Standard Operating Procedure

Tuberculosis Data Quality Checklist

Abbreviations

CENAT   National Center for Tuberculosis and Leprosy Control

DQR   Data Quality Review

HF   health facility

M&E   monitoring and evaluation

OD   operational district

TB   tuberculosis

VF   verification factor

WHO   World Health Organization

Purpose

The Tuberculosis (TB) Data Quality Review (DQR) tools will be routinely utilized by TB supervisors at the operational district (OD) level to conduct data quality checks on reporting completeness and timeliness, data accuracy, and internal consistency of reported data. Hence, this Standard Operating Procedure (SOP) is designed to provide guidance to the TB OD level supervisors on how to conduct the data quality reviews and implement the DQR tools (DQR data quality checklist and Excel spreadsheet).

Scope

The TB OD level supervisors will utilize the DQR Excel spreadsheet to monitor the quality of the data collected and reported monthly from the health facility (HF) under their administration areas.

Operations

On the fifth day of each month, TB OD level supervisors receive a monthly report from the HFs under their administration area. Once the report is received, the TB supervisors will go through the DQR checklist that focuses on reporting completeness, reporting timeliness, completeness of indicator data, verification of data accuracy, and internal consistency over time in the report. At this stage data in the report is not evaluated. The TB OD supervisor proceeds to count the number of reports received and reviews the data recorded in the report to determine if all needed sections the HF has to complete are filled in. This process is called **"reporting completeness."** If the report is complete, then the OD level TB supervisor will fill '1' in the Excel spreadsheet, or 'X' otherwise**.**

Then, the TB OD supervisor verifies the timeliness of the report, meaning whether the report was received per the scheduled date. In this case, it is the fifth day of each month. If OD receives the report according to schedule, then the OD will fill '1' in the Excel spreadsheet, or '0' otherwise. At this stage, only the timeliness is being evaluated. This is called **"reporting timeliness."**

The TB OD supervisor proceeds to check the list of the selected indicators to be reviewed and verifies that all the selected indicators are reported. If all the selected indicators have been entered, then the OD will fill '1' in the Excel spreadsheet, or '0' otherwise. This process of detailed verification is called **"completeness of reported indicator data."**

From this step onward, the OD verifies the data recorded field by field, comparing it to the data source available at the HF. The main sources of the data are the (1) TB registration book, (2) patient card, and (3) lab request 01.

A. Reporting Completeness

- Column 'A': Enter the name of the HF.
- Column 'B' to 'M': Enter the number '1' in the reporting month if the data recorded are complete, otherwise enter a '0.'
- Column 'N': The number of expected reports to be received at the OD in a year. So, column N remains as 12 (12 reports/year).
- Column 'O': Sums up the numbers from columns 'B to M' automatically.
- Column 'P': Calculates column 'O' into a percentage (%) and informs the OD of the completeness rate from each HF. It is calculated as the actual number of monthly reports received and the completed indicators data by the OD divided by the expected number of monthly reports to be sent to the OD. After completing the data entry from columns 'B' to 'M', Column 'P' automatically changes color. If the reporting completeness rate is below 75 percent, it will change to red; between 75 percent to 90 percent, it will change to blue; between 90 percent to 100 percent, it will change to green; and higher than 100 percent, it will change to orange.
- The 'summary results' contains the number and percentage of the HF's report completion rate: 100 percent, 75 percent to 99 percent, below 75 percent, and more than 100 percent. The HFs with a reporting completeness rate below 75 percent are marked red and considered to have poor report completeness. The HFs reporting more than 100 percent need to be further investigated to determine why the reporting rate is more than 100 percent.

Name of health facilities	Enter the number of monthly reports received by OD from the health facilities												Expected no. monthly reports to be sent to OD	Actual no. of monthly reports received by OD	Reporting completeness rate (%) (P=O/N*100)
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec			
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
វាលអង្គពោល	1												12	1	8
ពោធិ៍មាស	1	1	1	1	1	1	1	1	1	1	1	1	12	12	100
និទាន	0	1						1		1			12	3	25
ស្វាយចប	0	0	1						1				12	2	17
កក់ព្រះខែ				1									12	1	8
ពោធិ៍ចំរើន				1									12	1	8
ព្រៃញាតិ					1								12	1	8
ពោធិ៍អង្គង							1						12	1	8
U													12	0	0
V													12	0	0
OD totals	2	2	2	2	2	2	2	2	2	2	1	1	120	22	18

