications/tb_compendium_of_indicators/en/, accessed 9 November 2013).



STANDARDS AND BENCHMARKS
FOR TB SURVEILLANCE AND
VITAL REGISTRATION SYSTEMS:
CHECKLIST

PART A: CHARACTERISTICS OF THE TB SURVEILLANCE SYSTEM

Before completing the checklist, it is important to characterize the national TB surveillance system. Please provide answers to the following questions.

COUNTRY NAME: _____ DATE OF ASSESSMENT: _____

QUESTIONS	OUTCOMES (Best practices are in bold)	KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO IMPLEMENT KEY ACTION(S)
A1. How are data recorded for individual TB cases at the service delivery level, e.g. in TB diagnostic units, health centres, clinics? (Tick all that apply)	<input type="checkbox"/> Data are recorded electronically on a national internet-based system <input type="checkbox"/> Data are recorded electronically on a state/provincial/regional internet-based system <input type="checkbox"/> Data are recorded electronically on a local system <input type="checkbox"/> Data are recorded on paper <input type="checkbox"/> Data are not recorded		
A2. Do all service delivery points systematically use standardized TB data collection forms and tools?	<input type="checkbox"/> Yes, completely <input type="checkbox"/> Mostly <input type="checkbox"/> Partially <input type="checkbox"/> No, not at all		
A3. Which TB cases are included in the national TB surveillance data? (Tick all that apply and describe):	<input type="checkbox"/> All TB cases from all parts of the country <input type="checkbox"/> Some TB cases are excluded <input type="checkbox"/> Some part(s) of the country are excluded <input type="checkbox"/> Some case types are excluded <input type="checkbox"/> Some care providers, e.g. non-NTP providers, prisons, private practitioners, are excluded. <input type="checkbox"/> Others: _____ Describe: _____		

QUESTIONS	OUTCOMES (Best practices are in bold)	KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO IMPLEMENT KEY ACTION(S)
A4. What types of TB data are available at the national level? (Tick all that apply)	<input type="checkbox"/> Patient-level data that allow multiple episodes of TB in the same person to be identified are available <input type="checkbox"/> Case-level data are available for all of the country <input type="checkbox"/> Case-level data are available for parts of the country <input type="checkbox"/> Aggregated data are available, i.e. summaries for groups of cases.		
A5. What is the expected frequency of data transmission from the first sub-national administrative level to the national level? (Tick all that apply)	<input type="checkbox"/> Real-time <input type="checkbox"/> More often than monthly <input type="checkbox"/> Monthly <input type="checkbox"/> Quarterly <input type="checkbox"/> Less often than quarterly		
A6. At what levels of the system are TB data systematically verified for accuracy, timeliness and completeness? (Tick all that apply)	<input type="checkbox"/> From the service unit upwards <input type="checkbox"/> From the 1st administrative level upwards <input type="checkbox"/> From the 2nd administrative level upwards <input type="checkbox"/> Only at the national level <input type="checkbox"/> Not at any level		
A7. What types of quality assurance procedures are systematically undertaken for TB data? (Tick all that apply)	<input type="checkbox"/> Quality controls are in place for the electronic surveillance system (automated checks at data entry and batch checking, plus standard operating procedures (SOPs)) <input type="checkbox"/> Data are reviewed during supervisory monitoring visits to service units and sub-national levels (How often? _____) <input type="checkbox"/> Data are reviewed during meetings with TB staff (How often? _____) <input type="checkbox"/> Other (Specify: _____)		

QUESTIONS	OUTCOMES (Best practices are in bold)	KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO IMPLEMENT KEY ACTION(S)
A8. Is feedback on TB data quality systematically provided to all lower reporting levels?	<input type="checkbox"/> Yes, completely <input type="checkbox"/> Mostly <input type="checkbox"/> Partially <input type="checkbox"/> No, not at all		
A9. When are national TB case data for a given calendar year considered ready for national analyses and reporting?	<input type="checkbox"/> Before April the following calendar year <input type="checkbox"/> Before May the following calendar year <input type="checkbox"/> Before June the following calendar year <input type="checkbox"/> On or after beginning of June the following calendar year		
A10. Are there national guidelines for recording and reporting of TB data, e.g. documentation or instructions? (Tick all that apply)	<input type="checkbox"/> Yes. They are posted on the internet <input type="checkbox"/> Yes. They are available in a manual or other reference document, e.g. training materials <input type="checkbox"/> No		
A11. Does the national TB programme have a training plan that includes staff involved in data collection and reporting at all levels of the reporting process?	<input type="checkbox"/> Yes <input type="checkbox"/> No		

QUESTIONS	OUTCOMES (Best practices are in bold)	KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO IMPLEMENT KEY ACTION(S)
A12. How often do TB programme staff receive training specifically on TB surveillance, i.e. recoding and reporting of TB data? (Tick all that apply)	<input type="checkbox"/> Training is routinely received at national and sub-national levels (How often? _____) <input type="checkbox"/> Training is received on an ad hoc basis <input type="checkbox"/> Staff receive training when they are hired <input type="checkbox"/> No routine training is received		
A13. How many staff work on TB surveillance at the national level? (Tick all that apply)	<input type="checkbox"/> Epidemiologist: full-time (# _____) <input type="checkbox"/> Epidemiologist: part-time (# _____) <input type="checkbox"/> Statistician: full-time (# _____) <input type="checkbox"/> Statistician: part-time (# _____) <input type="checkbox"/> Data manager: full-time (# _____) <input type="checkbox"/> Data manager: part-time (# _____) <input type="checkbox"/> Data quality officers: full-time (# _____) <input type="checkbox"/> Data quality officers: part-time (# _____) <input type="checkbox"/> Other (specify: _____)		

QUESTIONS	OUTCOMES (Best practices are in bold)	KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO IMPLEMENT KEY ACTION(S)
A14. Is a national TB surveillance report routinely produced and disseminated on an annual basis?	<input type="checkbox"/> Yes <input type="checkbox"/> No		
A15. Are there written goals of the surveillance system?	<input type="checkbox"/> Yes <input type="checkbox"/> No		
A16. Are policies and procedures are in place to protect the confidentiality of all surveillance data e.g. records, registers?	<input type="checkbox"/> Yes, completely <input type="checkbox"/> Mostly <input type="checkbox"/> Partially <input type="checkbox"/> No, not at all		
A17. Is there a long-term financial plan and budget in place to support TB surveillance activities?	<input type="checkbox"/> Yes <input type="checkbox"/> No		
A18. When was the last time the TB surveillance system was evaluated?	<input type="checkbox"/> Within the past 5 years <input type="checkbox"/> Within the past 5-10 years <input type="checkbox"/> Never		

Additional Notes: _____

PART B (Section 1): CHECKLIST FOR TB SURVEILLANCE AND VITAL REGISTRATION SYSTEMS

For each standard, please assess whether the system is able to satisfy the associated benchmark(s), using the methods recommended in the user guide. Indicate 'Met', 'Partially met', "Not met" or 'Not applicable' in the results column. Describe the key results, any actions recommended to improve the quality of the system and the estimated budget to address these actions in the last two columns.

STANDARD	BENCHMARK(S)	RESULTS (See the User Guide for Interpretation)	RESULTS (DESCRIPTION) INCLUDING KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO ADDRESS KEY ACTION(S)
TB SURVEILLANCE SYSTEM DATA QUALITY				
B1.1. Case definitions are consistent with WHO guidelines	<p>All benchmarks should be satisfied to meet this standard:</p> <ul style="list-style-type: none"> • Laboratory-confirmed cases are distinguished from clinically diagnosed cases² • New cases are distinguished from previously treated cases • Pulmonary cases are distinguished from extrapulmonary cases. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		
B1.2. TB surveillance system is designed to capture a minimum set of variables for all reported TB cases	<p>Data are routinely collected for at least each of the following variables for all TB cases:</p> <ul style="list-style-type: none"> • Age or age group • Sex • Year of registration • Bacteriological results • History of previous treatment • Anatomical site of disease • For case-based systems, a patient identifier 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		

² i.e. by smear, culture or WHO-endorsed molecular test e.g. GeneXpert MTB/RIF

STANDARD	BENCHMARK(S)	RESULTS (See the User Guide for Interpretation)	RESULTS (DESCRIPTION) INCLUDING KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO ADDRESS KEY ACTION(S)
TB SURVEILLANCE SYSTEM DATA QUALITY				
B1.3. All scheduled periodic data submissions have been received and processed at the national level	<p><i>For paper-based systems:</i></p> <ul style="list-style-type: none"> 100% of expected reports from each TB BMU have been received and data aggregated at the national level <p><i>For national patient-based or case-based electronic systems that import data files from sub-national (e.g. provincial or regional) electronic systems:</i></p> <ul style="list-style-type: none"> 100% of expected data files have been imported. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met <input type="checkbox"/> Not Applicable		
B1.4. Data in quarterly reports (or equivalent) are accurate, complete, and internally consistent (<i>For paper-based systems only</i>)	<p>All benchmarks should be satisfied to meet this standard:</p> <ul style="list-style-type: none"> Sub-totals of the number of TB cases by age group, sex and case type matches the total number of reported TB cases in ≥95% of quarterly reports (or equivalent) from BMUs The number of TB cases in ≥95% of quarterly reports (or equivalent) matches the number of cases recorded in BMU TB registers and source documents (patient treatment cards and laboratory register) Data for a minimum set of variables are available for ≥95% of the total number of reported TB cases in BMU TB registers. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met <input type="checkbox"/> Not Applicable		

STANDARD	BENCHMARK(S)	RESULTS (See the User Guide for Interpretation)	RESULTS (DESCRIPTION) INCLUDING KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO ADDRESS KEY ACTION(S)
TB SURVEILLANCE SYSTEM DATA QUALITY				
B1.5. Data in the national database are accurate, complete, internally consistent, and free of duplicates <i>(For electronic case-based or patient-based systems only)</i>	<p>All benchmarks should be met to reach this standard:</p> <ul style="list-style-type: none"> • Data validation checks are in place at the national level to identify and correct invalid, inconsistent and/or missing data in the minimum set (Standard B1.2) • For each variable in the minimum set (Standard B1.2), ≥90% of case records are complete, valid and internally consistent for the year being assessed • <1% of case records in the national dataset for the year being assessed are unresolved potential duplicates. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met <input type="checkbox"/> Not Applicable		
B1.6. TB surveillance data are externally consistent	<ul style="list-style-type: none"> • Among new TB cases, the percentage of children diagnosed with TB is between 5–15% in low- and middle-income, and <10% in high-income countries 	<input type="checkbox"/> Met <input type="checkbox"/> Not met		

STANDARD	BENCHMARK(S)	RESULTS (See the User Guide for Interpretation)	RESULTS (DESCRIPTION) INCLUDING KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO ADDRESS KEY ACTION(S)
TB SURVEILLANCE SYSTEM DATA QUALITY				
B1.7. TB surveillance data are internally consistent over time	<p><i>If vital registration data are available, then the following benchmark should be satisfied for this standard to be met:</i></p> <ol style="list-style-type: none"> 1. Year-to-year change in the national number of reported TB cases is consistent with the year-to-year change in national TB mortality (HIV-negative, from national vital registration) i.e. trajectories with the same direction. <p><i>If vital registration data are not available, then the following benchmarks should be satisfied for this standard to be met. At the national level, evidence of internal consistency over the previous five years for the following benchmarks:</i></p> <ol style="list-style-type: none"> 2. Ratio of notified pulmonary to extrapulmonary TB cases 3. Ratio of male to female TB cases 4. Proportion of childhood TB cases out of all TB cases 5. Year-to-year change in the case notification rate for all forms of TB 6. Year-to-year change in the case notification rate for new smear-positive TB <p>and if data are available,</p> <ol style="list-style-type: none"> 7. Ratio of the number of people with presumptive TB to total notifications of TB cases. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		

STANDARD	BENCHMARK(S)	RESULTS (See the User Guide for Interpretation)	RESULTS (DESCRIPTION) INCLUDING KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO ADDRESS KEY ACTION(S)
TB SURVEILLANCE SYSTEM COVERAGE				
B1.8. All diagnosed cases of TB are reported	Both benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • TB reporting is a legal requirement • ≥90% of TB cases are reported to national health authorities, as determined by a national-level investigation (e.g. inventory study, conducted in past 10 years). 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		
B1.9. Population has good access to health care	Both benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • Under-five mortality rate (probability of dying by age 5 per 1000 live births) is <10 • <25% total health expenditure is out-of-pocket. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		
QUALITY AND COVERAGE OF VITAL REGISTRATION SYSTEM				
B1.10. Vital registration system has high national coverage and quality	Both benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • Cause of death documented in ≥90% of total deaths recorded in: a) national vital registration system or b) sample vital registration system • <10% of deaths have ICD codes for ill-defined causes (defined as ICD-9 780-799 and ICD-10 R00-R99). 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		

PART B (Section 2): SUPPLEMENTARY CHECKLIST FOR TB SURVEILLANCE

For each standard, please assess whether the system is able to satisfy the associated benchmark(s), using the methods recommended in the user guide. Indicate 'Met', 'Partially met', "Not met" or 'Not applicable' in the results column. Describe the key results, any actions recommended to improve the quality of the system and the estimated budget to address these actions in the last two columns.

STANDARD	BENCHMARK(S)	RESULTS	RESULTS (DESCRIPTION) INCLUDING KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO ADDRESS KEY ACTION(S)
SURVEILLANCE OF DRUG RESISTANT TB				
B2.1. Surveillance data provide a direct measure of drug-resistant TB in new cases	One of the two benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • Rifampicin susceptibility status (Positive/Negative) documented for ≥75% of new pulmonary TB cases • Rifampicin susceptibility status (Positive/Negative) documented for a nationally representative drug resistance survey of new pulmonary TB cases. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		
SURVEILLANCE OF TB/HIV				
B2.2. Surveillance data provide a direct measure of the prevalence of HIV infection in TB cases	One of the two benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • HIV status (Positive/Negative) is documented for ≥80% of all notified TB cases • HIV status is available from a representative sample from all TB cases notified in settings with a low-level epidemic state³ or where it is not feasible to implement routine surveillance. 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		

3 Low-level epidemic state: HIV prevalence has not consistently exceeded 5% in any defined sub-population.

STANDARD	BENCHMARK(S)	RESULTS	RESULTS (DESCRIPTION) INCLUDING KEY ACTION(S) REQUIRED TO ADDRESS THE GAPS	ESTIMATED BUDGET REQUIREMENTS TO ADDRESS KEY ACTION(S)
SURVEILLANCE OF CHILDHOOD TB				
B2.3. Surveillance data for children reported with TB are reliable and accurate, and all diagnosed childhood TB cases are reported	Both benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • Ratio of age groups 0–4 to 5–14 years is in the range 1.5–3.0 • ≥90% of childhood TB cases are reported to national health authorities, as determined by a national-level investigation (e.g. inventory study, conducted in the past 10 years) 	<input type="checkbox"/> Met <input type="checkbox"/> Partially met <input type="checkbox"/> Not met		

CHECKLIST SUMMARY (TICK THE BOXES WHERE APPROPRIATE)				
STANDARD	MET	PARTIALLY MET	NOT MET	NOT APPLICABLE
B1.1				
B1.2				
B1.3				
B1.4				
B1.5				
B1.6				
B1.7				
B1.8				
B1.9				
B1.10				
B2.1				
B2.2				
B2.3				



DATA, MATERIALS AND
PERSONNEL REQUIREMENTS TO
UNDERTAKE THE CHECKLIST

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
TB SURVEILLANCE SYSTEM DATA QUALITY			
B1.1. Case definitions are consistent with WHO guidelines	<p>All benchmarks should be satisfied to meet this standard:</p> <ul style="list-style-type: none"> • Laboratory-confirmed cases are distinguished from clinically diagnosed cases⁴ • New cases are distinguished from previously treated cases • Pulmonary cases are distinguished from extrapulmonary cases 	<ol style="list-style-type: none"> 1. NTP manuals and guidelines 2. National TB policy documents 3. National TB reporting forms, registers, treatment cards 4. NTP annual reports 5. Surveillance-related training documents 6. WHO recommended definitions: Definitions and reporting framework for tuberculosis (2013 revision). Geneva, World Health Organization, 2013 (http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf, accessed 9 November 2013). 	<p>NTP manager</p> <p>NTP programme officer</p> <p>NTP monitoring and evaluation officer</p> <p>NTP statistician/epidemiologist</p> <p>NTP data manager</p> <p>WHO TB programme officer</p>
B1.2. TB surveillance system is designed to capture a minimum set of variables for all reported TB cases	<p>Data are routinely collected for at least each of the following variables for all TB cases:</p> <ul style="list-style-type: none"> • Age or age group • Sex • Year of registration • Bacteriological results • History of previous treatment • Anatomical site of disease • For case-based systems, a patient identifier 	<ol style="list-style-type: none"> 1. NTP annual reports 2. Documentation of the surveillance system e.g. SOPs, data dictionary 3. National surveillance database listing the dataset of minimum variables 4. Paper data collection tools e.g. quarterly reports, sub-national reporting forms 5. National laboratory register 6. Reports or publications on data quality or surveillance evaluations. 	As previous

4 i.e. by smear, culture or WHO-endorsed molecular test e.g. GeneXpert MTB/RIF

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
TB SURVEILLANCE SYSTEM DATA QUALITY			
B1.3. All scheduled periodic data submissions have been received and processed at the national level	<p><i>For paper-based systems:</i></p> <ul style="list-style-type: none"> 100% of expected reports from each TB BMU have been received and data aggregated at the national level <p><i>For national patient-based or case-based electronic systems that import data files from sub-national (e.g. provincial or regional) electronic systems:</i></p> <ul style="list-style-type: none"> 100% of expected data files have been imported. 	<p><i>For paper-based systems:</i></p> <ol style="list-style-type: none"> Quarterly reports of TB cases sent to the NTP from BMUs over the period of one year Other reports are received outside the quarterly report system e.g. NGOs, non-NTP providers Possible requirement: if the national level only receives sub-national aggregates then each sub-national entity e.g. BMU, needs to be contacted and reviewed. <p><i>For electronic-based systems:</i></p> <ol style="list-style-type: none"> System logs that show which data files were imported for the reporting year and when they were imported. Possible requirement: If some sub-national data providers also manage their databases by importing data files extracted from databases managed lower down the administrative chain e.g. BMUs, then each of those sub-national data providers need to be contacted and asked to report on the total number of files expected and the total number of data files received and included in the datasets provided to the national level for the year being assessed National TB surveillance database. 	<p>As previous</p> <p>In addition:</p> <p>Provincial TB officers</p> <p>District TB officers</p>

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
TB SURVEILLANCE SYSTEM DATA QUALITY			
B1.4. Data in quarterly reports (or equivalent) are accurate, complete, and internally consistent (<i>For paper-based systems only</i>)	<p>All benchmarks should be satisfied to meet this standard:</p> <ul style="list-style-type: none"> • Sub-totals of the number of TB cases by age group, sex and case type matches the total number of reported TB cases in ≥95% of quarterly reports (or equivalent) from BMUs. • The number of TB cases in ≥95% of quarterly reports (or equivalent) matches the number of cases recorded in BMU TB registers and source documents (e.g. patient treatment cards and laboratory register) • Data for a minimum set of variables are available for ≥95% of the total number of reported TB cases in BMU TB registers 	<p>Method 1:</p> <p><i>Service Availability and Readiness Assessment (SARA):</i></p> <p>This assessment requires health facility visits with data collected based on key informant interviews and observation of key items. SARA can either be carried out as a sample or a census; the choice between these methodologies will depend on a number of elements including the country's resources, the objectives of the survey, and the availability of a complete listing of all health facilities (public and private) in the country (Master Facility List). The data quality module for TB should be part of SARA to fulfil the standard. More details: http://www.who.int/healthinfo/systems/sara_methods/en/index.html</p> <p>Method 2:</p> <p><i>Data quality audit of TB BMU:</i></p> <p>Review data for a specified period of time e.g. one quarter, including quarterly reports sent to the NTP from BMUs, TB registers from health facilities, patient treatment cards, laboratory registers and a list of all TB BMUs in the country.</p>	<p>As previous</p> <p>In addition:</p> <p>Provincial TB officers</p> <p>District TB officers</p> <p>Laboratory managers</p> <p>SARA: Small survey teams to do the census and health facility assessments that should include NTP staff.</p>

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
TB SURVEILLANCE SYSTEM DATA QUALITY			
B1.5. Data in the national database are accurate, complete, internally consistent and free of duplicates <i>(For electronic case-based or patient-based systems only)</i>	<p>All benchmarks should be met to reach this standard:</p> <ul style="list-style-type: none"> • Data validation checks are in place at the national level to identify and correct invalid, inconsistent and/or missing data in the minimum set (Standard B1.2) • For each variable in the minimum set (Standard B1.2), ≥90% of case records are complete, valid and internally consistent for the year being assessed • <1% of case records in the national dataset for the year being assessed are unresolved potential duplicates. 	<ol style="list-style-type: none"> 1. Records of notified TB cases in the national patient- or case-based database for the year of assessment. 2. Documentation and/ or SOPs for electronic surveillance systems <ul style="list-style-type: none"> • System logs that show which data files were imported for the reporting year and when they were imported • List of automated checks run at the time of data entry • List of data queries used to check data quality at the national level • SOPs for detection and removal of duplicate TB cases at national level. 	As previous
B1.6. TB surveillance data are externally consistent	<ul style="list-style-type: none"> • Among new TB cases, the percentage of children diagnosed with TB is between 5-15% in low- and middle-income, and <10% in high-income countries. 	<ol style="list-style-type: none"> 1. Reported TB case data from the national routine TB surveillance system disaggregated by age from the last year for which complete data are available 2. Country income grouping from the World Bank (http://data.worldbank.org/about/country-classifications/country-and-lending-groups, accessed 9 November 2013). 	As previous

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
TB SURVEILLANCE SYSTEM DATA QUALITY			
B1.7. TB surveillance data are internally consistent over time	<p>If vital registration data are available, then the following benchmark should be satisfied for this standard to be met:</p> <ol style="list-style-type: none"> 1. Year-to-year change in the national number of reported TB cases is consistent with the year-to-year change in national TB mortality (HIV-negative, from national vital registration) i.e. trajectories with the same direction. <p>If vital registration data are not available, then the following benchmarks should be satisfied for this standard to be met. At the national level, evidence of internal consistency over the previous five years for the following benchmarks:</p> <ol style="list-style-type: none"> 2. Ratio of notified pulmonary to extrapulmonary TB cases 3. Ratio of male to female TB cases 4. Proportion of childhood TB cases out of all TB cases 5. Year-to-year change in the case notification rate for all forms of TB 6. Year-to-year change in the case notification rate for new smear-positive TB <p>and if data are available,</p> <ol style="list-style-type: none"> 7. Ratio of the number of people with presumptive TB to total notifications of TB cases. 	<p>If VR data are available:</p> <ol style="list-style-type: none"> 1. TB mortality rates (HIV-negative TB) at the national level are obtained from vital registration (VR) systems <p>If VR data are not available:</p> <ol style="list-style-type: none"> 2. National level TB case data disaggregated by age (or age group), sex, type of disease, along with case notification rates (all forms and smear-positive TB cases) are obtained from routine TB surveillance for the past 5 years. Although not part of the benchmark assessment itself, similar subnational data should also be collected and examined 3. If available, TB suspect registry data at the national level. 	As previous

