WHO guidance on TB surveillance (2024)

Chapter 7: Data quality

TB Monitoring, Evaluation and Strategic Information Unit
Global Tuberculosis Programme

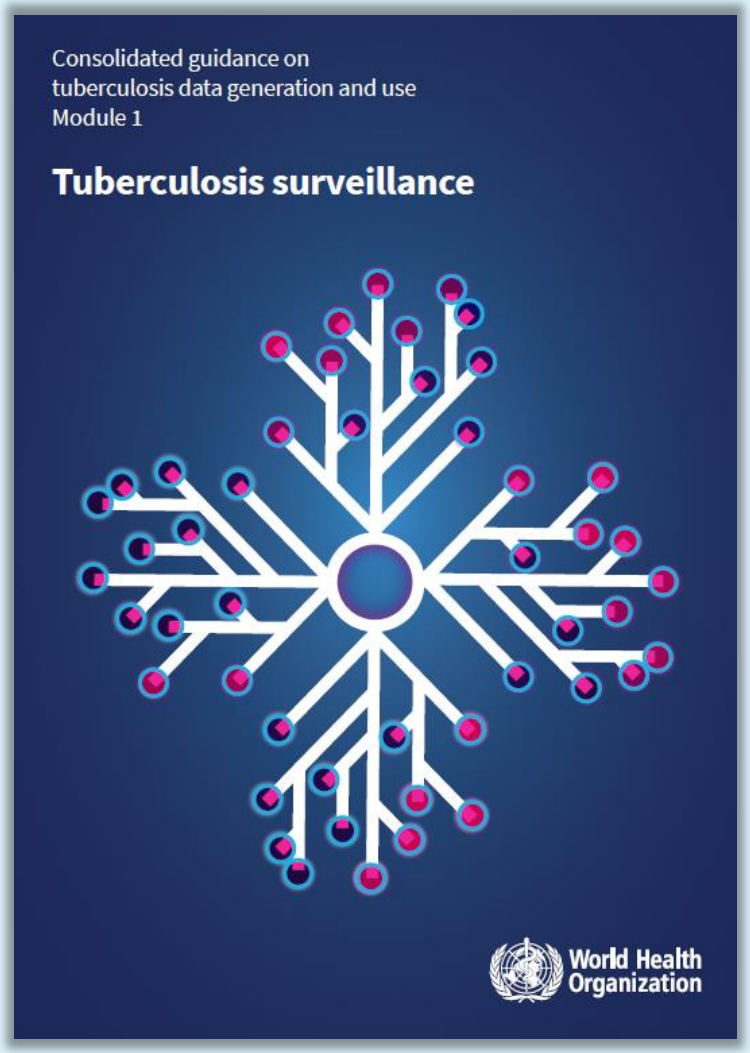
World Health Organization

Available May 2024

Consolidated guidance on tuberculosis data generation and use. Module 1:

Tuberculosis surveillance

https://iris.who.int/handle/10665/376612



World Health Organization

Chapter contents

Chapter 7 is divided in four major sections:

Section 7.1 Dimensions of data quality

Section 7.2 Governance and design features of a TB surveillance system that can help to ensure data quality

Section 7.3 Routine data validation checks

Section 7.4 Periodic evaluations of data quality

7.1 Dimensions of data quality

World Health Organization

Dimensions of data quality

Section 7.1 introduces the six dimensions of data quality that are commonly used in public health literature and provides a definition for each term.

1. **Accuracy:** data are a correct description of reality and free from errors.

2. **Completeness:** the extent to which expected data are available.

3. **Consistency:** data are both internally consistent (e.g. over time) and externally consistent (e.g. with expected benchmarks or values derived from research studies, or comparable with data from similar countries).

4. **Timeliness:** data are available at the time they are needed.

5. **Validity:** the extent to which data conform to the expected format, type and range.

6. **Uniqueness:** there is not more than one record for any given observation.

7.2 Governance and design features of a TB surveillance system that can help to ensure data quality

World Health Organization

Governance and design features to ensure data quality

Section 7.2 discusses seven key considerations for a well-governed and designed TB surveillance system and how it can help to ensure high data quality at all levels.

1. Data governance:

- The roles of the NTP (or equivalent), the entity responsible for the health information system, and other relevant agencies (e.g. those responsible disease control and prevention) should be clearly defined and agreed.
- How the TB surveillance system interacts or is integrated with the overall architecture for public health surveillance should be clearly specified.

2. Data sharing framework and agreements:

- The collection of data about people with TB disease and TB infection is not restricted to the NTP's TB surveillance system. Other sources of data include:
  - Databases for notifiable diseases
  - Databases managed by other national disease programmes
  - Databases managed by other ministries
  - Databases managed by health insurance schemes
- Legal frameworks and formalized agreements for sharing of data are important to facilitate cross-checking of TB surveillance data.