OD TB Supervisor's Checklist for Reporting Completeness Assessment

Metrics	Summary results	
	Number	**Percent**
Number and percent of HFs with reporting completeness rate between 75%-90%	0	0
Number and percent of HFs with reporting completeness rate below 75%	9	90
Number and percent of HFs with reporting completeness rate 90-100%	1	10
Number and percent of HFs with reporting completeness rate more than 100 %	0	0

Total number of health facilities   10

B. Reporting Timeliness

- Column 'A': Name of the HF. It auto-populates from the previous form (reporting completeness).
- Column 'B': Actual number of reports received by the OD during a year (12 months). It auto-populate based on the reporting completeness table column 'O' ('Actual no. of monthly reports received by OD').
- Column 'C' to 'N': Enter the number '1' if the report was received within the first five days of the month, otherwise enter a '0.'
- Column 'O': Sums up the numbers from columns 'C' to 'N' automatically.
- Column 'P': Calculates column 'O' into a percentage and informs the OD whether the report received is within the fifth day of the following month. If the received reports have not submitted on time (<75%), the cells in column 'P' will change to the color red. If most reports have been submitted on time (between 75% to 99%) the cells will change to color blue.
- The 'summary results' contains the number and percentage of the HF's reporting timeliness rate: 100 percent, 75 percent to 99 percent, and below 75 percent. The HFs with a reporting timeliness rate below 75 percent are marked red and considered to have poor on time reporting.

Reporting Timeliness Assessment Tool															
OD Supervisors Checklist for Assessment of Reporting Timeliness															
Names of health facilities	Actual numberf month reports receiv by OD during the year	Monthly reports received by OD by the report submission deadline												Total number of monthly reports received by OD by submission deadline	Report submission timeliness rate (P = O / B*100)
		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec		
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
វាលអង្គារពោល	1	1												1	100
ពោធិ៍មាស	12	1	1	1	1	1	1	1	1	1	1	1	1	12	100
និទាន	3	0	1					1		1				3	100

ស្វាយចចប	2	0	0	1						1				2	100
កក់ព្រះខែ	1				1									1	100
ពោធិ៍ចំរើន	1					1								1	100
ព្រែកញ្ញាតិ	1						1							1	100
ពោធិ៍អង្គង	1							1						1	100
U	0													0	#DIV/0!
V	0													0	#DIV/0!
OD Totals	22	2	2	2	2	2	2	2	2	2	2	1	1	22	100

Metrics	Summary Results	
	Number	**Percent**
Number and percent of HF with timeliness rate 75% or below	**0**	**0**
Number and percent of HF with reporting completeness rate between 75%-99%	**0**	**0**
Number and percent of HF with 100% reporting timeliness	**8**	**100**
Number and percent of HF more 100% reporting timeliness	0	0

Total health facilities   **8**

C. Completeness of Reported Indicator Data

- Column 'A': Name of the HF. It auto-populates from the previous form (reporting completeness).
- Columns 'B' to 'M': Enter the number '1' if the indicator data reported are complete in the report, otherwise leave it blank. The fields remain empty if the HF did not report the indicator regardless of availability of data (data source). Actual indicators of the reports with completed data (non-missing values) will be auto-populated in column 'O.' The number zero '0' does not indicate a missing value, but rather is the actual data reported in the report (ex: TB notification indicator - some HFs did not have TB notification data in month two ('Feb'), so a 'o' has been entered in the cell. In this case, the '0' has value and should be filled in as '0' in the 'Feb' column. Please see the HF (Svay Chorchorb).
- Column 'N': Total expected number of values (i.e., one for one indicator for 12 months).
- Column 'O': Sums up the numbers from columns 'B' to 'M' automatically.
- Column 'P': Calculates column 'O' into a percentage and informs the OD whether the report received has all of the values for the indicators. If the received report is not complete or is less than 90 percent, the number in column 'P' will change to the color blue; if it is 100 percent complete, then it will change to the color green; if it is 91 percent to 100 percent, it will change to the color red; and if it is more than 100 percent, then it will change to the color dark orange.
- At the beginning of each quarter, the TB OD supervisor should receive a list of indicators to conduct the data quality review. When the TB OD supervisor receives the report, they will verify the data reported for each indicator (all needed cells are complete or left blank). If

a zero ('0') is filled in, the TB OD supervisor should verify if that '0' is a true zero value or '0' where it is not zero (a non-true zero value). If their HF has a rate of 10 percent or less than 10 percent missing values, it is marked as red.