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
TB SURVEILLANCE SYSTEM COVERAGE			
B1.8. All diagnosed cases of TB are reported	Both benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • TB reporting is a legal requirement • ≥90% of TB cases are reported to national health authorities, as determined by a national-level investigation (e.g. inventory study, conducted in past 10 years). 	<ol style="list-style-type: none"> 1. Legal and regulatory frameworks, national TB health reports and policy documents 2. Inventory study reports 3. Reports or publications on data quality or surveillance evaluations. 	As previous
B1.9. Population has good access to health care	Both benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • Under-five mortality rate (probability of dying by age 5 per 1000 live births) is <10 • <25% total health expenditure is out-of-pocket. 	<ol style="list-style-type: none"> 1. Under-five mortality rate: WHO Global Health Observatory (www.who.int/gho/child_health/mortality/mortality_under_five/en, accessed 9 November 2013). 2. Proportion of national health expenditures that are out-of-pocket: WHO national health accounts database (www.who.int/nha/database, accessed 9 November 2013). WHO Global Health Expenditure Atlas (http://apps.who.int/nha/atlasfinal.pdf, accessed 9 November 2013). 	As previous

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
QUALITY AND COVERAGE OF VITAL REGISTRATION SYSTEM			
B1.10. Vital registration system has high national coverage and quality	Both benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • Cause of death documented in $\geq 90\%$ of total deaths recorded in: a) national vital registration system or b) sample vital registration system • $< 10\%$ of deaths have ICD* codes for ill-defined causes (defined as ICD-9 780-799 and ICD-10 R00-R99). 	1. Routine annual reports or periodic surveys about vital statistics 2. WHO Mortality Database (http://apps.who.int/gho/data/node.main.686?lang=en , accessed 9 November 2013).	As previous

(*ICD: International Classification of Diseases)

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
SURVEILLANCE OF DRUG-RESISTANT TB			
B2.1. Surveillance data provide a direct measure of drug-resistant TB among new cases	One of the two benchmarks should be satisfied to meet this standard: <ul style="list-style-type: none"> • Rifampicin susceptibility status (Positive/Negative) documented for $\geq 75\%$ of new pulmonary TB cases • Rifampicin susceptibility status (Positive/Negative) documented for a nationally representative drug resistance survey of new pulmonary TB cases. 	1. NTP report covering surveillance systems based on routine diagnostic testing of previously untreated TB cases 2. Report covering the results from special surveys of a representative sample of previously untreated TB cases e.g. results from a drug resistance survey conducted in the past five years including documentation of results of proficiency testing conducted at the supranational TB reference laboratory 3. National laboratory register.	As previous In addition: Laboratory manager

STANDARD	BENCHMARK(S)	DATA REQUIREMENTS (If Available)	PERSONNEL
SURVEILLANCE OF TB/HIV			
B2.2. Surveillance data provide a direct measure of the prevalence of HIV infection in TB cases	<p>One of the two benchmarks should be satisfied to meet this standard:</p> <ul style="list-style-type: none"> • HIV status (Positive/Negative) documented for $\geq 80\%$ of all notified TB cases • HIV status is available from a representative sample from all TB cases notified in settings with a low-level epidemic state⁵ or where it is not feasible to implement routine surveillance. 	<ol style="list-style-type: none"> 1. NTP report covering surveillance systems based on routine HIV testing of TB cases 2. Dataset from a standardized electronic system that may include an assessment of record linkages between TB and HIV surveillance systems 3. Report covering the results from a periodic survey of HIV infection among a sample of TB cases. 	<p>As previous</p> <p>In addition: Laboratory manager</p> <p>HIV surveillance officer</p>
SURVEILLANCE OF PAEDIATRIC TB			
B2.3. Surveillance data for children reported with TB are reliable and accurate, and all diagnosed childhood TB cases are reported	<p>Both benchmarks should be satisfied to meet this standard:</p> <ul style="list-style-type: none"> • Ratio of age groups 0–4 to 5–14 years is in the range 1.5–3.0 • $\geq 90\%$ of childhood TB cases are reported to national health authorities, as determined by a national-level investigation (e.g. inventory study, conducted in past 10 years). 	<ol style="list-style-type: none"> 1. NTP annual reports, manuals and guidelines 2. National and subnational datasets that are disaggregated by case type, age (including children <15 years) and sex 3. National TB reporting forms, registers, treatment cards specifically with reference to children <15 years of age. 	

5 Low-level epidemic state: HIV prevalence has not consistently exceeded 5% in any defined subpopulation.



STANDARDS AND BENCHMARKS
FOR TB SURVEILLANCE AND
VITAL REGISTRATION SYSTEMS:
USER GUIDE

Standard B1.1: Case definitions are consistent with WHO guidelines

Benchmarks:

All benchmarks should be satisfied to meet this standard:

- Laboratory-confirmed cases are distinguished from clinically diagnosed cases
- New cases are distinguished from previously treated cases
- Pulmonary cases are distinguished from extrapulmonary cases

Rationale for standard and benchmarks

TB case definitions are essential for effective TB surveillance. They are necessary to provide consistent information on epidemiological trends and control programme performance, and to make accurate local, regional and global comparisons. They are also used to guide treatment selection. Reported TB cases should be standardized to allow for meaningful monitoring of differences in rates between geographical areas and monitoring of trends in rates in reported TB cases over time. Case definitions that are unclear and not standardized, and changes or inconsistencies in case definitions, are detrimental to such monitoring. TB case definitions that are not comparable with universally recommended criteria for categorizing cases, e.g. laboratory⁶ versus clinically confirmed, new versus previously treated, pulmonary versus extrapulmonary, or are inconsistent over time make the analysis of surveillance data difficult, if not impossible.

Method to assess benchmarks

Data sources and data collection methods: Case definitions should be documented in NTP manuals and policies. A review of policy documents, programme manuals, annual reports and/or reporting forms are needed to assess this standard. National case definitions should be compared with WHO recommended definitions for logical consistency (rather than exact wording).⁷ If TB case counts based on national definitions equal TB case counts that would be obtained using WHO definitions, national definitions should be classified as being consistent with WHO guidelines.

⁶ By smear, culture or WHO-endorsed molecular test, e.g. GeneXpert MTB/RIF.

⁷ Definitions and reporting framework for tuberculosis (2013 revision). Geneva, World Health Organization, 2013 (http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf, accessed 9 November 2013).

Main limitations: The existence of standard case definitions in national policy documents does not necessarily translate into their adoption at all levels.

Interpretation of results: All three benchmarks should be satisfied to meet this standard. The standard is only partially met if one or two benchmarks are not satisfied. If none of the benchmarks are satisfied, the standard is not met.

Recommended actions: If this standard is not met, convene a national advisory group to review current case definitions, and then edit the national guidelines, adapt recording and reporting systems (paper forms and registers and/or electronic databases) and update curriculum and training materials accordingly.⁸ If required, key actions that address the gaps that currently prevent the standard from being achieved should be described. An estimated budget to support activities that could bridge these gaps would assist in developing an investment plan.

Examples

Kenya

In Kenya, the Division of Leprosy, Tuberculosis and Lung Disease has developed and distributed *Guidelines on the Management of Leprosy and Tuberculosis*. The guidelines were reviewed to assess whether the definitions are consistent with those recommended by WHO. Case definitions were clearly defined in the guidelines for TB management along with recording and reporting of TB case data, differentiating between laboratory-confirmed and clinical cases, new and previously treated cases (including relapses, failures and returnees after default), and pulmonary and extrapulmonary cases. Furthermore, a review of the curriculum and TB training manual for health workers demonstrated that these case definitions were part of the routine training for TB management. The definitions in Kenya are consistent with WHO guidelines, and the surveillance system in Kenya meets this standard.

United States of America

The electronic case-based TB surveillance system in the United States spans all 50 states, the District of Columbia (DC) and eight USA-affiliated island nations. Reporting occurs at the city, county and state level. Physicians and laboratories are responsible for reporting TB to the state or local health department, and likewise, the county and city areas report to their respective state TB programmes. State TB programmes,

8 See Parts A1, A3, A4, A10, A11, A15.

Washington DC, and New York City then directly report TB cases to the Centers for Disease Control and Prevention (CDC). Aggregate reporting was used from 1953 to 1984, but in 1985 a standardized report form, the Report of Verified Case of Tuberculosis (RVCT), was introduced and its mandated use allowed for case-based reporting. The RVCT has been revised several times since 1985, most recently in 2009. In addition, pulmonary TB cases are distinguished from extrapulmonary cases, and case classification or verification criteria are hierarchical, including: culture-confirmed, nucleic acid amplification test (NAAT)-confirmed, smear positive, clinical case and provider diagnosis. The USA TB surveillance system does not identify 'retreatment' cases, but does distinguish cases with a history of previous TB. Starting in 2009 and 2010, state TB programmes began reporting a previous episode of TB using a previously-reported patient identifier. Case definitions in the USA are slightly different from, but consistent with WHO guidelines. Therefore, the system in the USA satisfies all three benchmarks and meets this standard.

Standard B1.2: TB surveillance system is designed to capture a minimum set of variables for all reported TB cases

Benchmarks:

Data are routinely collected for at least each of the following variables for all TB cases:

- Age or age group
- Sex
- Year of registration
- Bacteriological results
- History of previous treatment
- Anatomical site of disease
- For case-based systems, a patient identifier

Rationale for standard and benchmarks

Surveillance systems need to collect data for a set of well-defined variables. It is important to collect uniform data to improve the comparability, consistency and relevance of surveillance information. The minimum set of variables selected represents the fundamental attributes necessary to assess data quality such that TB data can measure TB incidence and trends. For example, data on all cases disaggregated by age, sex, year, bacteriological results (i.e. laboratory⁹ versus clinically confirmed), history of previous treatment (i.e. new versus previously treated), and anatomical site of disease (e.g. pulmonary versus extrapulmonary), are needed to assess whether standards B1.5, B1.6 and B1.7 are met. In case-based systems, a unique patient identifier such as a TB registration number is needed to match and remove duplicate cases (see Section B1.5). Although this benchmark lists variables for which information must be routinely collected in order to meet the standard, it does not preclude NTPs from collecting additional data to meet their own specific needs.

9 i.e. by smear, culture or WHO-endorsed molecular test e.g. GeneXpert MTB/RIF.

Method to assess benchmarks

Data sources and data collection methods: This standard specifically targets data collected by providers within the national tuberculosis control network, including providers under the NTP and other public and private providers engaged with the programme. If an official national annual TB report from the previous year has complete data, is available for review and demonstrates that a minimum set of variables are collected for all TB cases, the report itself is sufficient to show the benchmarks have been met. Alternatively, if either a standardized electronic system or a set of paper data collection tools is used to capture each of the minimum sets of variables, then the benchmarks can be considered to have been met. If this is not the case, a review of the national surveillance database or, for paper-based systems, the TB quarterly reports can be undertaken.

Main limitations: This standard will miss providers that do not report to the NTP and/or are outside the reporting network. Most countries collect age- and sex-disaggregated data and these data are needed for all TB cases. However, paper-based systems following the 2006 WHO recording and reporting guidelines¹⁰ will fail both age and sex benchmarks, which are intended for smear-positive cases only.

Interpretation of results: For all benchmarks to be satisfied and to meet the criteria to satisfy the standard, data should be routinely captured by programme recording and reporting for each of the following variables for all cases: age or age group, sex, year of registration, and all three case types (new, pulmonary and laboratory-confirmed). For case-based systems, a patient identifier is also required. The standard is only partially met if at least one but not all criteria are satisfied. If none of the criteria are satisfied then the standard is not met.

Recommended actions: New programme recording and reporting forms and systems can be developed, and related trainings conducted to ensure the collection of the minimum set of data required to assess trends as well as internal and external consistency.¹¹ If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

¹⁰ Revised TB recording and reporting forms and registers. Geneva, World Health Organization, 2006 (http://whqlibdoc.who.int/hq/2006/WHO_HTM_TB_2006.373_eng.pdf, accessed 9 November 2013).

¹¹ See Parts A3, A4, A10, A11.

Example

Uganda

In Uganda, the surveillance system is paper-based. Patient data are collected at the facility level and then recorded in facility and district TB registers. Aggregated reports are compiled at the district level and sent via the province to the NTP office. A national annual TB report is routinely disseminated and standardized data collection forms are distributed and used in reporting units nationally. The annual and quarterly reports are reviewed in order to assess whether the minimal set of variables are being utilized to capture data for all TB cases in Uganda. Data are collected on sex, year of registration and disease site (pulmonary versus extrapulmonary TB), along with other variables prioritised by the National TB and Leprosy Programme. The age of each patient is collected in unit and district TB registers, but it is only routinely reported for new smear-positive cases and not for all cases. Laboratory confirmation is only collected for pulmonary cases. Based on the assessment, only some of the variables are captured for all cases, so this standard is *partially met*.

Standard B1.3:

All scheduled periodic data submissions have been received and processed at the national level

Benchmarks:

- For paper-based systems: 100% of expected reports from each TB BMU have been received and data aggregated at the national level
- For national patient-based or case-based electronic systems that import data files from sub-national (e.g. provincial or regional) electronic systems: 100% of expected data files have been imported

Rationale for standard and benchmarks

This standard focuses on the extent to which sub-national units are fulfilling their obligations to report their data to the national level. If not all reporting units at the sub-national level make their data available to the national level then nationally-available data are incomplete and, as a result, the TB surveillance system is not able to provide an accurate measure of the number of notified TB cases each year.

Paper-based systems, that follow WHO guidelines, rely on quarterly reports compiled at TB BMUs, where TB registers for a set of health facilities are aggregated. These quarterly reports are sent up an administrative chain, with aggregation sometimes occurring at each stage in the chain. The national aggregates should be based on reports from 100% of BMUs (that is, the lowest level at which aggregation occurs) in the country for the year being evaluated.

Some national patient-based or case-based electronic surveillance systems are not integrated real-time systems, i.e. BMUs or health facilities do not enter their individual TB case records directly into the national system. Instead, sub-national reporting units (for example, provinces) manage their own case-based electronic systems. The sub-national units extract TB case records from their databases into a file using an agreed format,¹² and send these data files according to an agreed timetable to the national level, where the files are imported into the national patient- or case-based database. In such situations, the

¹² For example an XML-based format such as HL7, or CSV files or native database files.

national database should include 100% of the expected data files from sub-national databases in the country for the year being evaluated. Using less than 100% of expected quarterly reports or data files can result in many cases not being included in the national statistics.

Method to assess benchmarks

Data sources and data collection methods:

- A. *Paper-based systems:* The main data sources are the quarterly reports of TB cases sent to the NTP from BMUs over the period of one year. If quarterly reports are received at the national level then the total number of reports received for the year being assessed can be counted directly. If other reports are received from outside the quarterly report system such as reports from non-NTP providers, these should also be included in the calculations. If the national level only receives sub-national aggregates, for example regional aggregates, then each sub-national entity must be contacted and asked to report on the total number of quarterly reports expected and the total number of reports received for each quarter from BMUs that have been included in the sub-national aggregates for the year being assessed.
- B. *Electronic systems that import data files extracted from sub-national databases, i.e. BMUs or health facilities that do not enter their individual TB case records directly into a national system:* The main data sources are the system logs that show which data files were imported for the reporting year and when they were imported. If some sub-national data providers also manage their databases by importing data files extracted from databases managed lower down the administrative chain, for example from BMUs, then each of those sub-national data providers need to be contacted and asked to report on the total number of files expected and the total number of data files received and included in the datasets provided to the national level for the year being assessed.

Alternatively, if no logs are kept, the national database can be queried to produce a table showing case notifications for each period covered by expected data files in the year being assessed, disaggregated by each of the sub-national data providers. While not perfect, this can show if any data files had not yet been imported from any of the data providers.

Main limitations:

1. Reporting units often do not submit reports of zero cases.
2. This standard does not quantify the extent of underreporting of cases within the submitted or non-submitted reports. This is addressed in standard B1.8 (*All diagnosed cases of TB are reported*).
3. This standard does not consider delays in accumulation of data at the national level.
4. This standard is limited to the units reporting to the NTP and therefore does not account for the units that should be submitting reports to the NTP but do not, such as private or public providers that treat TB cases.
5. The second benchmark applies to electronic systems that import data from sub-national feeder systems and therefore does not apply to integrated real-time systems, i.e. ones where BMUs, laboratories, health facilities and other providers inside and outside the NTP enter their individual TB case records directly into a national system. Assessing reporting completeness of these integrated systems is best done through inventory studies (see standard B1.8: *All diagnosed cases of TB are reported*). However, tabulation of TB cases reported by all BMUs and other registered providers in the year being assessed could provide an indication of missing cases, especially if expected case numbers per provider are not very low or if reporting of zero cases by providers is enforced.

Interpretation of results: Compare the number of routine periodic reports on TB cases included at the national level for the evaluation year to the number of reports expected to be included for the same period. If they are equal then the standard is *met*. If at least 50% but less than 100% of expected reports were received or imported at the national level, then the standard is *partially met*. If this is less than 50%, then the standard is *not met*.

Recommended actions: If some expected quarterly reports or datasets are missing then investigate the sources and causes of underreporting and take correctives measures. Routine cross-checking of TB registers at BMUs against reports during supervisory visits and meetings are crucial to minimize this source of data incompleteness. The importance of zero case reporting is recommended. SOPs should be in place to ensure sub-national entities have collected data from all of their constituencies before submitting their data to the national level. Completeness of reporting should be emphasized in the national M&E framework and

plan. Implementation of a national patient- or case-based electronic recording and reporting system should be considered if the current system is paper-based.¹³ If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Examples

Paper-based system

In Uganda, it is possible to know the number of districts submitting reports to the national level each year and compare this to the expected total. In 2011, 468 out of 468 expected reports were received at the national level. Therefore, Uganda meets the standard for 2011.

Standard practice in Uganda's paper-based system emphasizes the importance of monitoring reports. The BMU is the district. Facilities within the NTP as well as a number of private clinics and hospitals report within the TB surveillance network. Quarterly aggregate reports from districts are sent to the national level through the zonal TB and leprosy supervisors (ZTLSs), who validate the data before forwarding the reports to the national level. Also, the number of reports is routinely reviewed during quarterly meetings at the zonal level, where a ZTLS meets all district TB and leprosy supervisors and, at times, other key players such as laboratory focal people, chief administrative officers, heads of civil servants within the districts and secretaries of health. At the national level, the programme manager of the National TB and Leprosy Programme meets all ZTLSs and key partners. A few innovative districts also hold district quarterly meetings where the district health team meets all the sub-county health workers and nurses in charge of diagnostic treatment units.

Electronic system that imports data files from sub-national electronic systems

In a country with 10 provinces, each province sends a data file of TB cases four times a year to the national authorities (i.e. the NTP) who import the data into the national database. SOPs for submitting to the national database state that provinces should not send data files until they have included data from all sub-provincial data providers.

13 See Part A5, A9, A10.

The NTP also receives data files four times a year from the prison service and four times a year from a nongovernmental organization that provides TB diagnostic and treatment services outside the NTP.

Logs at the national level for 2011 show that all quarterly data files were received and imported from nine provinces, the prison service and from the nongovernmental organization, but data from only three quarters had been received from one province. Therefore, 47 out of 48 (98%) expected data files were received, this standard was *partially met* for 2011.

An alternative check: The national database was queried for the number of notified cases by source for each quarter, giving the following result:

Source	Year	Quarter	Cases notified
Province 1	2011	1	360
Province 1	2011	2	410
Province 1	2011	3	370
Province 1	2011	4	0
Province 2	2011	1	899
Province 2	2011	2	885
Province 2	2011	3	867
Province 2	2011	4	874
...*	0
Province 10	2011	3	703
Province 10	2011	4	638
Prison service	2011	1	27
Prison service	2011	2	33
Prison service	2011	3	26
Prison service	2011	4	29
NGO	2011	1	96
NGO	2011	2	70
NGO	2011	3	102
NGO	2011	4	91

*All provinces and quarters not shown in the table had zero cases notified.

The query results showed all provinces, the prison service and the nongovernmental organization had reported more than one case in each quarter, apart from province 1 that did not report any cases in quarter 4. This implies that one data file is missing from the national database. Therefore, confirming that this standard was partially met for 2011.

Standard B1.4:

Data in quarterly reports (or equivalent) are accurate, complete and internally consistent (For paper-based systems only)

Benchmarks:

All benchmarks should be satisfied to meet this standard:

- Sub-totals of the number of TB cases by age group, sex and case type equals the total number of reported TB cases in $\geq 95\%$ of quarterly reports (or equivalent) from BMUs
- The number of TB cases in $\geq 95\%$ of quarterly reports (or equivalent) matches the number of cases recorded in BMU TB registers and source documents (e.g. patient treatment cards and laboratory register)
- Data for a minimum set of variables are available for $\geq 95\%$ of the total number of reported TB cases in BMU TB registers

Rationale for standard and benchmarks

Based on WHO recording and reporting guidelines, registration of TB cases is completed on standardized recording and reporting forms, involving the collection of patient demographic and clinical data, as well as data on sputum examination results (TB and laboratory registers). Furthermore, data are collected to monitor patient progress towards the completion of treatment (treatment patient cards). Data from these standardized forms from the TB treatment facilities and laboratories are collated and consolidated at the level of the TB BMU to produce quarterly reports, which in turn feed into the production of nationwide quarterly reports. These reports aggregate and present data on case-finding and sputum examination results allowing for an assessment of: 1) the burden of TB disease in the country; and, 2) the performance and effectiveness of the surveillance system of the NTP. Erroneous conclusions could be drawn if the compilation of BMU data is inaccurate or incomplete.

Method to assess benchmarks

These benchmarks can be assessed through one of two means: 1) the Service Availability and Readiness Assessment (SARA); or, 2) an independent

national data quality audit mission. The preference is to use the results from SARA, however if this is not available then the latter method should be undertaken. The assessment of this standard places high demands on human resources and time. Undertaking either of these methods should be established by the NTP with appropriate technical guidance.

Service availability and readiness assessment (SARA)

SARA is a cross-cutting assessment tool, developed by the WHO Health Statistics and Health Information Systems department, that is designed to assess and monitor the service availability and readiness of the health sector to deliver key services, and to generate evidence to support the planning and management of a health system. The objective of the tool is to generate reliable and regular information on service delivery (such as the availability of key human and infrastructure resources), the availability of basic equipment, basic amenities, essential medicines, and diagnostic capacities, and the readiness of health facilities to provide basic health care interventions relating to family planning, child health services, basic and comprehensive emergency obstetric care, HIV, TB, malaria, and noncommunicable diseases. The SARA methodology was developed through collaboration between WHO and the United States Agency for International Development (USAID) to fill critical gaps in measuring and tracking progress in health systems strengthening. Repeat assessments with the tool can measure progress in health system strengthening over time.

One optional component of SARA is the TB data verification record review module (see examples in Appendix 1) that can be used to specifically assess standard B1.4 and its related benchmarks. This module should be specifically requested by the NTP and included in the overall SARA assessment if it is used as part of the checklist assessment. Completion of the SARA tool requires visits to BMUs and their respective health facilities¹⁴ by small teams with data collected based on key informant interviews and review of key documents for a certain period of time either through a nationally representative sample of public and private health facilities or a complete census. The sampling methodology of the BMUs and health facilities can be found online.¹⁵ Adaptation of the sampling methodology (see below about lot quality assurance sampling) – and logistical management that is required to undertake the mission – requires a knowledge of the TB surveillance system structure as the types of BMUs will vary between countries. Given that the entire SARA assessment encompasses more than solely TB, the complexity of data collection

14 Definition of a health facility: Any health institution with health care providers formally engaged in any of the following TB programme functions (DOTS): referring patients with presumptive TB or confirmed TB cases, laboratory diagnosis, TB treatment and patient support during treatment. (Source: Compendium of indicators for monitoring and evaluating national tuberculosis programmes. Geneva, World Health Organization, 2004 (http://www.who.int/tb/publications/tb_compendium_of_indicators/en/, accessed 9 November 2013).