Governance and design features to ensure data quality

Section 7.2 (continued)discusses seven key considerations for a well-governed and designed TB surveillance system and how it can help to ensure high data quality at all levels.

3. Data confidentiality and safety:

- Clear policies and processes must be in place to ensure the confidentiality of TB surveillance data.
- Recommended approaches that promote data confidentiality and safety:
  - Identify all legal requirements regarding data confidentiality
  - Implement role-based access permissions
  - Restrict access to physical data files which should be held on a server with no direct end-sure access
  - Ensure physical security of all devises to protect against theft
  - Use secure data encryption methods

Governance and design features to ensure data quality

Section 7.2 (continued)discusses seven key considerations for a well-governed and designed TB surveillance system and how it can help to ensure high data quality at all levels.

4. Provision of feedback among all system:

- To promote better acceptance of a digital surveillance system, data entry should be an element of engaging users in an interactive system and not be perceived as a mandatory chore.
- Feedback loops to engage users who enter data and create ownership can include:
  - Prompting users to act upon data in the system, such as tracing missing data or contacting patients who have not attended a visit.
  - Generating reports on standard performance indicators, such as case notifications, treatment outcomes and data completeness rates.
  - High administrative levels informing lower levels of data problems to be corrected.

Governance and design features to ensure data quality

Section 7.2 (continued)discusses seven key considerations for a well-governed and designed TB surveillance system and how it can help to ensure high data quality at all levels.

5. Software:

- The chosen software must allow for the creation of a database with functional capacity for storage, management, validation and analysis of data.
- Web- and case-based digital systems that use unique identifiers are the reference standard.
- Use of non-specialized software for data management (e.g. Microsoft Excel©) should be avoided.

6. Financial and human resources:

- Adequate and sustained financial and human resources is required for the design, roll-out and maintenance of a TB surveillance system.
- Data quality is strongly influence by adequate training of staff on data collection, management, reporting, analysis and use. Training of new staff and regular refresher training of existing staff is important.
- Expertise in data management, infectious disease epidemiology, statistics and information technology are essential, especially at the national level.
- Supervision and periodic data audits at the health facility should be carried out routinely.

Governance and design features to ensure data quality

Section 7.2 (continued)discusses seven key considerations for a well-governed and designed TB surveillance system and how it can help to ensure high data quality at all levels.

5. Clear specification of indicators and data items to be collected, reported and used, and associated standardized tools and operating procedures:

- The indicators to be reported and used, and the data items that need to be collected, should be clearly specified. (See Chapters 4 and 5 in the TB surveillance guidance).
- Standardized tools and SOPs should be developed and disseminated so that they are readily available by relevant staff. These include:
  - Standardized reporting tools for core data items which are supported by training materials and SOPs.
  - Guidance and protocols for data cleaning and how to deal with common sources of errors in a standardized and systematic way.

7.3 Routine data validation checks

World Health Organization

Routine data validation checks

Section 7.3 discusses the routine data validation checks that should be done at the health facility level and at higher administrative levels for both digital and paper-based TB surveillance systems.

At the health facility level:

- Pertains to data that are recorded at the health facility either on paper or entered into a digital system.

Digital systems	Paper-based systems
• Examples of data validation checks during data entry are provided in Table 7.1 of the TB surveillance guidance. • Automated data validation checks can be built directly into the system. • Error messages should alert users when a validation check fails, prompting them to investigate and take correct actions. • A well-designed set of data validation checks at data entry will help to ensure data are of high quality.	• Data validation checks are manual and can be cumbersome and are error-prone. • Manual checks for missing values, misclassifications and duplicate records for the same treatment episode. • Manual checks to ensure data are consistent across forms and registers. • Manual checks to ensure data that are then entered into a digital aggregate system are consistent with source paper records.

Routine data validation checks

Section 7.3 discusses the routine data validation checks that should be done at the health facility level and at higher administrative levels for both digital and paper-based TB surveillance systems. (continued)

At higher administrative levels:

- Pertains to the transmission of data towards the national level, which may require aggregation of the data up the administrative hierarchy or merging or appending multiple datasets to establish a consolidated dataset at the national level.

Digital systems	Paper-based systems
• Examples of data validation checks at higher administrative levels are provided in Table 7.2 of the TB surveillance guidance. • Automated data validation checks can be built directly into the system or can be carried out on the data using a statistical software by staff with relevant expertise. • Checks can help to identify common sources of error, to produce a list of data to check per facility, and to emphasize potential problems at the health facility.	• Data validation checks are manual and can be cumbersome and are error-prone. • Some of the example data validation checks provided in Table 7.2 of the TB surveillance guidance could be implemented. • Periodic audits of data collected in paper-based TB surveillance systems should be planned for and budgeted.