OD TB Supervisor's Checklist for Assessment of Completeness of Indicator Data **(List the name of the indicator used for reporting here….)**															
Name of health facilities	Enter the number of monthly Indicator report received by OD from the health facilities												Total expected no. of values for the indicator	Total actual no. of non-missing values for the given indicator	Completeness rate
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec			
A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
វាលអង្គពេល	1												12	1	8
ពោធិមាស	1	1	1	1	1	1	1	1	1	1	1	1	12	12	100
និទាន	0	1						1		1			12	4	33
ស្វាយចប	0	0	1					1					12	4	33
កក់ព្រះខែ				1									12	1	8
ពោធិចំរើន					1								12	1	8
ព្រៃញាតិ						1							12	1	8
ពោធិអង្គង							1						12	1	8
U													12	0	0
V													12	0	0
OD totals	0	1	3	1	1	0	1	2	0	2	0	0	120	25	21

Metrics	**Summary results**
	Number
Number of health facilities with completeness rate below 90 %	9
Number of health facilities with 100 % expected results	1
Number of health facilities between 91-99 %	0
Number of health facilities with reporting completeness rate more than 100 %	0

Total number of health facilities   **10**

D. Verification of Data Accuracy

- Column 'A': Name of the HF. It auto-populates from the previous form (reporting completeness).
- Column 'B': Enter the total number of indicators data reported in the reporting month.
- Column 'C': Enter the number referenced in the data source at the HF (e.g., TB patient registration book, TB patient treatment card).
- Column 'D': Divides automatic columns 'B'/'C' to calculate the verification factors (VF). This VF will inform the TB OD supervisors of whether the data reported and the data recorded in the data source is overreported (column 'E') or underreported (column 'F').
- Column 'E': Counts how many HFs underreported.
- Column 'F': Counts how many HFs overreported.
- Column 'G': Counts how many HFs reported data that matches the source data file exactly ('matching range').
- A HF a VF of less than 0.9 or greater than 1.1 are indicative of data quality problems and that HF should be investigated.

Data Accuracy Check						
Name of health facilities	Data reported in the monthly report	Figure recounted from the TB register	Verification Factor	VF < 0.90	VF > 1.10	VF = 1.0 (within +/- 10%)
			VF = C/B	(over-reporting)	(under-reporting)	(Exactly matches the reported data)
A	B	C	D	E	F	G
M	1	1	1.00	0	0	1
N	1	0	0.00	1	0	0
O	1	0	0.00	1	0	0
P	1	1	1.00	0	0	1
Q	1	1	1.00	0	0	1
R	1	1	1.00	0	0	1
S	1	1	1.00	0	0	1
T	1	1	1.00	0	0	1
U	1	1	1.00	0	0	1
V	1	1	1.00	0	0	1
Total number of HFs over-reporting				2		
Total number of HFs under-reporting					0	
Total number of HFs exactly matching						8

E. Internal Consistency Over Time

Examines the plausibility of the same reported indicators data based on the history of reporting in the previous months.
- First column: Name of the HF. It auto-populates from the previous form (reporting completeness).
- Columns 'A' to 'L': Enter the number of the indicators data recorded in the report of the preceding month.
- Column 'M': Enter the current month (ex: the current month is April, so fill in the numeric value for the month '4').
- Column 'N': An average of the preceding 12 months.
- Column 'O': Ratio of the current month to the average of previous months. Current month (column 'M' numerator)/average of the preceding months (column 'N' denominator).
- Column 'P': If there is a deviation from the average trend seen in the HF (more that 33% either way), then there can be a data quality issue. If there is a difference of more than 33 percent between the health facility and the OD, then it is highlighted in red and flagged as indicating potential data quality problems.

Name of health facilities	Preceding Months (Specify below)												Current month (Specify below)	Average of preceding 12 months in 2020 G = (A+B+C+D+E+F+G+I+J+K+L)/12	Ratio of current month to the average of preceding 12 months (O = M/N)	% Difference between health centre ratio and OD ratio [O (health centre) - O (OD)] / O (OD) X 100
	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec	Apr-23			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
វាលអង្គារពេល	1												4	0	48	675
ពោធិ៍មាស	2	5	6	2	3	6	7	1	1	7	2	2	4	4	1	-82
និទាន		0						0		1			4	0	48	675
ស្វាយចប			4						1				4	0	10	55
កក់ព្រះខែ					1								4	0	48	675
ពោធិ៍ចំរើន					3								4	0	16	158
ព្រៃញាតិ						4							4	0	12	94
ពោធិ៍អង្រុង							3						4	0	16	158
U														0	#DIV/0!	#DIV/0!
V														0	#DIV/0!	#DIV/0!
OD Totals	3	5	10	2	7	10	10	1	2	8	2	2	32	5	6	