15 More information about SARA sampling methodology: http://www.who.int/healthinfo/systems/sara_introduction/en/index.html.

is dependent on available human resources, therefore any undertaking should be planned in consultation with the SARA team at WHO.¹⁶

Independent national data quality audit mission

The same data verification record review data collection tool used in SARA can be used in the independent national data quality audit mission. Similar to SARA, the assessment is done at sampled BMUs and their respective health facilities. This should be done in a representative manner from an exhaustive list of all BMUs in the country, with a probability of selection for each BMU proportional to the number of TB cases they have reported for a certain period of time e.g. over the last calendar year. Sampling should be adapted to each country context, ensuring sampled units are drawn from all relevant strata as defined by variables associated with estimated TB burden in the country, e.g. urban versus rural, or low versus medium versus high TB reporting areas. Lot quality assurance sampling (LQAS)^{17,18} is recommended for the sample size determination of the number (n) of BMUs required to be sampled out of the total N in the country (see Appendix 2 for further guidance on sample size calculations. Ideally, LQAS methodology should also be used to estimate the sample size if SARA is also being implemented, therefore consultation with the SARA team and/or a statistician is strongly recommended). For small countries it might be required to include all BMUs in the assessment. Therefore, adaptation of the sampling methodology will require a good knowledge of the TB surveillance system structure as the types of BMUs will vary between countries.

Data sources and data collection methods: To assess the benchmarks, as part of SARA or as a separate data quality audit mission, data audits within each BMU should be done for a pre-specified period of time (verification period) e.g. one quarter from the past calendar year, as determined by the assessment team in consultation with the NTP.

As the transfer of TB surveillance data between different administrative levels, and the use of different source documents, may vary between countries, the logistics and feasibility of undertaking the assessment will need to be adequately planned. Staff from the NTP should be part of the team to undertake these assessments as they will have intimate knowledge of the documentation, and assist the logistics.

For the purposes of the first benchmark, for each selected BMU and its related health facilities, the BMU quarterly report from the verification

¹⁶ Please email Kathy O'Neill: oneillk@who.int

¹⁷ Sandiford P (1993). Lot quality assurance sampling for monitoring immunization programmes: cost-efficient or quick and dirty? *Health Policy and Planning*, 8:217–223.

¹⁸ Robertson SE & Valadez JJ (2006). Global review of health care surveys using lot quality assurance sampling (LQAS), 1984–2004. *Social Science and Medicine*, 63:1648–1660.

period should be checked for internal consistency. More specifically, this involves comparing disaggregated sub-totals in the quarterly report by: 1) age by sex; and, 2) case type against the reported totals (see Appendix 1.1 and 1.2). If the list of BMUs that will be assessed is already known, all the BMU quarterly reports could be assessed at the national level where they are all sent and collated.

For the purposes of the second benchmark, for each selected BMU and its related health facilities, the first data check consists of the BMU quarterly report from the verification period to be checked against the source documents – most commonly, BMU and health facility TB registers¹⁹, patient treatment cards and laboratory registers.²⁰ A second data check requires these source documents to be checked against each other. More specifically, the data checks involve the following (see Figure 1a, Appendix 1.1 and 1.3 for an example of the register verification tool):

- (i) Using the BMU TB register, count the number of TB cases for the verification period, and compare this against the number of cases reported in the BMU quarterly report,²¹ for the same verification period (Figure 1a).
- (ii) In the laboratory TB register(s) linked with the BMU, count the number of bacteriologically-confirmed cases²² listed for the verification period, and compare this against the number of cases reported in the BMU quarterly report for the same verification period (Ideally, laboratory-confirmed cases in the laboratory register should be individually matched with the referring BMU TB register.²³ The challenge of undertaking this task by matching limited identifiers from two different sources needs to be assessed in relation to the available time and human resources. In some countries, it may not be possible to identify a laboratory register for logistical reasons but

19 The BMU TB register (also sometimes called the district TB register) is intended primarily for recording the data needed to monitor BMU performance, using indicators and reports about TB patients. It is also commonly used to summarize testing results and treatment decisions in order to determine whether basic diagnostic and treatment guidelines are correctly implemented. No information that is beyond this monitoring scope should be included in the register. The register should contain the records of all patients diagnosed with TB and eligible for TB treatment, including those diagnosed with rifampicin-resistant TB or multidrug-resistant TB (MDR-TB), regardless of whether treatment was actually started. All of these cases are notifiable and should be included in the summary case notification reports sent to higher levels. The registration date is the date a patient is diagnosed with TB and is eligible for treatment. (Source: Compendium of indicators for monitoring and evaluating national tuberculosis programmes. Geneva, World Health Organization, 2004 (http://www.who.int/tb/publications/tb_compendium_of_indicators/en/, accessed 9 November 2013).

20 Depending on the organization of the NTP, some countries may have a TB register at the health facility level, which is their designated BMU, or they may have TB registers at the health facility and at a higher level, with the latter designated as the BMU.

21 Transferred-in patients that have been transferred from another TB register to continue treatment are not included in the receiving unit's quarterly and annual reports of case registrations and treatment outcomes. Misdiagnosed cases found not to have TB after they were registered as new TB cases, will need to be accounted for and deducted from the totals, that is, be de-notified.

22 In some surveillance systems, only smear-positive cases rather than bacteriologically-confirmed cases are recorded in the quarterly reports, therefore comparison with the laboratory register should be limited to these in order to make a direct comparison. Some laboratories may also serve other BMUs other than the one being assessed, so ensure only those for the assessed BMU are counted.

23 Patients who are diagnosed and registered in the laboratory of one particular BMU may be managed, and thus registered in another BMU.

every effort should be made to do so).

- (iii) Count the number of patient treatment cards at the health facility(-ies) reporting to the selected BMU, for the verification period, and compare the combined totals from each health facility against the number of cases reported in the BMU quarterly report, for the same verification period. (In some countries, TB patients keep the patient treatment cards and not the health facility. In these situations, this assessment cannot be undertaken. Ideally, all patients with treatment cards should be individually matched with the BMU TB register if time and human resources allow).
- (iv) In some countries, the BMU may be at the district level with many health facilities reporting to that BMU (Figure 1b). Each health facility may have its own TB register, in which case, for each health facility TB register(s) reporting to the assessed BMU, count the number of TB cases for the verification period, and compare the combined totals from each assessed health facility against the number of cases reported in the BMU TB register, for the same verification period. This combined total should be the same as that found in the BMU quarterly report for the corresponding verification period.

Figure 1a. Data flow: The data sources required to assess the second benchmark. The BMU is synonymous with the health facility. Solid lines represent primary assessments; dashed lines represent assessments if sources are available; dotted lines represent assessments if time and human resources allow.

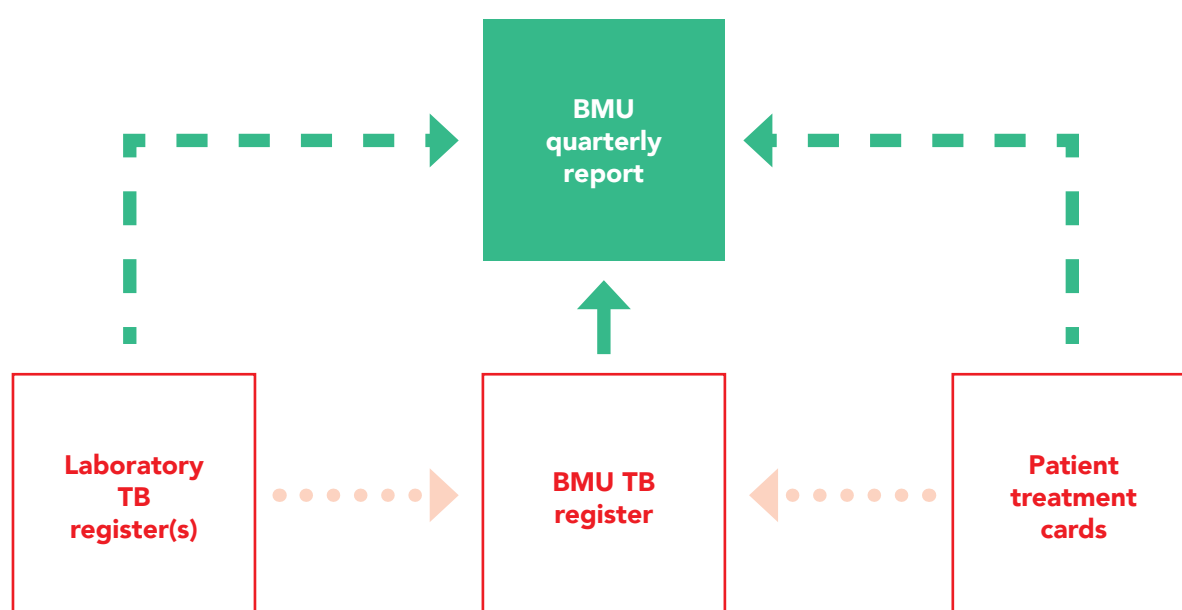
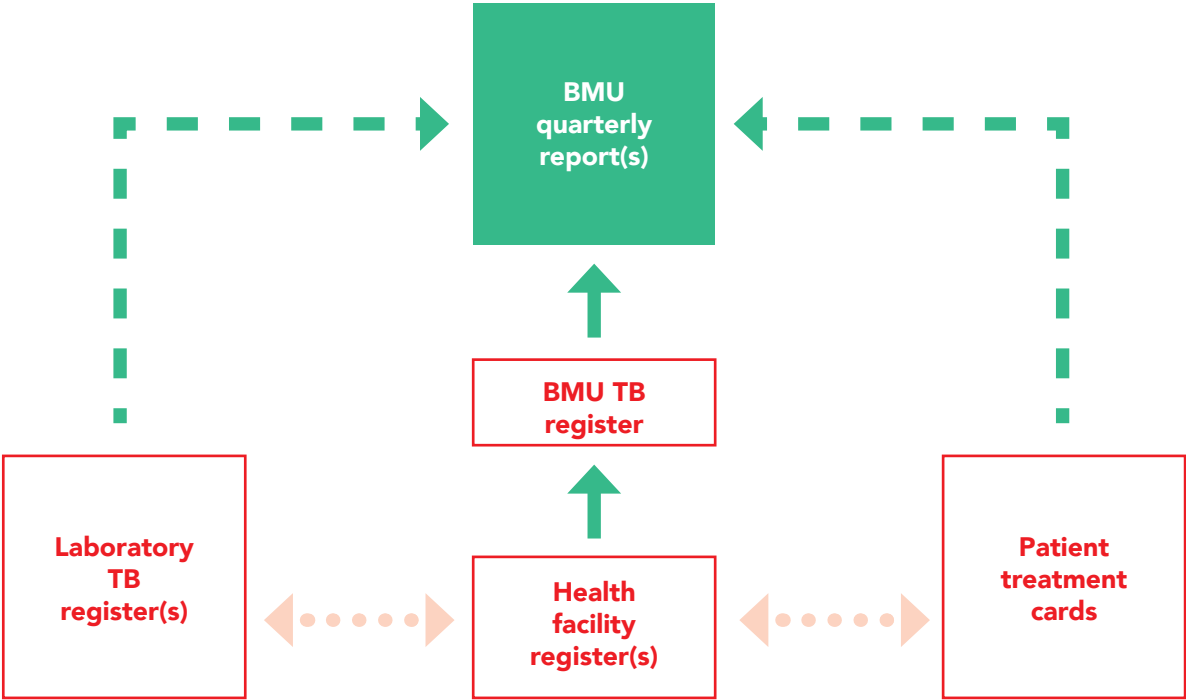


Figure 1b. Alternative data flow: In some countries, the BMU is at the district level with more than one health facility linked with it. This diagram reflects the data sources required to assess the second benchmark. Solid lines represent primary assessments; dashed lines represent assessments if sources are available; dotted lines represent assessments if time and human resources are available.



For the purposes of the third benchmark, for each selected BMU, the BMU TB registers for the verification period needs to be checked for completeness and data availability (see Appendix 1.2). As described for standard B1.2, these variables include year of registration, sex, age, disease classification, type of patient and bacteriological results. In the TB register, the proportion of TB cases with at least one of these six variables missing should be ascertained.

Main limitations: The application of LQAS theory in the context of this standard and its associated benchmarks is new and has not been extensively tested in the field. The time taken and human resources required to undertake the data source checks could be substantial. It may not be feasible to completely assess the second benchmark if the laboratory register(s) and/or treatment cards are not readily accessible

during the data audit. For logistical reasons, it may not be possible to assess all health facilities that report to the BMU.

Interpretation of results: There is a two-step process involved in assessing this standard. Firstly, each of the three benchmarks are *satisfied* if it fulfils the technical requirements for each assessed BMU, and if the total number of *unacceptable* BMUs for each benchmark is below a maximum allowable number, *d*, out of the total sampled.²⁴ Secondly, the standard is *met* if all three benchmarks are satisfied, *partially met* if only one or two of the benchmarks are satisfied, and *not met* if none of the three benchmarks are fully satisfied.

More specifically, for the purpose of the first benchmark, a BMU is classified as *acceptable* if from the selected quarterly report: 1) the sum of all age by sex sub-totals is exactly the same as the reported total; *and*, 2) the sum of all case type sub-totals is exactly the same as the reported total. The BMU is classified *unacceptable* in all other situations.

For the purposes of the second benchmark, a BMU is classified as *acceptable* if the total number of cases reported in the selected BMU quarterly report is exactly the same as: 1) the number of cases listed in the BMU TB register; *and, if available* 2) the number of bacteriologically-positive cases listed in the laboratory register(s) linked with that BMU; *and, if available* 3) the number of treatment cards of patients on treatment at the health facilities reporting to the BMU; *and, if applicable* 4) the combined total number of cases listed in all health facilities TB register(s) that report to the BMU. The BMU is classified *unacceptable* in all other situations. Therefore at a minimum, the BMU TB register should match the BMU quarterly report form.

For the purposes of the third benchmark, a BMU is classified as *acceptable* if for all variables in the minimum data set the proportion of TB cases without missing data is $\geq 95\%$. The BMU is classified *unacceptable* in all other situations.

Recommended actions: When benchmarks are not met, corrective actions should be taken such as M&E training and strengthening, in terms of frequency of visits and quality of supervision.²⁵ Given the inherent difficulties and limitations of paper-based systems, countries should consider moving to an electronic recording and reporting system.^{26,27} In addition, it may be necessary for the national guidelines for recording and reporting of data to be updated to reflect these benchmarks.²⁸ If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

²⁴ *d* is pre-determined during the sample size calculations for both the SARA and the independent data quality assessment (see Appendix 2 for details).

²⁵ See Parts A11, A12, A13.

²⁶ Electronic recording and reporting for TB care and control. Geneva, World Health Organization, 2011 (http://www.who.int/tb/publications/electronic_recording_reporting/, accessed 9 November 2013).

²⁷ See Parts A1, A2.

²⁸ See Part A10.

Examples

Country X is known to have a high burden of TB disease with an NTP network of 112 reporting BMUs e.g. districts. In this country, the BMU is at the health facility level with only one health facility reporting to each BMU. An independent data quality audit mission using LQAS was planned, with different scenarios for the required sample size to test benchmarks associated with the standard on data accuracy, completeness and internal consistency. The NTP chose a sample size of 50 BMUs, based on a decision interval of the number of allowable BMUs that would be deemed unacceptable, $d=0$, and a margin of statistical error of $\alpha=0.05$ (see Appendix 2).

From the exhaustive list of all 112 BMUs in the country, 50 were selected with a probability proportional to the number of TB cases diagnosed the previous calendar year. The audit team visited each of the selected BMUs and performed data audit activities - the results of which are summarized in the abridged Tables 1, 2 and 3.

Table 1. Assessing the 1st benchmark: comparing sex and case type disaggregated sub-totals with the reported total for the verification period in each of the visited BMUs' quarterly reports

	A	B	A+B	C	D	E	F	C+D+E+F	G	
BMU	Male	Female	Subtotal **	Smear- positive	Smear- negative	Smera not done/ NA	Extrapulmonary	Subtotal **	Total	Classification
1	10	11	21	8	7	2	4	21	21	Acceptable
2	150	101	251	98	35	64	54	251	251	Acceptable
3	71	50	121	30	22	44	25	121	121	Acceptable
...*
50	30	24	54	16	18	12	8	54	54	Acceptable

*All BMUs not shown in the table have been classified as acceptable.

**The two subtotals should be the same as the total from the quarterly report (i.e. column G).

Table 2. Assessing the 1st benchmark: comparing age disaggregated sub-totals with the reported total for the verification period in each of the visited BMUs' quarterly reports

Age group (years)										Classification
BMU	0-4	5-14	15-24	25-34	35-44	45-54	55-64	≥65	Total	
1	0	2	1	2	1	2	5	8	21	Acceptable
2	3	18	24	15	38	40	54	59	251	Acceptable
3	1	8	12	17	15	16	21	31	121	Acceptable
...*
50	0	3	7	5	8	12	9	10	54	Acceptable

*All BMUs not shown in the table have been classified as acceptable.

Out of the 50 sampled BMUs, none (less than the allowable decision interval $d=0$) were deemed unacceptable during the data quality assessment for the first benchmark. Therefore, the first benchmark was *satisfied*.

Table 3. Assessing 2nd benchmark for the verification period in each of the visited BMUs

BMU	n_1/n_2^*	n_3/n_4^{**}	n_5/n_2^{***}	Classification
1	21/21	8/8	21/21	Acceptable
2	251/251	98/98	251/251	Acceptable
3	121/121	30/30	115/121	Unacceptable
... ^{††}
50	54/54	16/16	48/54	Unacceptable

* n_1 =number of cases listed in the BMU TB register (in this example, this is also the health facility register) during the verification period; n_2 =number of cases reported in the BMU quarterly report during the same verification period.

** n_3 =number of all bacteriologically-confirmed cases listed in the laboratory register(s) linked with the BMU during the verification period, n_4 =number of bacteriologically-confirmed cases reported in the BMU quarterly report during the same verification period.

*** n_5 =number of treatment cards of patients on treatment at the health facility reporting to the BMU during the verification period.

†† All BMUs not shown in the table have been classified as acceptable.

Out of the 50 sampled BMUs, two were deemed unacceptable which are more than the allowable decision interval, d , therefore, the second benchmark was *not satisfied*.

Table 4. Assessing 3rd benchmark for the verification period in each of the visited health facilities of the BMUs. Percentage of unknown for each of the variables in the minimum set							
BMU	Age (%)	Sex (%)	Year (%)	Bacteriological result (%)	History of previous treatment (%)	Site of disease	Classification
1	0	1	2	8	5	2	Unacceptable
2	0	1	1	3	3	1	Acceptable
3	1	2	0	2	1	2	Acceptable
...*
50	0	2	2	4	1	2	Acceptable

*All BMUs not shown in the table have been classified as acceptable.

Out of the 50 sampled BMUs, one was deemed unacceptable, which is more than the allowable decision interval, d , therefore, the third benchmark was *not satisfied*.

The evidence produced from this data quality assessment meant that the country only *partially met* the standard because only one out of three benchmarks were *satisfied*.

Standard B1.5:

Data in the national database are accurate, complete, internally consistent and free of duplicates (For electronic case-based or patient-based systems only)

Benchmarks:

All benchmarks should be met to reach this standard:

- Data validation checks are in place at the national level to identify and correct invalid, inconsistent and/or missing data in the minimum set (Standard B1.2)
- For each variable in the minimum set (Standard B1.2), $\geq 90\%$ of case records are complete, valid and internally consistent for the year being assessed
- $<1\%$ of case records in the national dataset for the year being assessed are unresolved potential duplicates

Rationale for standard and benchmarks

The data provided by the national database needs to be trustworthy, unbiased and able to produce a reliable direct measure of TB incidence. Standard processes need to be in place to detect and fix errors and gaps in the minimum dataset (as per standard B1.2) so as to demonstrate that for the year being assessed, the dataset is free (or almost free) of problems. Outstanding unresolved errors need to be kept to a minimum otherwise the final dataset could produce biased and difficult-to-interpret results.

It is difficult to assess and fix data quality problems with paper-based aggregated data at the national level. Electronic patient- or case-based surveillance systems allow for better quality data because automated validation checks can be implemented for the minimum set of variables both at the time of data entry and also through subsequent querying of the database.

Definitions

The following is a brief description of some common data quality problems and common ways of detecting and preventing them.^{29,30,31}

29 Electronic recording and reporting for TB care and control. Geneva, World Health Organization, 2011 (http://www.who.int/tb/publications/electronic_recording_reporting/, accessed 9 November 2013).

30 Weiskopf NG, Weng C. Methods and dimensions of electronic health record data quality assessment: enabling reuse for clinical research. *Journal of the American Medical Informatics Association*, 2012 (<http://jamia.bmj.com/content/early/2012/06/24/amiajnl-2011-000681.full>, accessed 9 November 2013).

31 English L. Information Content Quality: Assessing the Quality of the Information Product. *International Association for Information and Data Quality Newsletter* 2 (3), 2006 (<http://iaidq.org/publications/doc2/english-2006-07.shtml>, accessed 9 November 2013).

Missing value: A fact about a TB case has not been recorded, for example a patient's date of birth or age. Essential fields such as those in the minimum dataset can be made mandatory fields at the time of data entry to prevent missing values, although there needs to be a balance between stringency and practicality to avoid making recording a case impossible when key data are not available. This difficulty can be reduced by having 'unknown' as an acceptable value in the database. However, for the purposes of this benchmark, 'unknown' is considered a missing value.

Inaccurate value: A recorded fact doesn't reflect reality, e.g. *date of birth* recorded as 01/01/1956 but patient's actual date of birth is 01/01/1965; name recorded as 'John Smyth' but the patient's actual name is 'Jonathan Smith'. If no other fields in the record can be used to check for inconsistencies (see below) then these errors are almost impossible to detect using algorithms, and would require a review of source documents. However, if an electronic system is actively used in care settings, by staff who know their patients, then these errors are likely to be noticed and fixed at the health care facility. It is much more difficult to spot such errors when data entry is a passive, one-way chore.³²

Invalid or unclear value: This is another type of inaccuracy. It often occurs when data are recorded in free text fields, for example if sex is recorded as 'under 15', which makes no sense. This is easily fixed in a data entry system by enforcing a closed set of data entry options, e.g. tick boxes or drop-down lists. Similarly, restrictions can be built to prevent invalid dates e.g. 31 February 2010, and to prevent text in numeric data fields e.g. 'ten' instead of 10.

Inconsistent or misclassified values: Here the combination of facts recorded about a case are contradictory and indicate that at least one of the facts cannot be true. Examples:

- a patient notified in 2012 with *date of birth* 01/01/1992 and with an age group recorded as over 65 years;
- a case with *date of notification* before the date of diagnosis;
- a TB patient classified with *case type* 'new' but a *date* is given for date of previous TB treatment;
- sex recorded as 'female' but patient's name or title (Mr) suggests the patient is male (this may be more difficult to spot using an algorithm).