7.4 Periodic evaluations of data quality

World Health Organization

Periodic evaluations of data quality

Section 7.4 discusses the importance of carrying out periodic evaluations of data quality through self-assessments and through an external, independent audit.

Table 7.3 of the TB surveillance guidance presents the main components of self-assessments and audits. These are summarized below:

Characteristic	Self assessments	Independent audits
Description	Assessment of data quality from reporting units	Assessment of data quality of the national surveillance system
Objective	Self-assessment of data quality and associated capacity in data recording and reporting, and data management	Comprehensive and independent assessment to evaluate the system's ability to record and report high-quality TB data
Frequency	Routine	Periodic
Implementation	By the NTP	By (programme-) independent, external auditors
Resource requirements	Usually low-to-medium resource intensive	Medium-to-high resource intensive
Output	Action plan detailing best practices, corrective measures, required resources and timelines	Formal report of system performance detailing recommendations for the NTP and stakeholders and often informs the development of national strategic plans, TB programme reviews, development of funding applications.
Resources	Manual on use of routine data quality assessment (RDQA) for TB monitoring	WHO data quality assurance toolkit WHO Surveillance Checklist (see Web Annex B).

For further information or in case of any questions, contact:

tbdata@who.int

World Health Organization

Links to the guidance on TB surveillance

Consolidated guidance on tuberculosis data generation and use. Module 1: Tuberculosis surveillance
https://iris.who.int/handle/10665/376612

Web annex A: Commonly observed problems and associated solutions.
https://iris.who.int/handle/10665/376481.

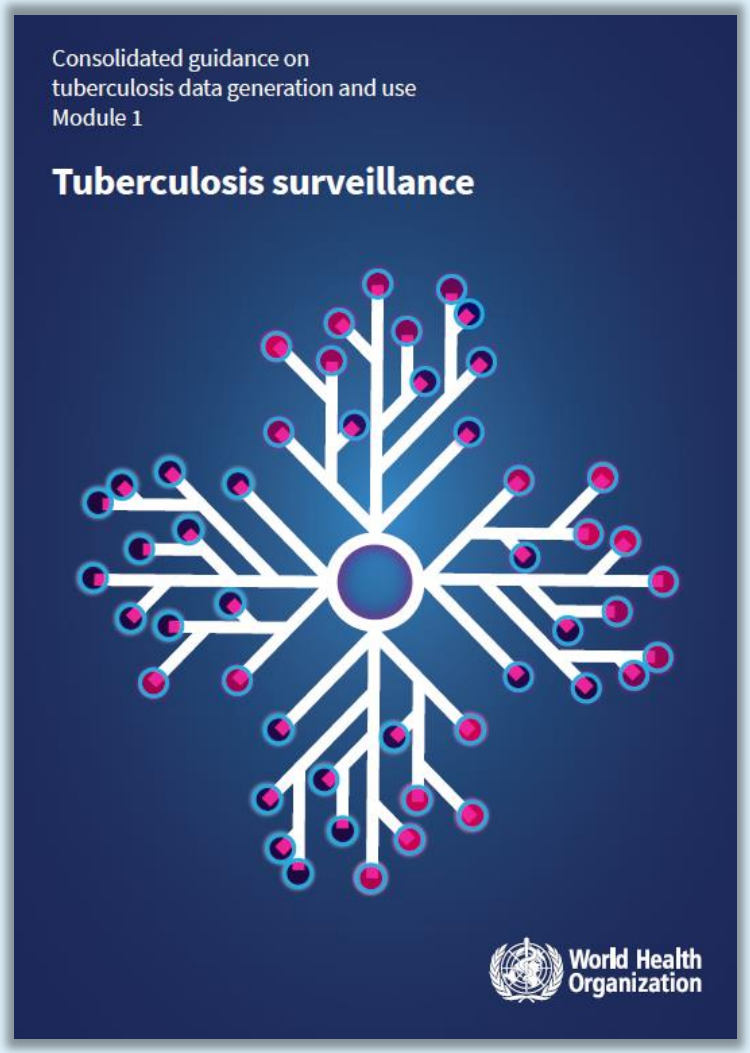
Web annex B: Standards and benchmarks for tuberculosis surveillance and vital registration systems: checklist, 2nd ed.
https://iris.who.int/handle/10665/376483

Web annex C: Record-linkage exercises.
https://iris.who.int/handle/10665/376484

Web annex D: Reporting of aggregated data and calculation of core indicators: templates and formulae.
https://iris.who.int/handle/10665/376486

Web annex E: Examples of how to report diagnosis, start of treatment and treatment outcomes.
https://iris.who.int/handle/10665/376489

Web annex F: Evaluation of the WHO DHIS2 case-based package for tuberculosis surveillance (TB tracker) in five pilot countries: summary of key findings.
https://iris.who.int/handle/10665/376490



World Health Organization