Metrics	Summary Results	
	Number	Percent
HC with at 33% or more difference between the HF and operational district ratio	7	87.5
HC with at less than 33% difference between the HF and operational district ratio	1	12.5

Total number of health facilities   8

Dashboard

The summary of the DQR is visualized in the last tab of the spreadsheet titled "Dashboard." The TB OD supervisor can utilize the dashboard to present the findings of their data quality review to the health facility or TB provincial supervisor.
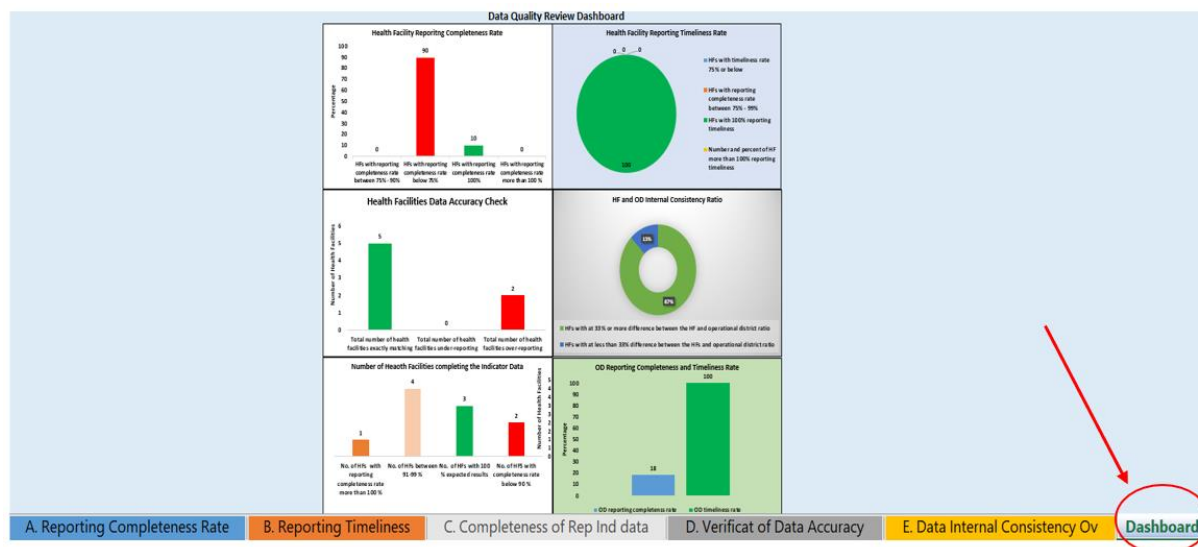
List of Data Sources

- Monthly reporting form
- TB registration book
- A copy of the TB patient treatment card
- Lab 01

References

World Health Organization. (2020). Data quality assurance (DQA). Retrieved from https://www.who.int/data/data-collection-tools/health-service-data/data-quality-assurance-dqa

World Health Organization (WHO). (2020). Data quality review: a toolkit for facility data quality assessment. Module 1. Framework and metrics. Geneva: WHO. Retrieved from https://cdn.who.int/media/docs/default-source/data-quality-pages/2021_dqa_module-1-framework-and-metrics-19-04-21.pdf?sfvrsn=13c95fb1_3&sequence=1&isAllowed=y

World Health Organization (WHO). (2020). Data quality review: a toolkit for facility data quality assessment. Module 2. Desk review of data quality. Geneva: WHO. Retrieved from https://cdn.who.int/media/docs/default-source/data-quality-pages/2021_-dqa_module-2_desk-review-of-data-quality.pdf?sfvrsn=7a0999e_9

World Health Organization (WHO). (2020). Data quality review: a toolkit for facility data quality assessment. Module 3. Data verification and system assessment. Geneva: WHO. Retrieved from https://cdn.who.int/media/docs/default-source/data-quality-pages/2021_dqa_module-3_site-assessment-framework.pdf?sfvrsn=4e936e2f_9&csf=1&web=1&e=DEfO5f

World Health Organization. (2012). Assessment of health facility data quality: Data quality report card Cambodia, 2012. Retrieved from https://cdn.who.int/media/docs/default-source/data-quality-pages/kh_dataqualityreportcard_2012.pdf?sfvrsn=6c66226c_5