Some of these problems can be prevented at the time of data entry

³² Electronic recording and reporting for TB care and control. Geneva, World Health Organization, 2011 (http://www.who.int/tb/publications/electronic_recording_reporting/, accessed 9 November 2013).

through the use of logical tests and by the use of automatically-calculated fields (for example age calculated from date of birth). It is always important to search for these errors by running a query on the national dataset (even though they can often be prevented by detecting them at the time of data entry).

Implausible value: An unrealistic value for a fact is recorded in a database record. This can happen with numeric and date fields, for example a patient in the year 2012 is recorded as having a *date of birth* of 01/01/1848. Acceptable ranges for numeric values and dates can be defined in the system and enforced at the time of data entry. It is always important to search for outliers and for values outside acceptable values in the national dataset, for example through plotting frequency charts or looking at the distribution of values in a dataset.

Duplicate records: More than one record in the national database refers to the same TB episode or case. Detecting duplicate case records can be made easier if national unique personal identifiers exist within a country, e.g. identity card numbers or social security numbers which are defined outside the TB surveillance system. If these are not available, sub-national or local patient identifiers can help. For example, a unique personal identifier could be made from a standard health facility code and a serial number specific to the health facility, e.g. the TB register number, but this must be accompanied by standard procedures to ensure: 1) that only one code is ever allocated to a TB case; and, 2) to manage the movement of a patient to another location that uses a separate numbering system. De-duplication algorithms will usually depend on 'fuzzy' or 'probabilistic' matching on other variables such as *name, sex, address, date of birth* and *date of diagnosis*.³³ Such algorithms take into account variations in spelling and recording of dates and rely on high completeness rates for such fields. They identify possible duplicates that then need to be checked manually to verify whether records of TB cases are indeed duplicates of other records documenting the same TB case/episode or are records of separate TB cases.

Methods to assess benchmarks

Data sources and data collection methods:

- Records of notified TB cases in the national patient- or case-based database for the year of assessment.
- System documentation, SOPs, the automated checks run at the time

³³ Assessing tuberculosis under-reporting through inventory studies. Geneva, World health Organization, 2012 (http://www.who.int/tb/publications/inventory_studies/en/, accessed 9 November 2013).

of data entry and the data queries used to check data quality at the national level.

- System documentation and/or SOPs for detection and removal of duplicate TB cases at national level.

Analysis: Examine whether standard procedures are used to identify and fix records with misclassified, implausible, inconsistent or missing data related to the minimum dataset (see the examples sections) to see how this is done in some countries. If not, additional tests may be needed when carrying out the data quality checks described in this section. Run the system data quality checks on the national dataset for the year of assessment. Calculate the percentage of records in which errors are flagged for each variable in the minimum dataset defined in standard B1.2 (age, sex, year of registration, bacteriological results, history of previous treatment and anatomical site of disease).

Examine the standard procedures in place for de-duplication. If feasible, run the duplicate detection algorithm used by the country on the national dataset for the year of assessment. Calculate the percentage of records that are flagged as extra duplicated case reports, e.g. in a pair of duplicated records only one is counted as the extra duplicate record. Note that this may not always be feasible because de-duplication is often a lengthy combination of automated and manual checks. An alternative is to examine the records kept of the de-duplication process and establish how many potential duplicates remain in the system after failing to get confirmation from other data sources or from reporting units. Alternatively, if a flagging system is used, calculate the percentage of records that remain flagged as potential duplicates.

Main limitations:

1. Variables and data structures in national databases may vary between countries, as may the type and rigor of data checks implemented within the surveillance system. Some examples of highly rigorous data quality checks from high-performing systems are provided below.
2. Duplicate cases can only be assessed in a case-based system that includes enough patient-specific information (name, date of birth, address, unique national identifier etc.) to allow duplicates to be identified and removed.

3. Not all data entry errors can be detected at the national level without checking for agreement with source documents or with patients themselves. For example, a date of birth of 01/01/1956 erroneously recorded as 01/01/1965 will be hard to identify at the national level (see discussion above).³⁴ A data audit could be conducted if considered useful (see Recommended actions below).

Interpretation of results: If standard data checking and fixing procedures are in place to identify records with misclassified, implausible, inconsistent or missing data then the first benchmark is satisfied. If $\geq 90\%$ records do not contain missing, inaccurate, invalid, inconsistent or implausible values for each of the variables in the minimum dataset (standard B1.2) then the second benchmark is satisfied. If $< 1\%$ of records are additional potential duplicate TB cases, then the third benchmark is satisfied.

If all benchmarks are satisfied then the standard is *met*. If the first is satisfied and only one of the other benchmarks are satisfied then the standard is only *partially met*. If the second *and* third benchmarks are not satisfied then the standard is *not met*; similarly, if the first benchmark is not satisfied irrespective of whether the second and/or third benchmarks are, then standard is also *not met*.

Recommended actions: If the standard is not met:

- Implement suggested data quality checks (see definitions section).³⁵
- Provide feedback to users at the sub-national level, for example share the results of analyses and show how the data are being used and how they benefit the NTP as a whole.³⁶
- Ensure data quality is highlighted and reviewed in quarterly or annual meetings.³⁷
- Develop data quality SOPs, guidelines and/or training materials and make sure staff are trained to understand and implement data quality SOPs.³⁸
- Secure senior management commitment to data quality and to good data governance.³⁹

Other options to consider include:

- Cross-check data from the national level dataset with sub-national level (district, provincial) data through periodic site visits to compare source data held on paper with data held electronically in order to ensure all data in the minimum dataset (in the electronic dataset) are

34 Electronic recording and reporting for TB care and control. Geneva, World health Organization, 2011 (http://www.who.int/tb/publications/electronic_recording_reporting/, accessed 9 November 2013).

35 See Part A7.

36 See Part A8.

37 See Part A7.

38 See Part A2, A11.

39 See Part A13.

accurate and complete.

- Establish a national data management unit, if one doesn't already exist, that would be responsible for data quality.
- Provide incentives for high quality data e.g. attach high quality data to funding contracts.
- Conduct operational research to evaluate and improve data quality in surveillance systems.
- If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Examples

The Netherlands

The Netherlands uses a single web-based TB notification system managed by The National Institute for Public Health and the Environment (RIVM). Clinicians at TB facilities can notify TB patients directly through the system. Data entry screens include data validation routines to prevent implausible or inaccurate entries, for example:⁴⁰

- *Date of diagnosis ≤ date of start of treatment ≤ date of end of treatment.*
- *Date of birth cannot be in the future.*
- *Date of birth cannot be before 1850.*
- Only pulmonary localization can be registered if the diagnosis is pulmonary TB and vice versa for extrapulmonary TB.

Data quality procedures are included in the programme's data management manual, which has been in use for many years.

Data managers in the national team check daily every newly registered case. If there is an inconsistency in the notification (for example if diagnosis is both pulmonary AND extrapulmonary TB but only a pulmonary localization is specified), a note is sent back to the local staff to adjust or clarify the notification before the notification is approved (In the previous example, the request would be to add the extrapulmonary localization or to change the diagnosis to *only* pulmonary). The surveillance system includes a *status of notification* field e.g. notified

40 Not an exclusive list. Source: Nico Kalisvaart, personal communication.

from municipal health centre to higher level; approved by national level. Treatment information is collected at a later stage and this information is also checked, e.g. is the length of treatment in line with the treatment outcome and the drug resistance status? If treatment is completed, treatment duration must be at least six months for drug-sensitive TB, or longer for drug-resistant TB.

De-duplication: Duplicate identification is carried out at the national level, using a combination of *date of birth*, *year of diagnosis* and *health care facility code* and if needed the search is expanded using other variables such as *sex*, *diagnosis* or postal code. Staff at the local level are contacted about possible duplicates. If local staff confirms that a record is actually a duplicate case, it is removed from the system by the municipal health centre and not by the data managers at the national level.

The national dataset for 2011 was examined and found to have:

- Age, sex, date of diagnosis, diagnosis (pulmonary or extrapulmonary TB) fields were 100% complete, valid and internally consistent.
- For the new/previously treated field: 100% of records had a value, but 7% of these had a value of 'unknown', therefore 93% of records are considered complete for this variable.
- For the culture status, 100% of records had a value, but 7% of these had a value of 'unknown', therefore 93% of records are considered complete for this variable.
- There was only 1 unresolved potential duplicate record, which is only (0.1%) of the total cases notified in 2011.
- The Netherlands system satisfied all three benchmarks and therefore the standard has been *met* for 2011.

United Kingdom

Public Health England (PHE)⁴¹ in the UK manages a national web-based surveillance system that collects TB data directly from TB clinics in real time, although London is not included and data from London are imported annually. PHE publishes an annual report which includes a section analysing data quality. A steering group with stakeholders' engagement oversees the national surveillance system. The principles of governance of the quality system include the commitment of senior management, a circle of local and national audits and evaluation, record-keeping to allow tracing of problems, customer input allowing

41 In April 2013, the Health Protection Agency formally changed its name to Public Health England.

for complaints and feedback, a documented development process that includes tests for quality and user satisfaction and mechanisms for implementing quality improvement through software development.⁴²

Completeness of case records: Mandatory fields include date of notification and date of birth so these fields should all be 100% complete. Sex, previous treatment, sputum smear status and site of disease are monitored for completeness. The proportion of records for which each of these variables was completed is calculated and published in the annual report, disaggregated by region. The target completion rate is 95% and where targets are not met the PHE investigates and takes action. For example, the PHE plans to increase the completeness of previous diagnosis and previous treatment fields by matching across previous years of data to identify those who have had a previous treatment history in the UK.

Consistency and accuracy: There are a limited number of free text fields within the system's data entry screens and most use pre-defined drop-down menus to ensure valid values, e.g. sex, site of disease. Each case is assigned a unique identifier by the system. Data validation checks are also implemented to prevent errors at data entry, for example:

- Date symptom onset is not before the date *first presented* which is not before the *treatment start date*.
- *Date of birth* cannot be in the future.
- Age is calculated directly from the *date of birth*.
- *Treatment outcome* is not available for the user to fill in less than six months after *Notification date*
- Name and other key patient identifiers are used to check if a similar record already exists in the system. If so, an alert is shown to the user to check that the record will not be a duplicate.

At a regional level, quarterly reports are produced by most regional coordinators according to an SOP that involves running standard queries (developed by the national team) on regional datasets to identify data problems and following these up with clinics. Although not all regions produce the quarterly reports, all are following up missing data. Examples of inconsistencies identified by the queries include:⁴³

- Age <16 years is inconsistent with occupation.
- Age >16 years but occupation is child.
- Age is less than the number of years since a previous diagnosis of TB.

42 See chapter 5 of Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK, 2012. London, Health Protection Agency, 2012 (http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1317134913404, accessed 9 November 2013).

43 Not an exclusive list. Source: Laura Anderson, personal communication.

- *Date of birth* is after case report date.
- *Year of arrival into the UK* is incompatible with age.
- *Previously treated* ticked but *previously diagnosed* not ticked.
- *Pulmonary site of disease* not ticked but case has a positive result from a *pulmonary specimen*.
- *Other non-pulmonary site* is ticked but accompanying free text suggests pulmonary disease.
- Case found not to have TB but has not been de-notified.

Matching laboratory and notification records: This is to ensure that all detected cases have been notified. Laboratory isolates are matched to notifications by direct automated matching in the system using *name* and *date of birth*. Cases can also be matched manually by the laboratory. Cases left unmatched at the end of the year are matched using probabilistic matching. Queries are run monthly to identify laboratory isolates with missing drug sensitivity test (DST) results and implausible values for specimen dates (future dates). These are then followed up with the reference laboratories and amended.

De-duplication: Duplicates are removed on an annual basis before data analysis begins. Although some duplicates can be prevented at data entry (see above), some do occur because of differences in spelling *name* or errors in *date of birth*. Records imported from London (which does not use the nationwide system) also need de-duplicating against records already in the national system. Duplicates are identified by direct matching using *name* and *date of birth* and notification date within 12 months.

Other regular checks conducted by the national team: These include:

- Follow-up of drug-resistant cases confirmed at the laboratory for notification.
- Follow-up of de-notified cases with confirmed laboratory results for contamination. If so, the case is recorded as a false positive on a special form; if not, it is discussed with the clinic. If wrongly de-notified then the case is re-notified. If treatment was inadequate, the national team recommends a follow-up appointment with clinical assessment/appropriate diagnostics.
- Follow up with laboratories for missing or future specimen dates and DST results.

- Cross-reference with prison health log and follow up of prisoners not notified or correction of the risk factors if they are not reported.
- Cross-referencing those with cause of death with the mortality registry.
- Follow up with clinics for missing data or notification for patients in molecular clusters (part of national strain typing service).

The UK system satisfied two out of three benchmarks. It was not possible to show whether the completeness benchmark for bacteriological confirmation was met or not and therefore the standard is considered to be *partially met*.

United States of America

There are 60 reporting areas in and out of the USA that now have their own software or can opt to use CDC-generated software to report their TB case data to the CDC. Those using one of the CDC applications have standard built-in electronic validation rules defined by CDC. Those using their own software may or may not have validation rules applied (but they are encouraged to use the CDC standardized validation rules). CDC receives very limited patient identifiers from the reporting area, such as date of birth, sex and a unique state case number issued by the reporting areas. Full identifying information is maintained at the state and local level, where legal custody and authority of such information resides.

CDC has developed data quality assurance resources such as guidelines, flowcharts and tools, and has developed training courses and self-study modules for staff in the reporting areas.⁴⁴

The 60 reporting areas send annual data files to the central database managed by CDC. Once data are received by CDC, a set of 211 validation rules are applied. Some of the rules related to the minimum dataset include:⁴⁵

- The *State case number* must be alphanumeric and unique.
- The *Date of birth* must be equal to or before many other dates reported for a 'verified case of tuberculosis' such as *Date reported*, *Date therapy started*, *Date sample collected* (e.g. for DST, sputum smear, sputum culture, etc.), *Date results reported* (again for various tests), *Date of death* and *Date arrived in USA*.
- The *Date of birth* must be equal to or after 01/01/1890.

44 Manangan LP, Tryon C, Magee E, Miramontes R. Innovative quality-assurance strategies for tuberculosis surveillance in the United States. *Tuberculosis Research and Treatment*, 2012. doi:10.1155/2012/481230.

45 Not an exhaustive list. Source: Lori Armstrong, personal communication.

- If *Sex at birth* is male then *Smear/Pathology/Cytology of tissue anatomic code* and *Culture of tissue anatomic code* must not equal Milk/ Vulva/ Vagina/ Uterus/Cervix/ Endometrium/Myometrium/ Fallopian/Ovary/Female Genital Fluids/Placenta/Fetus.
- If *Sex at birth* is female then *Smear/Pathology/Cytology of tissue anatomic code* and *Culture of tissue anatomic code* must not equal Penis/Prostate/Testis/Epididymis/ Male Genital Fluids.
- *Sex at birth* must be either male or female.
- *Date reported* must be equal to or after many other dates such as 01/01/1990, *Date counted*, *Year of previous diagnosis*, *Date arrived in USA* and *Date of birth*.
- *Date reported* must be equal to or before the current date.
- *Site of disease* must not contain a value not included in the list of valid values.
- *Site of disease* must be one of Pulmonary, Pleural, Lymphatic Intrathoracic, (other) Nose, Accessory Sinus, Nasopharynx, Laryngeal or Trachea if Sputum smear is positive or if *Sputum culture* is positive.
- *Previous diagnosis of TB* must be yes if *Year of previous diagnosis* is not blank.
- *Year of previous diagnosis* must be equal to or after 1900.
- *Year of previous diagnosis* must be equal to or before other years such as year of *Date FIRST isolate collected for DST*, year of *Date of the FIRST consistently negative culture*.
- *Sputum culture* must be positive if *Sputum culture conversion documented* is yes.
- *Culture of tissue* must be blank if *Site of disease* is blank or *Site* not stated.
- *Sputum culture* or *Culture of tissue* and *Other body fluids* must positive If *drug susceptibility testing* was done is yes.

When data do not meet any one of these rules an error report is sent back to the reporting jurisdiction so that they can fix the error and resend the data. After the revised data are received at CDC, a further set of rules are applied to produce a 'clean' dataset for analysis and reporting. The 'cleaning' replaces uncorrected data with blanks.

In addition, CDC provides feedback to the reporting jurisdictions via

two limited access websites. The first reports completeness indicators on selected data items based on the NTP Objectives and Performance Targets for 2015.⁴⁶ The second, accessible by TB programme managers at state level and CDC TB surveillance personnel, shows line-listed data of missing variables for individual TB cases covering all variables recorded for a TB case.

CDC requests that all variables in the standard reporting form (*Report of Verified Case of TB*) are 95% complete when they are reported to CDC at year-end.

In 2011, variables used to record the minimum dataset defined in standard B1.2, i.e. age, sex, year of registration, bacteriological results, history of previous treatment and anatomical site of disease, were all $\geq 95\%$ complete.

De-duplication is the responsibility of each jurisdiction and no checks for duplicates are conducted on the national dataset (the national dataset does not include patient names or addresses so de-duplication would be far more difficult). Each case has a unique number issued by a state and the software warns users if they try to assign a previously-assigned number to a case. Each jurisdiction also has a written protocol on how to alert each other when a patient moves from one jurisdiction to another so that the case is not reported twice at the national level.

The USA system satisfied two out of three benchmarks. It was not possible to show whether the benchmark on de-duplication at national level was met or not and therefore the standard is considered to be *partially met*.

46 <http://www.cdc.gov/tb/programs/Evaluation/Indicators/default.htm>.

Standard B1.6:

TB surveillance data are externally consistent

Benchmark:

- Among new TB cases, the percentage of children diagnosed with TB is between 5–15% in low- and middle-income, and <10% in high-income countries

Rationale for standard and benchmarks

The evaluation of the external consistency of the data should be an integral part of the assessment of any TB surveillance system. To assess external consistency national (and sub-national) TB surveillance data are compared against a plausible range of values that are based on what is known about the global epidemiology of TB. It is well known that, all other things being equal, a relatively small proportion of TB cases occur among children. The percentage of cases occurring among children is also an indicator of the amount of recent infection in a country and/or the performance of the surveillance system to capture diagnosed cases of TB among children.

Method to assess benchmark

Data sources and data collection methods: Reported TB case data from the national routine TB surveillance system disaggregated by age are needed. National surveillance data from the last year for which complete data are available should be analysed to determine the percentage of new TB cases that are aged <15 years among all new TB cases. Country income grouping can be found at the World Bank website.⁴⁷

Main limitations: The ranges of values used for the benchmark are based on recently observed values from national surveillance data of countries around the world and TB epidemiological studies, but it is conceivable that a few countries with accurate data and high-performing surveillance will fall outside these ranges.

Interpretation of results: Having values within the ranges stated in the benchmark values provides some reassurance of a well-functioning surveillance system, thus the benchmark is satisfied and the standard is considered to be *met*. Reasons why the standard has not been met

47 <http://data.worldbank.org/about/country-classifications/country-and-lending-groups>.

should be taken into consideration especially in settings with accurate data and high-performing surveillance systems, and explained in the “Results (Description)” column of the checklist.

Recommended actions: An unmet benchmark should be investigated further and potential reasons for discrepancies should be hypothesized, e.g. over- or under-diagnosis of childhood TB cases. Check for errors and correct them at all administrative levels, including source data verification and modification. (Although it will not be used to assess the benchmark itself, it may be worth exploring time trends in this benchmark in case there were concerns about internal consistency).⁴⁸ There are good examples of alternative explanations for internal and external consistency in the WHO Task Force Assessment of surveillance data workbook.⁴⁹ If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Examples

United Republic of Tanzania, Lao PDR and Italy

National TB surveillance data (new smear-positive, smear-negative and extrapulmonary cases) from United Republic of Tanzania, Lao PDR and Italy, which are low-, middle- and high-income countries, respectively, are presented below for the year 2011 (Table 1). At the time of the assessment, data from the last year for which complete data were available were analysed to obtain the percentage of cases aged <15 years among all new TB cases reported in 2011.

Table 1. National TB surveillance data from countries. United Republic of Tanzania, Lao PDR and Italy			
	United Republic of Tanzania	Lao PDR	Italy
Country income group	Low	Middle	High
Number of childhood/total new TB cases (SP, SN, EP)	4 883/58 278	52/4 136	113/12 511
Childhood TB (% over total number of new cases)	8.4%	1.3%	9.0%

48 See Parts A2, A4, A6, A10, A11.

49 Assessment of surveillance data workbook. Geneva, World Health Organization, 2009 (http://www.who.int/tb/advisory_bodies/impact_measurement_taskforce/meetings/ie_oct09_workbook.pdf, accessed 9 November 2013).

For **United Republic of Tanzania**, the percentage of reported TB cases that are among children in 2011 is 8.4%, which fits within the range of 5–15%, so this benchmark was satisfied and the standard was *met*.

For **Lao PDR**, the percentage of reported TB cases among children (1.3%) is considerably less than the plausible range of values (5–15%) in a middle-income country, therefore the standard was *not met*. This is most likely due to the under-diagnosis of childhood TB in the health centres and hospitals, and the under-reporting of cases to the NTP. However, since 2010, the NTP has been placing more focus on childhood TB reporting and is organizing more training sessions on the diagnosis and management of childhood TB with paediatricians throughout the country.

For **Italy**, 9.0% of reported TB cases are among children based on analysis of 2011 data, so this high-income country *met* the standard.

Standard B1.7: TB surveillance data are internally consistent over time

Benchmarks:

If vital registration data are available, then the following benchmark should be satisfied for this standard to be met:

1) Year-to-year change in the national number of reported TB cases is consistent with the year-to-year change in national TB mortality (HIV-negative, from national vital registration) i.e. trajectories with the same direction.

If vital registration data are not available, then the following benchmarks should be satisfied for this standard to be met. At the national level, evidence of internal consistency over the previous five years for the following benchmarks:

- 2) Ratio of notified pulmonary to extrapulmonary TB cases
 - 3) Ratio of male to female TB cases
 - 4) Proportion of childhood TB cases out of all TB cases
 - 5) Year-to-year change in the case notification rate for all forms of TB
 - 6) Year-to-year change in the case notification rate for new smear-positive TB
- and if data are available,
- 7) Ratio of the number of people with presumptive TB to total notifications of TB cases

Rationale for standard and benchmarks

TB is a relatively rare disease (case rates are typically expressed per 100 000 population per year) with an extremely wide range of incubation periods (from days to decades after infection), and a relatively low virulence (a limited proportion of infected individuals will develop the disease in their lifetime). The prevention of TB is based on a combination of: 1) a vaccination programme with limited epidemiological impact; 2) prophylaxis in infected individuals with very low overall population coverage; 3) infection control in health settings; and, 4) treatment of detected cases that prevent transmission within days to weeks of treatment initiation. This combination of tools complements improvements in economic and living conditions, but has a relatively low epidemiological impact as evidenced by historical data

on case rates changing fairly slowly over time. The 10% annual decrease of TB cases in Western Europe post-World War II was the “best ever performer” documented with both chemotherapy, social and economic improvements, while the “best recent performers” are Cambodia and China with a 3-5% annual decrease in TB cases.

All other things being equal, trends in case counts are expected to be in the same direction as trends in measured TB mortality. However, in 2013 there were approximately 120 countries with vital registration (VR) data of sufficient coverage and quality to measure trends in mortality, leaving almost 100 countries without such data (including most African nations). To enable the standard to be assessed in countries *without* VR data, other benchmarks for internal consistency have been included based upon data that most countries should be collecting. Given what we know about TB epidemiology, year-to-year changes in notification rates and case counts by type of TB, sex and age are not expected to be dramatically different.

Methods to assess benchmarks

Data sources and data collection methods:

For countries with available Vital Registration data:

Case notification rates (all forms of TB) at the national level are obtained from routine TB surveillance. TB mortality rates (HIV-negative TB) at the national level are obtained from VR systems.

For countries without available VR data:

For the past five years, reported national level TB case data disaggregated by age (or age group), sex, type of disease, along with notification rates (all forms and smear-positive TB cases) are required. Although, not a part of the benchmark assessment itself, similar sub-national data should also be collected and examined. Childhood TB cases (0–14 years) should be collected for all forms of TB. However in settings where data for all forms of childhood TB are not being reported, the proportion of new smear-positive cases in children out of all new smear-positive cases can be used. Some countries may have registries with listings of people with presumptive TB at the national level. Such data can provide an indication of the total number of people who were referred for diagnostic testing.

Analysis:

For countries with available Vital Registration data:

A five-year trend should be obtained by fitting a linear regression of log-transformed rates over time. The slope represents the average rate of change per unit of time (typically one year). The sign of the slope provides the direction of the trend. The benchmark is met if the slope for case notification rates is of the same sign as the slope for TB mortality rates. The test is applied to national data, but can also be applied to rates for first-level administrative areas. Raw TB mortality data from VR systems need to be adjusted to account for incomplete coverage and for ill-defined causes of death (i.e. ICD-9 code B46, ICD-10 codes R00–R99).⁵⁰ For VR-recorded deaths with ill-defined causes, it can be assumed that the proportion of deaths attributable to TB is the same as the observed proportion in recorded deaths with well-defined causes. The adjusted number of TB deaths, d_a , is therefore obtained from the VR report, d , as follows:

$$d_a = \frac{d}{c(1 - g)}$$

where c denotes coverage, i.e. the number of deaths with a documented cause divided by the total number of estimated deaths, and g denotes the proportion of ill-defined causes. These adjusted TB mortality rates from vital registration systems can be obtained online from WHO.⁵¹ If this benchmark can be assessed, there is no need to assess the other six benchmarks, although such analyses can help provide NTPs with other important insights into their TB surveillance systems.

For countries without available Vital Registration data:

The assessment of benchmarks 2 to 7 is based upon the analysis of data and expert opinion about the interpretation of results. Reliance on expert opinion is necessary because it has not yet been possible to define quantitative tests for internal consistency that are applicable in all settings. Even in countries with high-performance surveillance systems, assessment of the internal consistency of data is often based on expert opinion. Given this reliance on expert opinion to interpret findings, the assessment of whether or not benchmarks are met should be made by at least two independent assessors, along with input from the NTP. If all are in agreement that the data provide convincing evidence that the core benchmarks are satisfied, the standard can be considered to be met.

50 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.

51 The dataset can be obtained here: <http://www.who.int/tb/country/data/download/en/index.html>. The adjusted mortality rate (variable: "e_mort_exc_tbhiv_100k") for the selected years should only be for those countries where the source of the data (variable: "source_mort") are either "VR" or "VR imputed".

Main limitations:

Although all countries have TB surveillance systems for TB cases, a VR system with sufficient coverage and quality (in which causes of death are recorded according to the ICD system; see standard B1.10) is still lacking in approximately 100 countries. For example, in the 2012 *Global Tuberculosis Report*, direct measurements of trends in mortality from VR systems were available for 119 out of 217 countries and territories. The VR benchmark will fail more often if: 1) there are small numbers of cases, which result in large stochastic fluctuations; 2) slopes are very close to zero compared with slopes showing a greater departure from zero; and, 3) trends in mortality or in case notifications change direction, e.g. rise and fall with a peak in the middle of the studied period. It should be noted that changes in the case fatality rate (resulting in more TB deaths per 100 notified cases on average), e.g. due to a rise in drug-resistant TB, are not expected to result in mortality and case notification rates with opposition directions.

Despite considerable efforts, it was not possible to identify suitable quantitative tests (benchmarks) for internal consistency that are universally applicable for countries without a VR system. As every country is in a different state of the epidemic, it is currently not feasible to set a standardized limit for the variation of internal consistency. Therefore, expert opinion is used to interpret the results of analyses and to determine whether the standard is met. Examples of alternative methods that may assist in guiding expert opinion as used by other countries to ensure and improve internal consistency of national level TB surveillance data are shown elsewhere (Appendix 3).

Interpretation of results:

For countries with available Vital Registration data:

Regarding the VR benchmark, if trends in case notification rates and TB mortality rates have opposite directions, the benchmark is not satisfied. If the VR benchmark is satisfied then the standard is *met*.

For countries without available Vital Registration data:

All five benchmarks (or six, if presumptive TB data are available) need to be satisfied in order for the standard to be *met*. A benchmark can be satisfied if data are inconsistent but explainable. For example, a dramatic, yet inconsistent decline in the year-to-year change in case notification that was investigated to be a true decline satisfies the

benchmark. However, if a dramatic change in the sex ratio from one year to the next was found to be due to staff not submitting their reports, then this is explainable but the lack of internal consistency is due to a surveillance artefact rather than a true change. Hence in this situation, the benchmark is not satisfied. Such explanations should be documented in the checklist. If at least one but not all benchmarks are satisfied, then the standard is *partially met*. If no benchmarks are satisfied, then the standard is *not met*.

Recommended actions:

When trends are considered inconsistent, countries should attempt to understand why this is the case. Reasons for rapid time changes, or for inconsistent trends in case reports, should be investigated. Check for errors and correct them at all levels, including source data verification and modification.⁵² Keep track of programmatic changes and changes in case definitions.⁵³ Recent changes in the coverage of routine TB surveillance may account for opposite trajectories in TB cases and deaths and this possibility should be investigated. Explore if differences are due to true differences in the TB epidemic nationally over time or sub-nationally (TB determinants such as HIV prevalence, urbanization and socioeconomic situation, or effect of TB control activities, etc.), or due to differences that are primarily a result of changes in case definitions or in the recording and reporting system (e.g. structure, coverage or performance of the notification system). If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge the gap(s) would assist in developing an investment plan.

52 See Parts A6, A7

53 See Parts A10, A15.

Examples

Countries with Vital Registration data

In country X and country Y, case notification rates and adjusted TB mortality rates are as below:

Year	Country	Case notification rate (per 100 000)	Adjusted TB mortality rate (per 100 000)
2007	X	41	3.1
2008	X	40	3.0
2009	X	39	2.9
2010	X	38	2.9
2011	X	37	2.8
2007	Y	56	8.6
2008	Y	54	8.7
2009	Y	51	8.8
2010	Y	46	8.9
2011	Y	41	8.9

A linear regression model is fitted to log-transformed case notification rates and to log-transformed adjusted TB mortality rates. The slope is then extracted from each model. The slope describes the average rate of change per year. We compare the sign of the two slopes.

In the R statistical environment,⁵⁴ the data are entered and the test is done as follows:

```
# Country X
>notif<- c(41, 40, 39, 38, 37)
>death<- c(3.1, 3.0, 2.9, 2.9, 2.8)
>time<- 2007:2011
>slope.notif<- coef(lm(log(notif) ~ time))[2]
>slope.death<- coef(lm(log(death) ~ time))[2]
>print(c(slope.notif, slope.death))

yryr
-0.02566 -0.02374
```

For country X, the average yearly change in case notification rate is -0.02566, i.e. a fall of 2.6% per year, and the average yearly change in mortality rate is -0.02374, i.e. a fall of 2.4% per year. The slopes are both negative and this benchmark is therefore *satisfied*.

For country Y, the average yearly change in case notification rate is -0.07839, i.e. a fall of 7.8% per year, and the average yearly change in mortality rate is 0.00913, i.e. an increase of 0.9% per year. The slopes have opposing trajectories (one negative, the other positive) and therefore this benchmark is *not satisfied*.

```
# Country Y
>notif<- c(56, 54, 51, 46, 41)
>death<- c(8.6, 8.7, 8.8, 8.9, 8.9)
>time<- 2007:2011
>slope.notif<- coef(lm(log(notif) ~ time))[2]
>slope.death<- coef(lm(log(death) ~ time))[2]
>print(c(slope.notif, slope.death))

yryr
-0.07839 0.00913
```

54 <http://www.r-project.org>.

Using Stata (v12.1) statistical package, the same results are also generated with the following commands:

```
input year rate_x death_x rate_y death_y
2007 41 3.1 56 8.6
2008 40 3.0 54 8.7
2009 39 2.9 51 8.8
2010 38 2.9 46 8.9
2011 37 2.8 41 8.9
end

gen lograte_x=ln(rate_x)
regress lograte_x year

*// Coefficient = -0.0256602

gen logdeath_x=ln(death_x)
regress logdeath_x year

*// Coefficient = -0.0237467

gen lograte_y=ln(rate_y)
regress lograte_y year

*// Coefficient = -0.0783902

gen logdeath_y=ln(death_y)
regress logdeath_y year

*// Coefficient = 0.0091306
```

An example of countries without a vital registration system

Standard B1.7 can be examined by analysing routine surveillance data from the previous five years. In countries that are still developing a robust VR system, analysis of all the benchmarks excluding the one on VR is sufficient to assess this overall standard. Uganda's assessment of internal consistency is presented below as a guiding example for standard B1.7.

Benchmark 2: Ratio of notified pulmonary to extrapulmonary TB cases

Data requirements: The total number of notified pulmonary (PTB) and extrapulmonary (EPTB) cases reported nationally for each of the previous five years are required. All new and relapsed TB cases should be included in the analyses.⁵⁵

Conducting the analysis: Calculate the ratio of PTB to EPTB cases for each year (Table 1).

Table 1. The ratio of the number of notified TB pulmonary cases to extrapulmonary TB cases in Uganda, 2008 – 2012	
Year	Pulmonary: Extrapulmonary TB
2008	7.96
2009	7.52
2010	8.38
2011	8.26
2012	7.72

Interpretation: During 2008–2012 the ratio of pulmonary to extrapulmonary cases notified nationally ranged from 7.52–8.38, with a slight increase in pulmonary cases during 2010 and 2011. This suggests that these data are generally internally consistent. This benchmark is considered to be satisfied. If the data are available, this analysis may also be conducted at the sub-national level.

⁵⁵ Relapse patients have previously been treated for TB, were declared cured or treatment completed at the end of their most recent course of treatment, and are now diagnosed with a recurrent episode of TB (either a true relapse or a new episode of TB caused by reinfection). See: Definitions and reporting framework for tuberculosis – 2013 revision. Geneva, World health Organization, 2013 (http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf, accessed 9 November 2013).

Benchmark 3: Ratio of male to female TB cases

Data requirements: The total number of notified TB cases reported nationally for males and for females for each of the five previous years are required. All new and relapsed TB cases should be included in the analysis.

Conducting the analysis: Calculate the ratio of male to female TB cases for each year (Table 2):

Table 2. The ratio of the number of male to female TB cases (new and relapsed) in Uganda, 2008–2012

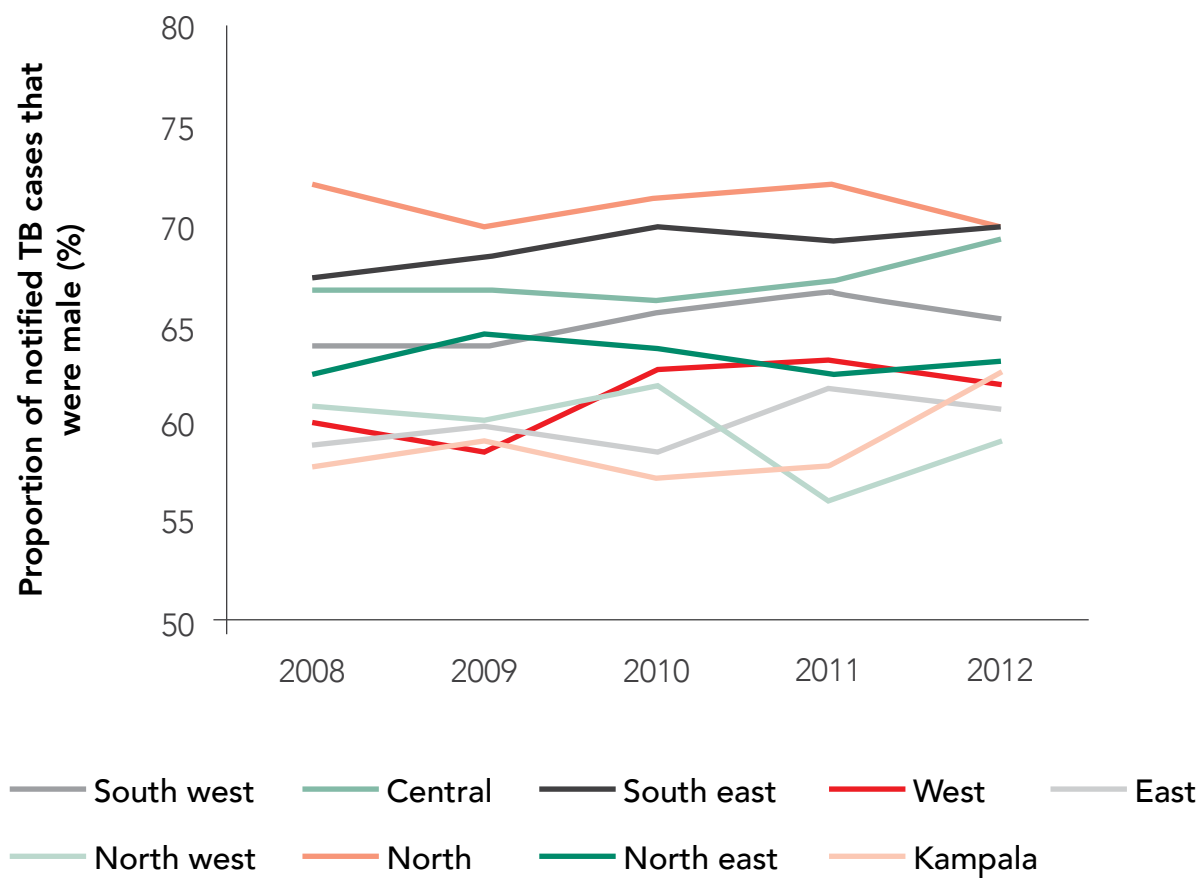
Year	Male: Female TB cases
2008	1.47
2009	1.52
2010	1.58
2011	1.63
2012	1.64

Interpretation: During 2008–2012 the male to female ratio of notified TB cases was generally consistent, but there was a notable increase in the proportion of male cases over time. The increasing ratio could suggest women were less likely to seek health care as the years progressed and/or were underdiagnosed when they did seek care; that men became more likely to seek health care and thus comprise a greater percentage of the TB cases; or that epidemiologically, there was a real increase in men (or decrease in women) who developed TB in Uganda over these years. Overall, this analysis suggests that the data have good internal consistency, but it would be beneficial to look further into why this ratio was increasing over time. This benchmark was deemed to be satisfied.

Additional analyses: If the data are available, this analysis may also be conducted at the sub-national level. For the most part, data are internally consistent when observed at the sub-national level (Figure 1). However some zones have greater variation (e.g. in the North west zone, where the number of male TB cases decreased almost 10% from 2010 to 2011, before increasing 5% the following year). In contrast, in Kampala, the

number of male TB cases saw a large unexplainable increase from 2011 to 2012. Investigating the reasons behind these inconsistencies would be beneficial to the NTP (i.e. are these real changes in notifications and if so, why? Or are they data entry errors?)

Figure 1a. The proportion of reported TB cases that were male by TB reporting zone, Uganda, 2008–2012



Benchmark 4: Proportion of childhood TB cases out of all TB cases

Data requirements: The total number of nationally notified childhood TB cases (0–14 years) and the total number of all notified TB cases (any age) for each of the five previous years are required. While sub-optimal, the total number of new smear positive childhood TB cases could be used in conjunction with the total number of all smear-positive TB cases in countries where only smear-positive childhood TB data are available. All new and relapsed TB cases should be included in the analyses.

Conducting the analysis: In Uganda, data were available only for new smear-positive childhood TB cases. Therefore, the following proportions of smear-positive childhood TB cases out of all smear-positive TB cases were calculated for each year (Table 3):

Interpretation: Although the proportion of children reported to have

Table 3. The number of new smear-positive childhood TB cases out of all smear-positive TB cases (any age), Uganda, 2008 – 2012

Year	Pulmonary: Extrapulmonary TB
2008	0.031
2009	0.025
2010	0.028
2011	0.027
2012	0.026

smear-positive TB out of all smear-positive TB cases decreased over time, there was a larger decline between 2008 and 2009, before rising again in 2010. This variation requires further investigation, and as such, this benchmark is *not* deemed to have been satisfied.

Benchmark 5: Year-to-year change in the case notification rate for all forms of TB

Data requirements: National case notification rates for all forms of TB for the previous five years are required. National population data are also required for each year. All forms of TB should be included in this analysis (Table 4).

Conducting the analysis: In Uganda, data were available only for new smear-positive childhood TB cases. Therefore, the following proportions of smear-positive childhood TB cases out of all smear-positive TB cases were calculated for each year (Table 3):

Table 4. Total number of notified TB cases (all forms) out of the total population, Uganda, 2008 – 2012

Year	Pulmonary: Extrapulmonary TB	% change from the previous year
2008	144.7	-
2009	137.7	-5.1
2010	134.1	-2.3
2011	140.6	+4.9
2012	130.2	-7.4

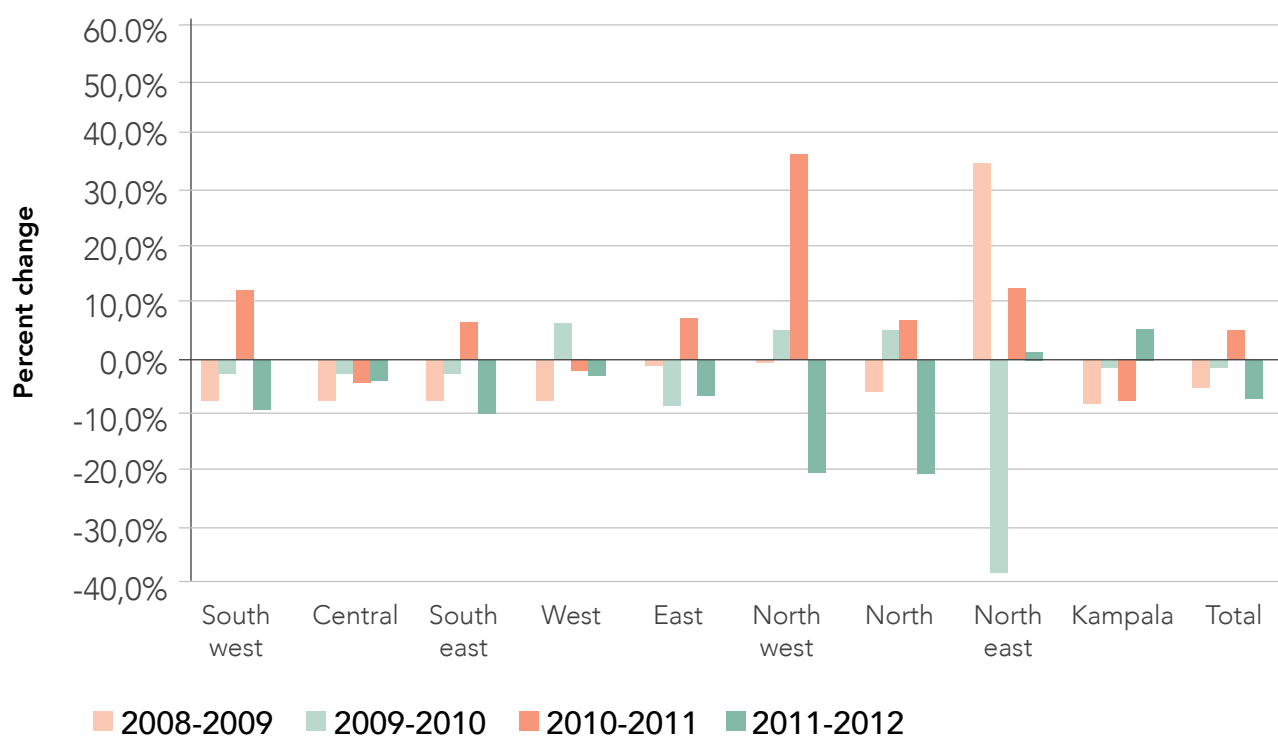
Analysis and Interpretation: These data show that the TB case notification rates at a national level are decreased between 2008 and 2012 in Uganda. However, the percentage change in TB case notification rate at the national level shows much variability by year over time, which therefore suggests that the data are not internally consistent, and thus this benchmark is *not* satisfied

Additional analysis: The analysis can also be conducted if sub-national level notification and sub-national population data are also available. In Uganda, TB notification rates were also compared from 2008 to 2012 by geographical zone (Figure 2). These data suggest that reporting is generally consistent in most zones, with most variability occurring in a few zones, e.g. North west, North east and Kampala. However, the percentage change in TB case notification rate at the sub-national level shows much variability by year over time and by reporting region (Figure 3), which helps to identify the areas that require further investigation and intervention.

Figure 2. TB notification rates by TB Zone, Uganda, 2008–2012



Figure 3. Percentage change in TB case notification rate by TB reporting region, Uganda 2008–2012



Benchmark 6: Year-to-year change in the case notification rate for new smear-positive TB

Data requirements: National case notification rates for smear-positive TB for the previous five years are required. Only new smear-positive cases are included in this analysis. As this is a rate, national population data are also needed for each year.

Conducting the analysis: Calculate smear-positive notification rate per 100 000 (Table 5).

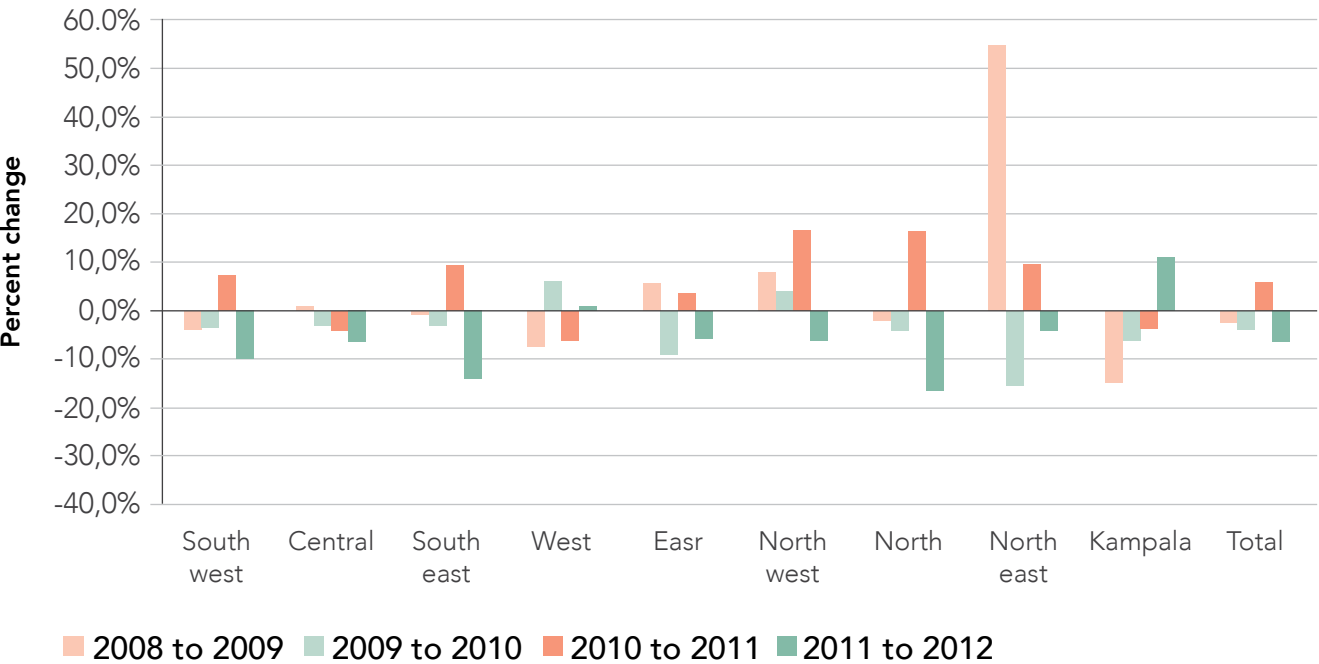
Table 5. Number of smear-positive TB case rate in Uganda, Uganda, 2008 – 2012		
Year	Smear positive case notification rate (per 100 000)	% change from previous year
2008	78.1	-
2009	76.1	-2.5
2010	73.3	-3.6
2011	77.8	+6.1
2012	72.7	-6.5

Analysis and Interpretation: These data show that at a national level there is a decline in smear-positive notification rates between 2008 and 2012 throughout Uganda. Much like in Benchmark 4, there was a sharp rise in 2011 with an unexpected subsequent decline the following year. Overall, the percentage change in TB case notification rate by years at the national level shows much variability by year over time. Therefore this suggests that the data are not internally consistent, and thus the benchmark is *not* satisfied.

Additional analysis: The analysis can also be conducted if sub-national level notification and sub-national population data are also available. When such data are observed at the sub-national level, 6 out of 9 regions showed an increase in the percentage change of smear-positive case notifications from 2010 to 2011 (Figure 4). Also of note is the large percentage change from 2008 to 2009 in the North east zone, which does not appear to impact the overall national percentage change. This suggests that such changes are due to the small number of cases from a small population being notified. Nonetheless, such

inconsistencies should be further investigated. Overall, the percentage change in TB case notification rate by years at the sub-national level shows much variability over time and by geographical area, and supports the conclusion that the data are not internally consistent.

Figure 4. Percentage change in TB case notification rate by TB reporting region, Uganda 2008–2012



Benchmark 7: Ratio of the number of people with presumptive TB to total notifications of TB cases

Data requirements: The number of people with presumptive TB and the number of new and relapsed TB cases (all forms) are required. Ideally, people with presumptive TB include all those who have been suspected of having TB, irrespective of whether or not a smear (or other laboratory test) was performed.

Conducting the analysis: For each year, the ratio of TB suspects to notifications can be calculated (Table 6).

Table 6. Ratio of the number of people with presumptive TB to the total number of notified TB cases, Uganda, 2008–2012

Year	Number of people with presumptive TB	Number of notified cases, all forms of TB	Ratio
2008	88 576	42 197	2.10
2009	103 572	41 703	2.48
2010	126 437	42 885	2.95
2011	147 583	46 306	3.19
2012	155 852	44 437	3.51

Interpretation: The ratio of people with presumptive TB to notifications in Uganda increased steadily between 2008 and 2012. As such, these trends do satisfy the benchmark requirement for consistency, and it suggests that the reach of the NTP is steadily expanding over this time. Despite an increase in the number of tests being performed with time, the number of notified cases has not increased at the same rate, and even declined in 2012 compared with 2011. It is unlikely that the number of notified cases has plateaued given their latest case detection rate of 69% (2012), but further investigations may find a change in laboratory testing methods, a change in case definitions or some other artefact of the surveillance system.

Summary

There were no available VR data, therefore benchmarks 2 to 7 were applied in this assessment. Data trends satisfied benchmarks 2, 3 and 7, but were not satisfied for benchmarks 4, 5 and 6. Therefore, based on this assessment, standard B1.7 for internal consistency was only *partially met*.

Standard B1.8:

All diagnosed cases of TB are reported

Benchmarks:

Both benchmarks should be satisfied to meet this standard:

- TB reporting is a legal requirement
- $\geq 90\%$ of TB cases are reported to national health authorities, as determined by a national-level investigation (e.g. inventory study, conducted in the past 10 years)

Rationale for standard and benchmarks

TB under-reporting is the proportion of detected TB cases not reported to national health authorities. If many people with TB are diagnosed but not reported, then the actual burden of TB is unknown, the number of patients that may be receiving sub-standard care is uncertain, and resources may not be best targeted to those most in need. Under-reporting may be a significant problem in settings where TB is endemic, reporting of cases is not mandatory, and people seek care for TB symptoms from a wide variety of health care providers. TB under-reporting should be minimized in order to effectively monitor the burden of TB and its trends through routine surveillance of TB cases.^{56,57}

TB case reporting should be legally mandated to ensure TB cases are captured by the TB surveillance system. However, even if a legal framework is in place, some under-reporting is inevitable and the level of TB under-reporting should be known by the NTP. Inventory studies can be used to obtain a direct measurement of under-reporting within TB surveillance.

Inventory studies aim at assessing the number of detected TB patients during a defined period of time by actively observing health providers' practice, and then computing the proportion of detected cases not reported to health authorities. The number of TB cases meeting standard case definitions in all or in a sample of public and private health facilities (including laboratories) is compared, through record linkage, with the records for cases reported to local and national authorities.

To date, there is limited evidence for the benchmark because few countries have conducted national level studies, and as such, the cut-off may evolve over time. However, inventory studies in countries with high performing TB surveillance systems, such as the UK and The Netherlands, have estimated under-reporting to be less than 10% after

56 Tuberculosis in the UK: Annual report on tuberculosis surveillance in the UK, 2010. London, Health Protection Agency Centre for Infections, 2010 (http://www.hpa.org.uk/webc/HPAwebFile/HPAweb_C/1287143594275, access 9 November 2013).

57 van Hest NA, Smit F, Baars HW, De Vries G, De Haas PE, Westenend PJ, Nagelkerke NJ, Richardus JH. Completeness of notification of tuberculosis in The Netherlands: how reliable is record-linkage and capture-recapture analysis? *Epidemiology and Infection*, 2007;135(6):1021–9.

detailed review of and adjustment for unmatched records and/or false-positive cases.

Method to assess benchmarks

Data sources and data collection methods: For the first benchmark, an assessment of legal and regulatory frameworks should be made through the study of health reports.

For the second benchmark, TB under-reporting should be directly measured through inventory studies. Results from recent (conducted 10-years ago or less) inventory studies should be used as the primary source of data. If no recent inventory studies have been done, one should be implemented. Guidelines for conducting inventory studies are available from WHO.⁵⁸

Main limitations: Even though a legal framework may be in place for reporting TB, in some settings, cases may still not be reported or the laws may not be adequately enforced.

To conduct inventory studies, adequate financial and human resources are needed for optimum results. Usually, more data sources and bigger sample sizes will increase accuracy, but will also increase the cost of data collection.⁵⁹ Results obtained from small-scale studies not based on random sampling of all geographical areas in the country may not provide representative results. Cases with no or limited access to diagnostic services will not be detected. If unreported TB cases are not captured in the data sources used in the study but, for example, are captured elsewhere, an underestimate of the level of under-reporting may result. If the most recent inventory study was <10 years ago and there have been major changes to the surveillance system since then, those study results may no longer reflect the level of under-reporting, and therefore a new study should be conducted.

Interpretation of results: Both benchmarks should be satisfied for this standard to be *met*; which suggests most detected TB cases are reported. If TB reporting is a legal requirement for all providers, the first benchmark is satisfied. If at least 90% of cases were found to have been reported from a recently conducted inventory study (based on a best estimate, after any adjustments for unmatched records and/or false-positive cases), then the second benchmark is satisfied. If only one benchmark is satisfied then the standard is *partially met*. If neither is satisfied, then the standard is *not met*.

58 Assessing tuberculosis under-reporting through inventory studies. Geneva, World health Organization, 2012 (http://www.who.int/tb/publications/inventory_studies/en/, accessed 9 November 2013).

59 The costs of recent studies in Iraq and Yemen that covered approximately half of these countries and lasted six months were in the range US\$ 120 000–300 000.

Recommended actions: If TB is not a reportable disease, NTPs should work with governing bodies to advocate for and implement mandatory reporting.⁶⁰ Requiring complete information for all TB patients from public or private health care providers helps governments ensure proper distribution of resources and TB control measures proportionate to the real burden of disease.

If an inventory study has not been conducted to directly assess under-reporting of TB cases within the past 10 years, begin planning a survey using recommended WHO guidelines. If an inventory study shows that a large proportion of diagnosed cases go unreported, urgent action is needed. A close examination of the reasons why under-reporting happens should be undertaken. Also, targeted strategies should be developed and implemented to strengthen national surveillance systems and control strategies.⁶¹ For example, public-private and public-public mix approaches, such as incentive-based and reimbursement schemes and mandatory reporting of cases, can help ensure provider practices are in line with national guidelines and international standards. Furthermore, data from inventory studies can be used to update and make corrections in surveillance estimates. If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Examples

United Kingdom

In the United Kingdom, notification of TB cases first became a legal requirement in 1913. All forms of TB are statutorily notifiable by the physician making or suspecting the diagnosis under the Public Health (Control of Disease) Act 1984.

Public Health England uses inventory methods to assess whether all cases of TB are reported to the national programme, based on matching with other data sources, such as the laboratory database, national HIV surveillance and bespoke surveys. These comparisons generally show between 5 to 17% under-reporting to the surveillance system.¹ In 2006, a detailed audit of the unmatched laboratory *Mycobacterium tuberculosis* isolates was conducted. This further in-depth review of the data showed that under-reporting was actually about 5%.¹ Lessons learned from inventory studies are applied to the national web-based case reporting system to improve data quality.

60 See Parts A10, A15.

61 See Part A18.

In the case of the UK, both benchmarks are satisfied and the standard is *met*.

Iraq

TB is a notifiable disease in Iraq.

An inventory study was implemented in Iraq in 2011 in which the level of TB under-reporting was assessed.⁶² Prospective longitudinal surveillance was implemented for all eligible public and private non-NTP providers in a random sample of eight of the eighteen Iraqi governorates for three months (May to July). Laboratory and TB register forms identical to those used by the NTP were introduced in non-NTP public and private facilities to record demographic, diagnostic, referral and treatment information and NTP verification. Record-linkage of data sources was then conducted. A total of 1985 TB cases were identified and of these, the NTP registered 1677 patients (observed completeness 84% i.e. 16% under-reporting). Findings show that TB surveillance needs to be strengthened to reduce levels of under-reporting in Iraq.

In the case of Iraq, the first benchmark was satisfied but the second benchmark was not, so this standard was *not met*.

62 Huseynova S, Hashim DS, Tbeni MR, Harris R, Bassili A, Abubakar I, Glaziou P, Floyd K, van Hest NA. Estimating tuberculosis burden and reporting in resource-limited countries: a capture-recapture study in Iraq. *International Journal of Tuberculosis and Lung Disease*, 2013; 17(4):462-7

Standard B1.9: Population has good access to health care

Benchmarks:

Both benchmarks should be satisfied to meet this standard:

- Under-five mortality rate (probability of dying by age 5 per 1000 live births) is <10
- <25% total health expenditure is out-of-pocket

Rationale for standard and benchmarks

If standards B1.1 to B1.7 are met, national data on TB notifications are of high quality. If standard B1.8 is also met, then national notification data are providing a direct measure of the number of TB cases that were diagnosed. For notification data to provide a direct measurement of TB incidence, however, a further condition must be met: the number of *undiagnosed* cases must be a small or negligible fraction of the total number of TB cases. Whether or not this condition is met is not directly within the influence of the NTP and its surveillance system itself, but it does directly influence the ability of a surveillance system to account for all cases of TB.

For all (or almost all) cases of TB to be diagnosed, health care services that include staff qualified to recognize TB signs and symptoms and TB diagnostic capability (laboratory capacity and radiography) must be widely available and barriers to using them (including geographic access, financial costs, and perceptions about quality and acceptability) must be low or non-existent. Broadly speaking, a country's population must have good access to a well-functioning health care system.

There are no universally agreed benchmarks that define whether or not a country's population has good access to a well-functioning health care system. However, examples of indicators that are used to assess health system coverage and performance and that can be reported for all or most countries include: the under-five mortality rate (per 1000 live births); life expectancy at birth; the maternal mortality ratio (per 100 000 live births); the number of doctors and nurses per 10 000 population; antenatal care coverage; and out-of-pocket expenditures on health as a percentage of all expenditures on health.⁶³

For the purposes of the TB surveillance checklist, two of these indicators have been chosen. The under-five mortality rate has been selected because it provides a very broad, overall indication of the quality and

63 World Health Statistics 2013. Geneva, World health Organization, (http://www.who.int/gho/publications/world_health_statistics/en/, accessed 9 November 2013).

coverage of health care. The percentage of health expenditures that are out-of-pocket, i.e. expenditures that are paid directly to health care practitioners and are not reimbursable, was also selected because it provides a broad indication of whether there are major financial barriers to accessing health care. The benchmarks – that the under-five mortality rate should be less than 10 per 1000 live births and that out-of-pocket expenditures should be less than 25% of total national expenditures on health – have been defined based on current values for these indicators in countries considered to have health systems of high coverage and quality. These cut-offs may evolve over time.

It is recognized that this standard and associated benchmarks are outside the purview of the NTP. However, to assess the surveillance system's capacity to accurately estimate TB burden, this standard and the associated benchmarks are deemed necessary

Method to assess benchmarks

Data sources and data collection methods: The overall assessment of this standard should be done at the national level. The latest country-specific estimates of the under-five mortality rate can be found in the WHO publication *World Health Statistics* (issued annually) and on the WHO Global Health Observatory website.⁶⁴ The latest country-specific estimates of the proportion of national health expenditures that are out-of-pocket can be found in WHO's national health accounts database and also in the WHO Global Health Expenditure Atlas.^{65,66}

Main limitations: The benchmarks provide an approximate indication of access to high-quality health care, with the rationale that this is necessary to ensure that people with TB signs and symptoms will seek care and be diagnosed. It is possible that there are countries that do not meet the benchmarks but in which virtually all cases of TB are reaching health care services and are being diagnosed. There is also evidence that cases of TB can be missed in countries where the population has good access to high-quality health care, although this is expected to be rare.

Interpretation of results: If both benchmarks are satisfied this standard is *met*. If standards B1.1 to B1.8 are also *met*, then a country's TB surveillance system can be considered to provide a direct measure of

64 Under-five mortality (section of the WHO Global Health Observatory website): www.who.int/gho/child_health/mortality/mortality_under_five/en.

65 Select the "Out of pocket expenditure as a % of THE" for the country and years of interest (units: % of total health expenditure). Available from: www.who.int/nha/database.

66 See the specific country profiles: <http://apps.who.int/nha/atlasfinal.pdf>.

TB cases occurring each year. If only one of the benchmarks is satisfied then the standard will be classified as *partially met*. If neither benchmark is satisfied, the standard is *not met*. In such a situation, it is likely that there are people with TB who are not being diagnosed with the disease and that TB notification data are thus not a good proxy for TB incidence.

Recommended actions: The actions needed if one or both benchmarks are not met are well beyond the influence of a TB surveillance system and are likely long-term in nature. The first requires strengthening health care services, for example through increased financial investments, along with recruitment and retention of well-trained staff and their deployment throughout the country. Part of which should be a long-term financial plan to support TB surveillance activities.⁶⁷ Out-of-pocket expenditures can be lowered by introducing mandatory national health insurance in which required levels of co-payments are low (and where there is a ceiling on the absolute amount of expenditure that is out-of-pocket) or by introducing a tax-funded health care system that is free at the point of access. If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Examples

Brazil

In 2010, the under-five mortality rate was 19 per 1000 live births and out-of-pocket health expenditures were 31% of total health expenditures. Brazil did not satisfy either of the two benchmarks and thus the standard was *not met*.

China

In 2010, the under-five mortality rate was 18 per 1000 live births and out-of-pocket health expenditures were 37% of total health expenditures. China did not satisfy either of the two benchmarks and thus the standard was *not met*.

India

In 2010, the under-five mortality rate was 63 per 1000 live births and out-

⁶⁷ See Part A17.

of-pocket health expenditures were 61% of total health expenditures. India did not satisfy either of the two benchmarks and thus the standard was *not met*.

Kenya

In 2010, the under-five mortality rate was 85 per 1000 live births and out-of-pocket health expenditures were 43% of total health expenditures. Kenya did not satisfy either of the two benchmarks and thus the standard was *not met*.

The Netherlands

In 2010, the under-five mortality rate was 4 per 1000 live births and out-of-pocket health expenditures were 5% of total health expenditures. The Netherlands satisfied both benchmarks and thus the standard was *met*.

South Africa

In 2010, the under-five mortality rate was 57 per 1000 live births and out-of-pocket health expenditures were 17% of total health expenditures. South Africa satisfied the second benchmark but is a considerable distance from satisfying the first benchmark, and thus the standard was *not met*.

Thailand

In 2010, the under-five mortality rate was 13 per 1000 live births and out-of-pocket health expenditures were 14% of total health expenditures. Thailand satisfied the second benchmark but was just short of the first. It was very close, but only *partially met* the standard.

United Kingdom

In 2010, the under-five mortality rate was 5 per 1000 live births and out-of-pocket health expenditures were 10% of total health expenditures. The UK satisfied both benchmarks and *met* the standard.

United States of America

In 2010, the under-five mortality rate was 8 per 1000 live births and out-of-pocket health expenditures were 12% of total health expenditures. The USA satisfied both benchmarks and *met* the standard.

Standard B1.10:

Vital registration system has high national coverage and quality

Benchmarks:

Both benchmarks should be satisfied to meet this standard:

- Cause of death documented in $\geq 90\%$ of total deaths recorded in: a) national vital registration system or b) sample vital registration system
- $< 10\%$ of deaths have ICD codes for ill-defined causes (defined as ICD-9 780-799 and ICD-10 R00-R99)

Rationale for standard and benchmarks

There is a wide variation in the coverage and quality of cause of death data across countries. High coverage and accurate information on TB deaths is essential for monitoring the magnitude, distribution and trends of the burden of TB, as well as for informing policy decisions and distribution of resources. Mortality attributable to TB can be measured directly where a good vital registration system, with high coverage and accurate coding of cause-of-death, exists. The number of deaths among patients on TB treatment is not an accurate measure of TB mortality because it includes all causes of death and does not include deaths from TB among people not registered on treatment. If there is low coverage of registration of TB deaths, then an under-reporting of mortality can be the result. If the proportion of ill-defined causes of death is high, then estimates of the distribution of TB deaths are incorrect, hindering effective resource allocation.

Coverage is defined as the number of deaths reported in a given year as a percentage of estimated deaths in the resident population for a country. Quality of data within a vital registration system can be defined as the proportion of deaths assigned to ICD codes for symptoms, signs and ill-defined conditions (ICD-9 780-799 and ICD-10 R00-R99). The benchmarks of $\geq 90\%$ coverage and $< 10\%$ poor quality data (i.e. deaths with ICD codes for ill-defined causes) are based on WHO quality criteria for vital registration data in the 2012 *Global Health Report* and a review of WHO data of 122 country-year data points.⁶⁸

68 World Health Statistics 2012. Geneva, World Health Organization, 2012 (http://www.who.int/gho/publications/world_health_statistics/EN_WHS2012_Part2.pdf, accessed 9 November 2013).

It is recognized that this standard and benchmarks are outside the purview of the NTP. However, to assess the capacity of the surveillance system to accurately estimate TB burden, this standard and associated benchmarks are deemed necessary.

Method to assess benchmark

Data sources and methods of data collection: Information about vital statistics and the vital registration system are often available through the national statistics office. The quality of data and coverage of the system are usually reported through routine annual reports or periodic surveys e.g. every 10 years. Information can also be found at the WHO Mortality Database;⁶⁹ the data at this website comprise deaths registered in national vital registration systems, with underlying cause of death as coded by the national authority.

Main limitations: The vital registration system and thus the performance of this standard are often outside of the direct control of the NTP. Many countries have no or sub-optimal vital registration systems.

Interpretation of results: Low coverage indicates not all deaths are registered. The proportion of ill-defined to total deaths gives an indication of the quality of certification and the application of ICD rules in the selection of the underlying cause of death. Both benchmarks should be satisfied to meet this standard. If only one benchmark is satisfied, then the standard is *partially met*. If neither is satisfied, then the standard is *not met*.

Recommended actions: While vital registration systems are often outside the NTP structure, they are within the programmes' sphere of influence. TB programme directors should use the results of the assessment as hard evidence to show changes are needed in another sector in order to accurately measure TB burden. If a vital registration system with a high level of accuracy and coverage of deaths does not exist, then sample vital registration systems or mortality surveys, and demographic surveillance systems using verbal autopsy to determine cause of death are a potential interim source of improved estimates of mortality attributable to TB (financial plans to support TB surveillance activities should possibly consider investments in such activities).⁷⁰ Though they may only partially cover deaths in a country, they can be an important intermediate solution to obtain mortality and cause of death information. To improve quality, assessments of vital registration

69 Mortality and global health estimates (section of the WHO Global Health Observatory website): (<http://apps.who.int/gho/data/node.main.686?lang=en>, accessed 9 November 2013).

70 See Part A17.

systems and analyses of their data are needed. Further improvements can happen by increasing the percentage of deaths that are certified by attending and non-attending physicians, training personnel to ensure the medical condition leading to death is accurately identified and recorded, and emphasising the importance of accurate reporting on death certificates. If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Example

Thailand

In a study in Thailand,⁷¹ investigators used the Health Metrics Network tool⁷² in combination with a systems analysis and literature review to evaluate the collection and flow of mortality and cause-of-death data to ascertain and improve weaknesses in their vital registration system. Based on the *Survey of Population Change*,⁷³ a nationally representative household survey conducted every decade, coverage of the vital registration of death in Thailand was determined to be 95%. But the proportion of ill-defined deaths was approximately 40%. Based on this, the standard would have been *partially met*. Recommendations for improvement included training of physicians and data coders, harmonization of death certificates and registries, and increasing public awareness of the importance of registering all deaths.

71 Tangcharoensathien V, Faramnuayphol P, Teokul W, Bundhamcharoen K, Wibulpholprasert S. A critical assessment of mortality statistics in Thailand: Potential for improvements. *Bulletin of the World Health Organization*, 2006; 84(3):233–8.

72 Health Metrics Network Framework. Geneva, Health Metrics Network, 2008 (<http://www.who.int/healthmetrics/tools/en/>, accessed 9 November 2013).

73 Population Change Survey. Bangkok, National Statistics Office, 1996 (<http://web.nso.go.th/survey/popchan/popchan.htm>, accessed 9 November 2013).

Standard B2.1:

Surveillance data provide a direct measure of drug-resistant TB in new cases

Benchmarks:

One of the two benchmarks should be satisfied to meet this standard:

- Rifampicin susceptibility status (Positive/Negative) documented for $\geq 75\%$ of new pulmonary TB cases
- Rifampicin susceptibility status (Positive/Negative) documented for a nationally representative drug resistance survey of new pulmonary TB cases

Rationale for standard and benchmarks

Surveillance of drug-resistant TB is critical to describe TB epidemiology, analyse trends over time, assess performance of NTPs, forecast the need for patient treatments, design standardized regimens for the treatment of drug-resistant TB, and promptly identify and respond to outbreaks. Failure to monitor drug-resistant TB contributes to poor treatment outcomes for individual patients and transmission of drug-resistant strains in the community. The group of previously untreated TB cases is the ideal target of surveillance of drug resistance. By definition, previously untreated TB patients have not taken anti-TB medication for more than one month, therefore the identification of drug-resistant TB in this group of patients indicates transmission of drug resistance in the community. In accordance with the WHO *Guidelines for surveillance of drug resistance in tuberculosis*⁷⁴ to assess drug-resistant TB in previously untreated TB cases, two approaches for data collection can be used: a) surveillance systems based on routine diagnostic testing of previously untreated (new) TB cases; or, b) special surveys of a representative sample of previously untreated (new) TB cases. Surveillance systems are the preferred approach to monitor drug resistance, but special surveys represent a valid alternative in settings where laboratory capacities for culture and DST are limited.

When surveillance systems are used, to assess the representativeness of drug-resistant TB results it is critical that a high proportion of new

74 Guidelines for surveillance of drug resistance in tuberculosis. Fourth edition. Geneva, World Health Organization, 2009 (http://whqlibdoc.who.int/publications/2009/9789241598675_eng.pdf, accessed 9 November 2013).

pulmonary TB cases ($\geq 75\%$) have DST results documented for at least rifampicin.⁷⁵ Evidence from surveillance systems in North American and European countries (Western and Eastern) has been used to set the threshold of $\geq 75\%$.

A WHO-endorsed laboratory method for DST should be used for both surveillance systems and special surveys. Furthermore, the quality of DST results should be assured by a Supranational TB Reference Laboratory (SRL).⁷⁶

Method to assess benchmarks

Data sources and data collection methods: This standard can be assessed at the national level through two different methods of data collection:

- a) *Surveillance systems based on routine diagnostic testing of previously untreated TB cases.* If an official national annual TB report from the previous year for which data are complete is available for review and contains information to assess the first benchmark (% of new pulmonary cases with rifampicin susceptibility status documented), the report itself is sufficient to assess the standard. In the absence of such a report, completeness of data can be examined in a case-based electronic surveillance system or a national laboratory register used to capture these variables.
- b) *Special surveys of a representative sample of previously untreated TB cases.* The report of the most recent survey should be reviewed, and compliance of the survey with the WHO *Guidelines for surveillance of drug resistance in tuberculosis* should be assessed. The survey should have been performed in the past five years.

The quality of the DST results is assessed by reviewing the results of the proficiency testing conducted with the SRL. Concordance of these results should be $\geq 95\%$ for rifampicin. However, it should be noted that this standard seeks to assess the quality of data but not that of the underlying testing methods.

Main limitations: This standard will miss providers that do not report to the NTP and/or are outside the reporting network, e.g. private health care providers. If the standard is met by utilizing special surveys of a representative sample of previously untreated TB cases, then an accurate time trend analysis of frequencies of drug resistance could be

75 Documented as either a 'positive' or 'negative' result if a molecular test is used and a 'resistant' or 'susceptible' result if conventional DST is used. Note that the 1st benchmark requires a result for at least $\geq 75\%$ of cases; indeterminate results are not included.

76 More information is available at: <http://www.stoptb.org/wg/gli/srln.asp>

problematic due to the very large sample size required. The quality of the laboratory may influence the outcome, but this can be minimized as long as the laboratories meet the assured standard.

Interpretation of results: In order to meet this standard either the first, second or both benchmarks should be satisfied. If rifampicin susceptibility status is documented for $\geq 50\%$ but $< 75\%$ of new pulmonary TB cases, and/or a nationally representative drug resistance survey is currently underway then this standard is *partially met*. If neither of the two benchmarks is satisfied, then the standard is *not met*.

Recommended actions: Countries with no surveillance data on drug resistance should plan to conduct a survey of a nationally representative sample of previously untreated TB cases. Such surveys should be repeated every five years to monitor time trends in drug resistance. Countries with experience in conducting drug resistance surveys should move towards the establishment of a surveillance system based on routine diagnostic testing of TB cases, which represents the best approach to monitor drug resistance. Routine drug resistance testing should then be part of the NTP strategy and integrated into their long-term financial plans. This will require strengthening of the country's capacity to perform culture and DST, according to WHO recommendations.⁷⁷ If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Examples

United Kingdom

In Latvia, an electronic information system for TB surveillance is used. To assess whether at least the minimal data elements were captured for previously untreated TB cases, the TB surveillance information system was investigated. Of the 703 previously untreated pulmonary TB cases reported in 2011, 579 had a positive culture and among them, 562 had available DST results for rifampicin (80%, 562/703). Concordance for rifampicin in the most recent round of proficiency testing conducted with the SRL was $\geq 95\%$. The standard can be considered to have been met in Latvia because the first benchmark was satisfied.

Bangladesh

In Bangladesh, a surveillance system based on routine diagnostic testing

77 Guidelines for surveillance of drug resistance in tuberculosis. Fourth edition. Geneva, World Health Organization, 2009 (http://whqlibdoc.who.int/publications/2009/9789241598675_eng.pdf, accessed 9 November 2013).

78 Guidelines for surveillance of drug resistance in tuberculosis. Fourth edition. Geneva, World Health Organization, 2009 (http://whqlibdoc.who.int/publications/2009/9789241598675_eng.pdf, accessed 9 November 2013).

of previously untreated TB cases has not yet been established, but a survey of a nationally representative sample of previously untreated TB cases was completed in 2011 according to WHO recommendations.⁷⁸ Study sites across the country were selected using a probability proportional to size approach. All consecutive new pulmonary TB cases presenting to the 39 selected health centres were enrolled in the study. Culture and drug sensitivity results were available for 1050 of the 1080 enrolled patients. The standard B2.1 was therefore considered to have been *met* because the second benchmark had been satisfied.

Standard B2.2:

Surveillance data provide a direct measure of the prevalence of HIV infection in TB cases

Benchmarks:

One of the two benchmarks should be satisfied to meet this standard:

- HIV status (Positive/Negative) is documented for $\geq 80\%$ of all notified TB cases
- HIV status is available from a representative sample from all TB cases notified in settings with a low-level epidemic state or where it is not feasible to implement routine surveillance

Rationale for standard and benchmarks

Surveillance of HIV in TB patients is critical for individual patient management, proper planning of integrated TB and HIV diagnostic and treatment services, and for understanding the TB epidemic (i.e. TB incidence, prevalence and mortality) regardless of the HIV epidemic state. WHO recommends that all patients with presumptive or diagnosed TB should receive HIV testing and counselling to ensure early case detection and rapid initiation of treatment.⁷⁹ In line with the WHO policy on collaborative TB/HIV activities, data on HIV status among TB cases should be collected through routine surveillance in all settings regardless of the HIV epidemic state (i.e. generalized,⁸⁰ concentrated⁸¹ or low-level⁸² epidemic states). However, in settings with a low-level epidemic state, where it is not feasible to implement routine HIV testing for the vast majority of TB cases, data could be collected through periodic surveys of a sample of TB cases. When routine surveillance systems are used to assess the representativeness of HIV testing results, it is crucial that a high proportion of all notified TB cases ($\geq 80\%$) has had their HIV status documented. Routine HIV testing and counselling provides critically important benefits to people living with HIV, including better access to testing, early case detection and rapid initiation of treatment.

Method to assess benchmarks

Data sources and data collection methods: This standard can be

79 WHO policy on collaborative TB/HIV activities (WHO/HTM/TB/2012.1). Geneva, World Health Organization, 2012 (http://whqlibdoc.who.int/publications/2012/9789241503006_eng.pdf, accessed 9 November 2013).

80 Generalized epidemic state: HIV prevalence consistently $> 1\%$ among pregnant women.

81 Concentrated epidemic state: HIV prevalence is consistently $> 5\%$ in at least one defined sub-population, and is $< 1\%$ among pregnant women in urban areas.

82 Low-level epidemic state: HIV prevalence has not consistently exceeded 5% in any defined sub-population.

assessed at the national level through two different methods of data collection:

1. *Surveillance systems based on routine HIV testing* of TB cases in all settings. If an official annual TB report from the previous year for which data are complete is available for review and demonstrates the proportion of TB patients tested for HIV, the report itself is sufficient. Alternatively, if data are available in a standardized electronic system, data from the most complete calendar year can be used for the analysis. In countries with separate systems for TB and HIV surveillance, where the HIV status cannot be routinely documented in the TB surveillance system for confidentiality or other reasons, record linkage between the TB and HIV surveillance systems should be performed to assess whether the benchmark has been met.⁸³
2. *In countries with a low-level HIV epidemic state or where it is not feasible to implement routine surveillance for the vast majority of TB cases, periodic (special) or sentinel surveys are recommended every 2–3 years.* Surveys should be conducted in accordance with WHO policy on collaborative TB/HIV activities.⁸⁴

Main limitations: Some countries have separate systems for TB and HIV surveillance. In such circumstances, record linkage between the two systems should be performed. This operation can be complex and time-consuming, especially in settings where unique identifiers for patients are not routinely used. Record linkage may not be possible for legal reasons, therefore undertaking periodic surveys may be a suitable alternative.

Interpretation of results: The standard is met if 80% or more TB cases have a documented HIV status. If the first benchmark is satisfied, and if standard B1.8 is also met, then a country's TB surveillance system can be considered to provide a direct measure of the prevalence of HIV infection among TB cases. In settings with a low-level HIV epidemic state, where it is not feasible to implement routine HIV testing for the majority of TB cases, if HIV status is available from a representative sample of all notified TB cases using periodic surveys conducted according to the recommendations of WHO, the standard is *met*. If HIV status is documented for $\geq 50\%$ but $< 80\%$ of all notified TB cases, or a periodic HIV survey of TB cases is currently underway then this standard is *partially met*. If neither of the two benchmarks is satisfied, then the standard is considered *not met*.

83 Assessing tuberculosis under-reporting through inventory studies. Geneva, World Health Organization, 2012 (http://www.who.int/tb/publications/inventory_studies/en/, accessed 9 November 2013).

84 WHO policy on collaborative TB/HIV activities (WHO/HTM/TB/2012.1). Geneva, World Health Organization, 2012 (http://whqlibdoc.who.int/publications/2012/9789241503006_eng.pdf, accessed 9 November 2013).

Recommended actions:

- Surveillance of HIV among TB patients and surveillance of active TB disease among people living with HIV should be conducted in all countries irrespective of national adult HIV and TB prevalence rates. This should be done using standardized reporting and recording formats, to understand the trends of the epidemic and develop sound strategies to address the dual TB/HIV epidemic. Strategies should also be documented as part of the national TB guidelines.
- HIV testing and counselling of all patients with presumptive or diagnosed TB should form the basis of routine surveillance in all settings. To better ensure completeness of HIV data among TB cases, other patient monitoring systems – such as HIV care/antiretroviral therapy (ART), maternal and child health (MCH)/prevention of mother-to-child transmission (PMTCT), and TB/HIV – can be used to cross-check and reconcile data between HIV and TB programmes at local and country levels.
- In settings with a low-level HIV epidemic state where routine HIV surveillance is not feasible and HIV status is not available from a representative sample of notified TB cases, a survey to elucidate HIV status among TB patients is recommended at least every 2–3 years.
- If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Examples

Kenya

Since 2008, the Ministry of Health in Kenya – a country with a generalized HIV epidemic – has managed to record the known HIV status of over 80% of TB patients. In 2011, 103 981 TB cases (all forms) were reported in Kenya. Among them, 97 136 (93%) had their HIV status recorded in the TB register. Standard B2.2 has been *met* because the first benchmark is satisfied. To achieve such high coverage of HIV surveillance among TB patients, Kenya successfully scaled up collaborative TB/HIV activities, especially the integration of HIV testing and ART in TB facilities. This comprehensive strategy also contributed to the improved uptake of ART for HIV-infected TB patients in the country.

South Africa

In South Africa, a country with a generalized HIV epidemic state, a total of 389 974 TB cases (all forms) were reported in 2011. Among them 323 440 have a known HIV status recoded in the TB register (83%; 323 440/389 974). Standard B2.2 has been *met* because the first benchmark is satisfied.

Standard B2.3:

Surveillance data for children reported with TB are reliable and accurate, and all diagnosed childhood TB cases are reported

Benchmarks:

Both benchmarks should be satisfied to meet this standard:

- Ratio of age groups 0–4 to 5–14 years is in the range 1.5–3.0
- ≥90% of childhood TB cases are reported to national health authorities, as determined by a national-level investigation (e.g. inventory study, conducted in the past 10 years)

Rationale for standard and benchmarks

Registration of TB cases is completed on standardized recording and reporting forms, involving the collection of patient demographic, clinical, diagnosis, and treatment outcome data. Improving the diagnosis of TB continues to be the greatest challenge in children,⁸⁵ where obtaining a sputum specimen is often not even attempted. Diagnosis depends on the level of health care services, and the availability of standard TB diagnostic tests and human resources. It is based on a composite set of clinical criteria and diagnostic test results ranging between the key features suggestive of childhood TB in a WHO-recommended algorithm⁸⁶ for the straightforward pulmonary cases, to the more complicated tertiary care diagnosis of different disease manifestations, which are seen more often in children than adults. Furthermore, in high TB endemic settings, the majority of childhood TB cases will occur in young children who tend to have paucibacillary disease. As the emphasis of NTP reporting has traditionally been on sputum smear-positive disease, which is very uncommon in children, many childhood TB cases are not diagnosed and recorded. Such diagnostic challenges, combined with the lack of routine recording and reporting of childhood TB, make it difficult to know the true burden of TB disease in children, which is most likely underestimated from the TB surveillance data.

Given the challenges of diagnosing and reporting childhood TB cases, the level of TB under-reporting of childhood TB (as well as all other cases) should be known by NTPs. Inventory studies can be used to obtain a direct measurement of under-reporting within TB surveillance (please refer to standard 1.8).

⁸⁵ Defined as between the ages 0–14 years.

⁸⁶ Guidance for national tuberculosis programmes on the management of tuberculosis in children. Geneva, World health Organization, 2006 (http://whqlibdoc.who.int/hq/2006/WHO_HTM_TB_2006.371_eng.pdf, accessed 9 November 2013).

Method to assess benchmarks

Data sources and data collection methods: For the first benchmark, data from national routine surveillance systems disaggregated by case type, age and sex should be examined and routinely checked annually.

For the second benchmark, TB under-reporting should be directly measured through inventory studies. Results from recent (conducted 10-years ago or less) inventory studies should be used as the primary source of data. If no recent inventory studies have been done, one should be implemented. Guidelines for conducting inventory studies are available from WHO.⁸⁷

Main limitations: The ranges of values used for the benchmarks have been based on data from a limited number of countries (please also refer to standard 1.8).

Interpretation of results: Accurate benchmark values within the suggested ranges compared with what is known about childhood TB epidemiology particularly within the country over time, provide some reassurance of a well-functioning surveillance system.^{88,89} These benchmarks are linked to standard 1.4 and 1.6, whereby the under-diagnosis of childhood TB in health centres and hospitals, and the under-reporting of cases to the NTP will similarly affect these benchmarks. If at least 90% of childhood TB cases were found to have been reported from a recently conducted inventory study (based on a best estimate, after any adjustments for unmatched records and/or false-positive cases), then the second benchmark is satisfied. Meeting both benchmarks is required for this standard to be *met*. If only one of the benchmarks is satisfied, then the standard is only *partially met*. If none of the benchmarks are satisfied, the standard is *not met*.

Recommended actions: If the first benchmark is not satisfied, it should be investigated further and potential reasons for these discrepancies should be hypothesized and discussed with the health workers who make and report the diagnosis of childhood TB, and the TB officers who record the childhood TB cases. As such, corrective actions (such as M&E training, and the strengthening of such training, in terms of frequency of visits and quality of supervision) may be required.

If an inventory study has not been conducted within the past 10 years to directly assess under-reporting of cases, begin planning a survey using recommended WHO guidelines. If an inventory study shows that a large proportion of diagnosed cases go unreported, urgent action is needed. A close examination of the reasons why under-reporting happens should be undertaken.

87 Assessing tuberculosis under-reporting through inventory studies. Geneva, World Health Organization, 2012 (http://www.who.int/tb/publications/inventory_studies/en/, accessed 9 November 2013).

88 BJ Marais et al. The clinical epidemiology of childhood pulmonary tuberculosis: a critical review of literature from the pre-chemotherapy era. *International Journal of Tuberculosis and Lung Disease*, 2004; 8(3):278–285.

89 AD Harries et al. Childhood tuberculosis in Malawi: nationwide case-finding and treatment outcomes. *International Journal of Tuberculosis and Lung Disease*, 2002; 6(5):424–431.

If required, key actions that will address benchmark gaps that limit achievement of the standard should be described. An estimated budget to support activities that could bridge this gap(s) would assist in developing an investment plan.

Example

A review of childhood TB data in a high-burden country identified incidence of all forms of TB disease to be more than 200/100 000 with 30% of the population being children (0–14 years). Available data report that childhood TB accounts for 1.7% of all TB cases recorded, with 30% of childhood TB cases aged between 0 and 4 years, and 70% aged between 5 and 14 years. As such, the ratio of the 0–4 age group to the 5–14 age group was 0.4, which is not within the desired range, therefore this benchmark was not *satisfied*. Since a national level investigation (e.g. inventory study) has not been undertaken, the second benchmark was also *not satisfied*. As a result, the benchmark was *not met*. The data provided above also support this decision as they strongly suggest a marked under-representation of childhood TB cases, especially those within the 0–4 year age group – the age group for which TB diagnoses are most challenging.



GLOSSARY OF TERMS

4 Glossary of terms

Age of patient: Patient age in years, or according to age group.

All forms of TB: Pulmonary (smear-positive and smear-negative) and extrapulmonary TB.

Anatomical site of disease: *Extrapulmonary TB, pulmonary TB or both.*

Bacteriological results: Bacteriology refers to the smear status of pulmonary cases and the identification of *M. tuberculosis* for any case by culture or newer methods.

Basic management unit: A BMU is defined in terms of management, supervision, and monitoring responsibility. A BMU for TB control may have several treatment facilities, one or more laboratories, and one or more hospitals. The defining aspect is the presence of a manager or coordinator who oversees TB control activities for the unit and who maintains a master register of all TB patients being treated, which is used to monitor the programme and report on indicators to higher levels. Typically, the units correspond to the government's second subnational administrative division, which might be called, for example, a "district" or "county". It is internationally recommended that a BMU cover a population between 50 000 and 150 000 or up to 300 000 for large cities.

Basic management unit TB register: The BMU TB register (also sometimes called the district TB register) is intended primarily for recording the data needed to monitor BMU performance, using indicators and reports about TB patients. It is also commonly used to summarize testing results and treatment decisions in order to determine whether basic diagnostic and treatment guidelines are correctly implemented. No information that is beyond this monitoring scope should be included in the register. The register should contain the records of all patients diagnosed with TB and eligible for TB treatment, including those diagnosed with rifampicin-resistant TB or multidrug-resistant TB (MDR-TB), regardless of whether treatment was actually started. All of these cases are notifiable and should be included in the summary case notification reports sent to higher levels. The registration date is the date a patient is diagnosed with TB and is eligible for treatment.

Case-based records: This is where a case of TB and associated care and treatment information is recorded. Individual cases or episodes

of disease are recorded in case-based records. There will typically be a case identifier e.g. TB treatment number, held in the system to link various items of information related to the TB case, and this identifier may be confined to a limited part of the system (such as a treatment facility).

Clinically diagnosed TB case: A case in which a health worker (clinician or other medical practitioner) has diagnosed TB and has decided to treat the patient with a full course of anti-TB treatment, not necessarily meeting the definite TB case criteria.

Culture (or equivalent): Culture or WHO-endorsed molecular diagnostic test for TB e.g. Line Probe Assays, GeneXpert MTB/RIF.

Completeness of data: Data quality attribute to assess if data within a given field are comprehensive; proportion of cases with complete data on key variables.

Completeness of case reporting: Data quality attribute to assess completeness of registration and if TB cases are reported to national authorities via the surveillance system.

d: “Decision rule” in the LQAS literature (appendix 2); or number of TB deaths (B1.7).

d_a: Adjusted number of TB deaths.

Definite TB case: A patient with M. tuberculosis complex identified from a clinical specimen, either by culture or by a molecular method. In countries that lack the laboratory capacity to routinely identify M. tuberculosis, a pulmonary case with one or more initial sputum smear examinations positive for acid-fast bacilli (AFB) is also considered to be a “definite” case, provided that there is a functional external quality assurance (EQA) system with blind re-checking.

Duplicate records: More than one record in the national database refer to the same TB episode or case.

Extrapulmonary case of TB: A case of TB involving organs other than the lungs, e.g. pleura, lymph nodes, abdomen, genitourinary tract, skin, joints and bones, meninges.

g: Denotes the proportion of ill-defined causes.

Health expenditures that are out-of-pocket: Expenditures that are paid directly to health care practitioners and are not reimbursable; any direct outlay by households, including gratuities and in-kind payments,

to health practitioners and suppliers of pharmaceuticals, therapeutic appliances, and other goods and services whose primary intent is to contribute to the restoration or enhancement of the health status of individuals or population groups.

Health facility: A health facility is any health institution with health care providers formally engaged in any of the following TB programme functions (DOTS): referring patients with presumptive TB or confirmed TB cases, laboratory diagnosis, TB treatment and patient support during treatment.

Health Metrics Network tool: Provides the guidance on how to assess a country health information system.⁹⁰ Table IIIB – Assessing National HIS Data Sources: Civil Registration, provides 13 items within vital registration systems to be assessed, and the necessary criteria to be considered ‘highly adequate’ for each of these.

History of previous treatment: At the time of registration, each patient meeting the case definition is also classified according to whether or not he or she has previously received TB treatment or not, and if so, the outcome (if known).

Ill-defined causes ICD codes for symptoms, signs and ill-defined conditions (ICD-9 780-799 and ICD-10 R00-R99), also known as ‘garbage codes’.

Implausible value: An unrealistic value for a fact is recorded in a database record.

Inaccurate value: A recorded fact doesn’t reflect reality, e.g. *date of birth* recorded as 01/01/1956 but patient’s actual date of birth is 01/01/1965; *name* recorded as ‘John Smyth’ but the patient’s actual name is ‘Jonathan Smith’.

Inconsistent or misclassified values: Here the combination of facts recorded about a case are contradictory and indicate that at least one of the facts cannot be true.

Invalid or unclear value: This is a type of inaccuracy. It often occurs when data are recorded in free text fields, for example if sex is recorded as ‘under 15’, which makes no sense.

Inventory study for TB: Study conducted with the aim at assessing the number of detected TB patients during a defined period of time by actively observing health providers’ practice, and then computing the proportion of detected cases not reported to health authorities.

90 Assessing the national health information system: an assessment tool. Geneva, Health Metrics Network, 2008 (http://www.who.int/healthmetrics/tools/Version_4.00_Assessment_Tool3.pdf, accessed 9 November 2013).

Laboratory-confirmed TB case: A TB case diagnosed by smear, culture or other WHO-endorsed molecular test e.g. GeneXpert MTB/RIF.

Lot quality assurance sampling (LQAS): Sampling technique used to assess performance by determining whether a group selected units e.g. individuals, clinics, forms have achieved a given *pre-specified* performance standard, enabling it to be classified as acceptable or unacceptable.

Minimum set of variables: Age or age group, sex, year of registration, bacteriological results, history of previous treatment, anatomical site of disease, and a patient identifier (for case-based systems).

Missing value: A fact about a TB case that has not been recorded, for example a patient's date of birth or age.

Mortality due to TB: Estimated number of deaths attributable to TB in a given time period, expressed per 100 000 population per year, including deaths from all forms of TB.

Mortality rate: Ratio of the number of people dying in a year to the total mid-year population in which the deaths occurred i.e. crude death rate.

Multidrug-resistant tuberculosis (MDR-TB): Tuberculosis caused by strains of *M. tuberculosis* that are resistant to at least isoniazid and rifampicin.

New case of TB: A patient who has never had treatment for TB or who has taken anti-TB drugs for less than one month. New patients may have positive or negative bacteriology and may have disease at any anatomical site.

Notified TB case: Case of TB is reported to national authorities.

P: In LQAS, this denotes the pre-specified "acceptable" level or threshold of the allowable number of sampled "lots" that are unacceptable in terms of each of the outcomes of interest.

Patient-based records: This is where a *patient* (an individual person rather than a TB case) is the basic unit of recording. Cases or episodes of disease are recorded, linked to separate records for each individual based on a unique personal identifier. If there is a reliable, nationally-unique identifier for all or at least an appreciable proportion of TB patients when they present for care, such as an ID card number or social security number, then records can truly be patient-based.

Patient identifier, e.g. numerical ID: A code or a value or a combination of values associated with an individual that is guaranteed to be unique to that individual. For example, an ID card number, social security number or TB registration number. A combination of more than one identifier can be used with or without the identity card number, according to the type of setting (presence of nominal system, availability of a coding system for health facilities, and extent of population coverage for identity card or social security number).

Presumptive TB This refers to a patient who presents with symptoms or signs suggestive of TB (previously known as a TB suspect).

Previously treated case of TB: A patient who has received 1 month of more of anti-TB drugs in the past, may have positive or negative bacteriology and may have disease at any anatomical site.

Probabilistic matching: A method used to estimate the probability that two TB cases are a true match. Two records are compared based on a specified number of fields. Each field is assigned a weight and potential matches are identified by calculating the sum of the strength of agreement or disagreement between fields.

Pulmonary case of TB: A case of TB involving the lung parenchyma. A patient with both pulmonary and extrapulmonary TB should be classified as a case of pulmonary TB.

Sub-national level: The administrative level below the national level e.g. states, provinces.

Tuberculosis case: A definite case of TB (see above) or one in which a health worker (clinician or other medical practitioner) has diagnosed TB and has decided to treat the patient with a full course of anti-TB treatment. *Note.* Any person given treatment for TB should be recorded as a case.

TB under-reporting: Proportion of detected TB cases not reported to national health authorities.

Under-five mortality rate: The probability of dying before the age of 5 years (per 1000 live births).

Verification period: A pre-specified period of time e.g. one quarter from the past calendar year, used for data quality assessment.

Vital registration coverage: Number of deaths reported in a given year as a percentage of estimated deaths in the resident population for a country i.e. the number of deaths with a documented cause divided by the total number of estimated deaths, denoted by 'c'.

Vital registration quality of data: Proportion of deaths assigned to ICD codes for symptoms, signs and ill-defined conditions (ICD-9 780-799 and ICD-10 R00-R99).

Year of registration: Year of registration of TB patient by provider.



APPENDICES

Appendix 1: Data verification tools for standard 1.4

1.1: Quarterly report verification tool – example

Number	Question	Result	Skip
TB QUARTERLY REPORT: TB BMU REPORTING VERIFICATION			
BMU IDENTIFICATION⁹¹			
N01	Name of BMU		
N02	BMU code	BMU CODE <input type="text"/> <input type="text"/> <input type="text"/>	
REVIEW THE TB QUARTERLY REPORT FOR THIS BMU COVERING MONTH X TO MONTH X 201X AND ANSWER THE FOLLOWING QUESTIONS.			
N03	What is the total number of TB cases (all types)? [See Box B1 of appendix 1.2 = B1]	A = <input type="text"/> <input type="text"/> <input type="text"/>	
N04	What is the sum of the pulmonary (bacteriologically confirmed), pulmonary (clinically diagnosed), extrapulmonary (bacteriologically confirmed or diagnosed)? [B2]	B = <input type="text"/> <input type="text"/> <input type="text"/>	
N05	Does the total number of TB cases (all types) equal the sum of the number of TB cases for smear positive pulmonary, smear negative pulmonary, no smear done pulmonary, and extrapulmonary TB (i.e. does A equal to B)? PLEASE CIRCLE YES OR NO. If no, please describe the discrepancy.	Yes	No, Describe:
N06	What is the total number of new and relapsed TB cases who are male? [B3a]	C = <input type="text"/> <input type="text"/> <input type="text"/>	
N07	What is the sum of the age-disaggregated new and relapsed TB cases by who are male? [B3]	D = <input type="text"/> <input type="text"/> <input type="text"/>	

⁹¹ Depending on the country situation, completion of these tools may require adjustment depending on whether the BMU is at the district level or at the health facility level. Consultation with the NTP and technical partners is strongly recommended prior to undertaking this assessment.

Number	Question	Result	Skip
TB QUARTERLY REPORT: TB BMU REPORTING VERIFICATION			
BMU IDENTIFICATION			
N08	Does the total number of new and relapsed TB cases who are male equal the sum of the age-disaggregated new and relapsed TB cases who are male (i.e. does C equal to D)? PLEASE CIRCLE YES OR NO. If no, please describe the discrepancy.	Yes	No, Describe:
N09	What is the total number of new and relapsed TB cases who are female? [B4a]	E = <input type="text"/> <input type="text"/> <input type="text"/>	
N10	What is the sum of the age-disaggregated new and relapsed TB cases by who are female? [B4]	F = <input type="text"/> <input type="text"/> <input type="text"/>	
N11	Does the total number of new and relapsed TB cases who are female equal the sum of the age-disaggregated new and relapsed TB cases who are female (i.e. does E equal to F)? PLEASE CIRCLE YES OR NO. If no, please describe the discrepancy.	Yes	No, Describe:
N12	What is the sum of the sex-disaggregated totals for all new and relapsed TB cases? [B3a,B4a]	G = <input type="text"/> <input type="text"/> <input type="text"/>	
N13	What is the sum of all new TB cases (B5) and all relapse cases (B6)?	H = <input type="text"/> <input type="text"/> <input type="text"/>	
N14	Does the sum of the sex-disaggregated totals for all new and relapsed TB cases match the sum of all new TB cases and all relapse cases? (i.e. does G equal to H)? PLEASE CIRCLE YES OR NO. If no, please describe the discrepancy	Yes	No, Describe:
N15 ⁹²	For each facility in this BMU, for the verification period covering MONTH X TO MONTH X 201X , fill in the following:		

92 If the country's BMU is at a level higher than the health facility level e.g. district, then question N15 could be completed by using the BMU TB register.

Number	Question	Result	Skip
TB QUARTERLY REPORT: TB BMU REPORTING VERIFICATION			
BMU IDENTIFICATION			
Facility Code (a)	Facility Name (b)	Recount the total number of TB cases (all types) (c)	TB cases transferred in (d)
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

1.2 Example quarterly report on TB case registration in the basic management unit. *The bold letters and numbers refer to the quarterly report verification tool.*

Name of BMU:		Facility:		Patient registered during		Quarter of year	
Name of TB Coordinator:		Signature:		Date of completion of this form:			

Block 1: All TB cases registered during the quarter					
	New	Relapse	Previously treated (excluding relapse)	Previous treatment history unknown	Total
Pulmonary, bacteriologically confirmed	B2	B2	B2	B2	
Pulmonary, clinically diagnosed	B2	B2	B2	B2	
Extrapulmonary, bacteriologically confirmed or clinically diagnosed	B2	B2	B2	B2	B2
Total (all types)	B5	B6			B1

Block 2. All new and relapse cases (bacteriologically confirmed or clinically diagnosed) registered during the quarter by age group and sex								
	0-4	5-14	15-24	25-34	35-44	55-64	>65	Total
Male	B3	B3	B3	B3	B3	B3	B3	B3a
Female	B4	B4	B4	B4	B4	B4	B4	B4a

1.1: Quarterly report verification tool – example

Number	Question	Result	Skip
FACILITY HEALTH OFFICE: TB REPORTING VERIFICATION			
HEALTH FACILITY IDENTIFICATION⁹³			
1	Health facility name		
2	Health facility code	CODE <input type="text"/> <input type="text"/> <input type="text"/>	
3	Does this health facility send quarterly TB reports to the national TB programme?	YES..... NO.....	1 2 → END
ASK THE TB PROGRAMME OFFICER IN THE HEALTH OFFICE TO PROVIDE YOU WITH THE SOURCE DOUMENTS THAT ARE USED TO COMPILE THE TB QUARTERLY REPORTS COVERING [QUARTER X – MONTH 1 TO MONTH 3 2012] .			
4	Are the TB quarterly report, health facility TB register, laboratory TB register and patient treatment cards covering [QUARTER X – MONTH 1 TO MONTH 3 2012] available for review?	Yes, all documents are available PARTLY, some documents are available DESCRIBE NO, none of the documents are available	1 2 3 → END
5	From the TB quarterly report , count the total number of TB cases (all types) for the verification period (MONTH1 TO MONTH3 2012). IF THE TB QUARTERLY REPORT IS NOT AVAILABLE, RECORD "N/A".	TOTAL TB CASES (ALL TYPES) REPORTED IN TB QUARTERLY REPORT A = <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	
6	From the health facility TB register , count the total number of TB cases (all types) for the verification period (MONTH1 to MONTH3 20XX). IF THE TB REGISTER IS NOT AVAILABLE, RECORD "N/A".	TOTAL TB CASES (ALL TYPES) COUNTED FROM HEALTH FACILITY TB REGISTER B = <input type="text"/> <input type="text"/> <input type="text"/> <input type="text"/>	

⁹³ Depending on the organisation of data flow, completion of these tools may require adjustment depending on whether the BMU is at the district level or at the health facility level. Consultation with the NTP and technical partners is strongly recommended prior to undertaking this assessment.

Number	Question	Result
FACILITY HEALTH OFFICE: TB REPORTING VERIFICATION		
HEALTH FACILITY IDENTIFICATION		
7	<p>From the health facility TB register, count the number of TB cases that were transferred in for the verification period (MONTH1 to MONTH3 20XX).</p> <p>IF THE TB REGISTER IS NOT AVAILABLE, RECORD "N/A".</p>	<p>TB CASES TRANSFERRED IN</p> <p>C = <input type="text"/><input type="text"/><input type="text"/><input type="text"/></p>
8	Calculate	<p>TB CASES FROM HEALTH FACILITY TB REGISTER</p> <p>D = B - C = <input type="text"/><input type="text"/><input type="text"/><input type="text"/></p>
9	<p>Count the total number of bacteriologically-confirmed TB cases from the laboratory TB register <i>linked</i> with the BMU for the verification period (MONTH1 TO MONTH3 2012).</p> <p>IF THE LABORATORY REPORT IS NOT AVAILABLE, RECORD "N/A".</p>	<p>TOTAL BACTERIOLOGICALLY-CONFIRMED TB CASES REPORTED IN THE LABORATORY TB REGISTER</p> <p>E = <input type="text"/><input type="text"/><input type="text"/><input type="text"/></p>
10	<p>Count the total number of patient treatment cards for the verification period (MONTH1 TO MONTH3 2012).</p> <p>IF TREATMENT CARDS ARE NOT AVAILABLE, RECORD "N/A".</p>	<p>TOTAL PATIENT TREATMENT CARDS REPORTED</p> <p>F = <input type="text"/><input type="text"/><input type="text"/><input type="text"/></p>
11	<p>Copy the total number of TB cases from the HMIS facility monthly reports for the verification period (MONTH1 TO MONTH3 20XX).</p> <p>IF THE HMIS MONTHLY REPORT IS NOT AVAILABLE, RECORD "N/A".</p>	<p>TOTAL TB CASES REPORTED FROM HMIS MONTHLY REPORTS</p> <p>G = <input type="text"/><input type="text"/><input type="text"/><input type="text"/></p>
12	<p>What are the reasons for the discrepancy (if any) observed between A WITH B, D, E, F and G? e.g. data entry errors, arithmetic errors, gaps or missing information in source documents, etc.</p>	
UG_TB	<p>Missing data: Count the number of cases in [QUARTER X – MONTH 1 TO MONTH 3 20XX] with missing information for each of the following columns in the BMU TB register⁹⁴:</p>	

94 This is assuming the health facility register is equivalent to the BMU TB register. If both a BMU TB register and health facility register exists, then the BMU TB register should be assessed.

Number	Question	Result
FACILITY HEALTH OFFICE: TB REPORTING VERIFICATION		
HEALTH FACILITY IDENTIFICATION		
01	Year of registration	<input type="text"/> <input type="text"/> <input type="text"/>
02	Sex	<input type="text"/> <input type="text"/> <input type="text"/>
03	Age	<input type="text"/> <input type="text"/> <input type="text"/>
04	Disease classification	<input type="text"/> <input type="text"/> <input type="text"/>
05	Type of patient	<input type="text"/> <input type="text"/> <input type="text"/>
06	Bacteriological results	<input type="text"/> <input type="text"/> <input type="text"/>
07	Number of TB cases missing data in at least one of the 6 columns	<input type="text"/> <input type="text"/> <input type="text"/>

Appendix 2:

Applying the lot quality assurance sampling (LQAS) theory

In the context of this standard, LQAS theory can be used to test whether the country overall has a *pre-specified* “acceptable” level p of: 1) internal consistency; and, 2) data accuracy and completeness. The “acceptable” threshold p is in turn translated into a threshold (denoted d) of the allowable number of sampled BMUs (which are the “lots”) that are unacceptable in terms of each of the outcomes of interest (each corresponding to a benchmark). This value d is referred to as the “decision rule” in the LQAS literature. If the number of sampled BMUs found to be unacceptable is the same or below the pre-defined critical value d then the country overall is considered to be acceptable and meets the benchmark.

The selection of the value of d is as much an informed choice as it is a calculation. The value of d is based partly on statistical considerations, through the selected values for the pre-specified acceptable threshold level of the percentage of matched records to source documents and the statistical error that is allowed; partly on logistical and feasibility considerations; and partly on what investigators consider acceptable or unacceptable in relation to the total sample size of required BMUs. The stricter the study team decides to be in terms of selecting small values for d , the smaller the required sample size n will be. The larger d becomes the larger the total sample size n also becomes, hence weakening the advantage of small sample sizes that LQAS produces. Examples of sample size calculations for different values of n and d based on the LQAS approach are shown in Table 1.

Table 1. Example sample size calculations for n BMUs to be included in the investigation, for a pre-specified level of an “acceptable” percentage of $p=95\%$ and varying: 1) total numbers N of BMUs in the country, 2) values of decision interval d , and 3) values of the statistical type I error α

Total number of BMUs in the country N		$d = 0$	$d = 1$	$d = 2$	$d = 3$	$d = 4$	$d = 5$
100	$\alpha = 0.05$	45	65	80	92	99	100
	$\alpha = 0.1$	37	57	74	88	98	100
150	$\alpha = 0.05$	52	76	96	114	129	142
	$\alpha = 0.1$	42	65	86	105	122	137
200	$\alpha = 0.05$	51	76	97	117	135	152
	$\alpha = 0.1$	41	64	84	104	123	141
300	$\alpha = 0.05$	54	80	103	125	145	165
	$\alpha = 0.1$	42	66	87	108	128	147
500	$\alpha = 0.05$	56	83	108	131	153	175
	$\alpha = 0.1$	43	68	90	111	131	151
1,000	$\alpha = 0.05$	57	86	112	136	159	181
	$\alpha = 0.1$	44	69	92	113	134	153
10,000	$\alpha = 0.05$	59	89	115	140	164	187
	$\alpha = 0.1$	45	71	93	115	135	155

Table 1 shows that the number of BMUs n that are required for the study increases as: 1) the total number of BMUs in the country N increases and 2) the decision interval d increases. On the other hand, the larger the allowable statistical error α is, the lower the required sample size becomes.

The sample sizes in Table 1 were calculated using the R computer code shown below. Different scenarios can be explored according to different country contexts using this code. R is freely available for download at: <http://www.r-project.org>

Box 1. R code that can be used to explore different sample size scenarios

```
nsiz<- function (p = 0.95, N = 1000, d = 1, alpha = 0.05){  
  s <- N  
  for (n in N:1){  
    m <- N - n  
    k <- trunc ((1-p) * N)  
    if (dhyper (d, n, m, k) > alpha) break  
    s <- n  
  }  
  return (s)  
}
```

The code shown in Box 1, when entered in the R console, it generates a function named **nsiz**. The function code is minimalist and does not include checks for improper parameter values. The function can be used for different values of the acceptable threshold level of matched audited records to source documents p , the decision threshold d , the total number of BMUs in the country N , and the type I error α . In the first example in Box 2 below, **nsiz** returns a sample size of BMUs $n=76$ under the default assumptions $p=0.95$, $N=200$, $d=1$ and $\alpha=0.05$. In the second example in Box 2, the returned sample size $n=57$ corresponds to assumptions $p=0.95$, $N=1,000$, $d=0$ and $\alpha=0.05$.

Box 2. Examples of sample size calculations using the code provided in Box 1

```
nsiz(p=0.95, N=200, d=1, alpha=0.05)  
[1] 76  
nsiz(p=0.95, N=1,000, d=0, alpha=0.05)  
[1] 57
```

Appendix 3:

Examples of methods used in The Netherlands, UK and USA to ensure and improve internal consistency of TB surveillance data at the national level

The Netherlands

Within The Netherlands, the number of TB cases are compared against external and internal registration systems and trends of TB cases over time are also examined:

- Aggregated TB mortality within the surveillance system is compared annually with the aggregated death rates reported by the Central Bureau of Statistics (CBS).⁹⁵ Increases and decreases are routinely monitored in both systems, and trends are explored, if needed.
- All culture-positive TB cases (approximately 70%) examined by the National Reference Laboratory are linked to the patients reported in the surveillance system on a weekly basis.
- Treatment outcome results are integrated in the case-based surveillance system.
- The national data management unit validates all new notified cases on a daily basis (real-time surveillance) and notifies involved authorities in case of a significant increase in new TB cases. In the case of a significant decrease, the primary data sources are checked to ensure there has been no under-reporting.
- Comparison of quarterly and annual reports with previous reporting periods to look changes in trends within TB surveillance data. This is done using the 'eye-ball test'; no statistical tests are used.

⁹⁵ See <http://www.cbs.nl/nl-NL/menu/home/default.htm>

United Kingdom

There are a several different strategies used to ensure the number of reported TB cases is internally consistent in the UK:

- Public Health England, which manages the TB surveillance system (ETS) in the UK, obtains mortality data from the Office of National Statistics (ONS)⁹⁶ on a weekly basis and these data are used to validate TB death data from the ETS on an annual basis using aggregated numbers.
- Trends in TB mortality over time are examined separately for ETS data (TB caused, contributed to, was incidental to or unknown relationship to death) and ONS data (underlying cause is TB). In 2012, data from the two systems were linked, using all ONS records in which TB was mentioned (rather than just those where TB was the underlying cause of death on the death certificate). Based on this analysis, cases with an unknown relationship to death in ETS were updated using ONS data, and this procedure is now a routine consistency check. In the future, if there are more TB deaths found in ONS than in ETS, the regional co-ordinators will be contacted to follow these cases up for notification and query the treatment outcomes recorded in ETS with the clinics. Furthermore, in the future, laboratory isolates that are not notified on a regular basis (monthly) will be matched with the mortality data, based on previous studies that have shown that a higher proportion of not notified laboratory cases had died compared to those that had been notified.
- Patients reported in the ETS are linked with culture positive cases examined by the National Reference Laboratory. Since 2010, all drug resistant isolates from the lab are followed-up for notification with the clinics. If the isolates are not matched to cases by clinics, the isolate data remain in the system indefinitely, prompting the regional coordinators to follow-up the clinics to notify the case.
- Treatment outcome results are integrated in the ETS. Regional coordinators follow-up 12 month treatment outcomes on an annual basis and, as of 2013, will begin following-up all outcomes over 12 months.
- Within the UK, regions compare their own regional data over time and notify the national team if sudden changes are seen in the data. At the national level, to make sure there is nothing unexpected in the data, they are compared over time using statistical tests, e.g. chi-square test for trend to examine proportions of drug resistance over

96 See <http://www.statistics.gov.uk/hub/index.html>

time, and analyses are conducted on sub-groups (chi-square tests for the proportion of cases that are, for example, male, pulmonary, UK-born, and by age group, region, history of previous treatment) comparing the current and previous year. The ratio of children to adults over time is also examined. A sudden increase or decrease in the frequency of certain variables that is not as expected would prompt an investigation into the validity of the data with the reference laboratories and the software development unit. Also, a dashboard was developed with the Department of Health, which examines the proportion of cases completing treatment by the health protection unit and whether it is above or below the target of 85%. The dashboard also compares the incidence rate in the current and previous year using confidence intervals. If rates exceed the previous year or treatment outcome falls below the target the health protection unit (for areas with over 20 cases per year) is alerted to look into this in more detail.

United States of America

In the USA, there are two ways to assess changes in TB case rates. The first method is to compare the reported TB cases by state within the National Tuberculosis Surveillance System (NTSS) with the verbal count each state gives to the Centers for Disease Control and Prevention (CDC) each year. This ensures that case counts match at both state and national levels. The second method is to compare the change in rate each year against what is the expected rate for the nation. If the change in the national rate is above or below the expected rate, more vigorous statistical methods and investigations may be undertaken.

Winston et al⁹⁷ described such an investigation in a recently published paper. In 2009, an unexpected and considerably steeper decline in reported TB case count was observed in the USA compared with recent years. Based on NTSS provisional data in March 2010, national TB case rates declined -11.4% in 2009 compared to an average annual -3.8% decline since 2000. CDC worked with partners to investigate if the decline reflected changes in surveillance reporting or diagnosis or other factors. Trends from multiple sources on TB treatment initiation, medication sales, and laboratory and genotyping data on culture-positive TB were examined. Over 142,000 incident TB cases reported to the US NTSS during January 1, 2000-December 31, 2009 were analysed. The investigation also included analysis of TB control programme data from 59 public health reporting areas, self-reported data from 50 CDC-

97 Winston CA, Navin TR, Becerra JE, et al. Unexpected decline in tuberculosis cases coincident with economic recession - United States, 2009. , 2011;11:846.

funded public health laboratories, monthly electronic prescription claims for new TB therapy prescriptions, and complete genotyping results available for NTSS cases. Accounting for prior trends using regression and time-series analyses, the deviation between observed and expected TB cases in 2009 was calculated according to patient and clinical characteristics, and the point in time the deviation occurred was assessed. The overall deviation in TB cases in 2009 was -7.9%, with -994 fewer cases reported than expected ($P < 0.001$). Other independent information systems (TB prescription claims, and public health laboratories) reported similar patterns of declines. Genotyping data did not suggest sudden decreases in recent transmission. The assessments show that the decline in reported TB was not an artefact of changes in surveillance methods. While the steady decline of TB cases before 2009 suggested ongoing improvement in TB control, no substantial changes in TB control activities or TB transmission that would account for the abrupt decline in 2009 were identified. Other multiple causes coincident with economic recession in the United States, including decreased immigration and delayed access to medical care, could be related to TB decline.

ISBN 978 92 4 150672 4

