Data Quality Review

A toolkit for facility data quality assessment

Working document
September 2015

Acknowledgements

This toolkit is the result of collaboration between the World Health Organization, the Global Fund, GAVI and USAID/MEASURE Evaluation. The toolkit proposes a unified approach to data quality. It integrates and builds upon previous and current tools and methods designed to assess facility-level data quality, taking into account best practices and lessons learned from many countries.

Kavitha Viswanathan (WHO Geneva) oversaw the technical development of the toolkit, with overall guidance provided by Kathryn O'Neill (WHO Geneva) and with technical input from Ties Boerma, David Boone (John Snow, Inc.), Robert Pond, Olav Poppe, Claire Preaud, Ashley Sheffel and Marina Takane.

Important contributions were provided by technical experts from a broad range of public health programmes and include Thomas Cherian, Marta Gacic-Dobo, Jan Van Grevendonk (WHO Immunization, Vaccines and Biologicals); Richard Cibulski, Michael Lynch (WHO Malaria); Katherine Floyd, Philippe Glaziou, Hazim Timimi (WHO Tuberculosis); Isabelle Bergeri, Txema Callejas, Chika Hayashi (WHO HIV); Ryuichi Komatsu, John Puvimanasinghe, Annie Schwartz, Alka Singh, Nathalie Zorzi (The Global Fund); Peter Hansen, Hope Johnson, Chung-won Lee (GAVI).

Particular thanks are due to the country partners that tested and implemented different components of the approach and provided valuable feedback, namely: Benin, Burkina Faso, Cambodia, Democratic Republic of Congo, Kenya, Mauritania, Sierra Leone, Togo, Uganda and Zanzibar.

Financial support was provided by GAVI, the Global Fund and the Norwegian Agency for Development Cooperation (Norad).

Contents

Part 1:
Framework & metrics

1.1 Background

Data are widely used for a variety of purposes – including health sector reviews, planning, programme monitoring, quality improvement and reporting. For this purpose, it is critical to have high-quality data on performance in the health sector available at least routinely.

The national health management information system (HMIS) and other parallel reporting systems (where they exist) collect data on routine health activities and health problems that are reported from health facilities in the national health-care system. These health facility data are a primary source for assessing health sector performance – i.e. the Ministry of Health (MOH) compiles the data on a regular basis to report on achievements and trends in key health performance indicators. However, HMIS data often exhibit problems of quality, and many users do not trust these data.

All data are subject to quality limitations such as missing values, bias, measurement error, and human errors in data entry and computation. Data quality assessments should always be undertaken to understand how much confidence can be placed in the health data that is used to assess health sector performance and to understand the relative strengths and weaknesses of the various data sources. In particular, it is important to know the reliability of national coverage estimates and other results derived from health facility data.

Various data quality assessment mechanisms are used by national authorities and other partners to examine the quality of health facility data. In addition, electronic health information systems have some built-in checks of data quality. However, the different tools and approaches have certain limitations, including the following:

- National disease programmes carry out assessments of data quality and of quality of services individually, making it difficult to assess the capacity of health facilities in a comprehensive way. Data issues often cut across programmes and it is more efficient to examine them holistically.
- Data quality assessment efforts have not usually been part of a regular data quality monitoring system which feeds into the health sector review. Also, the assessments are often not conducted according to a well-defined, standardized methodology.
- The sample size used by these methods is often not large enough to be representative of all health facilities, thus making it difficult to reach broad conclusions about reporting accuracy. Small sample sizes can also affect the level of confidence one places in the results.

This framework and toolkit represent a collaborative effort of WHO, the Global Fund and GAVI to promote a harmonized approach to assess the quality of data reported from the level of health facilities to the national level.

This Data Quality Review (DQR) methodology builds on existing data quality assurance mechanisms; the methodology and indicators have been developed and selected on the basis of broad consultation with international health programme experts from leading donor and technical assistance agencies. It is expected that individual programmes will use the findings of a completed DQR to inform their respective detailed assessments of data quality and programme-specific systems. The DQR in no way replaces routine monitoring, supervision and evaluation intended to strengthen individual programmes.

The ultimate goal of the DQR is to contribute to the improvement of the quality of data used by countries for reviews of progress and performance – such as annual health sector reviews, programme planning, and monitoring and evaluation – in order to facilitate decision-making.

1.2 | Overview

Sound decisions are based on sound data; therefore it is essential to ensure that the data are of good quality. Health facility data constitute a primary data source for assessing the performance of the health sector. Ministries of health therefore compile data regularly to track progress towards goals and objectives, plan for future needs, and set priorities for the health system. However, data of poor quality result in a lack of trust among users.

A comprehensive and holistic review of the quality of data collected from health facilities requires a multi-pronged approach, including:

- routine and regular (i.e. monthly) reviews of data quality built into a system of checks of the HMIS or other programme reporting systems as part of a feedback cycle that identifies errors in near real-time so they can be corrected as they occur;
- an annual independent assessment of core tracer indicators in order to identify gaps and errors in reporting and the plausibility of trends in health facility data used for annual health sector planning, monitoring and evaluation reported during the previous year; and
- periodic in-depth programme-specific reviews of data quality that typically focus on a single disease/programme area, conducted as required by specific programmes – e.g. prior to programme reviews.

Scope of DQR

The scope of this version of the DQR is to support annual and periodic independent assessments of facility-reported data. The periodicity depends on the focus of the review – i.e. whether it is to look at common cross-cutting data quality issues that must be taken into consideration and addressed when preparing annual health analytical reports, or whether it is to look in greater depth with a programme-specific approach in advance of programme reviews. A subsequent version of the DQR will address routine and regular checks and validation of data.

The DQR does not preclude the need for ongoing routine checks and validations of data that are internal to the HMIS and are part of routine data management procedures. Routine and regular reviews of data are critical components of health information and reporting systems. While this version of the DQR does not specifically address routine and regular data quality checks within the HMIS, the proposed DQR metrics can be used and can be incorporated in the routine internal checks and controls of data quality. Work is underway to incorporate DQR metrics in the District Health Information System 2.0 (DHIS2[1]) . For countries that have the DHIS2 system as their routine HMIS, this addition will greatly facilitate regular

[1] DHIS2 is a web-based, open-source software that is used by countries chiefly as their health information system for data management and monitoring of health programmes. It has also been used for logistics management, mobile tracking and facility registers. More information can be found at: https://www.dhis2.org/ (accessed 29 May 2015).

data quality assessments. In addition, there are other existing tools that can be used for routine data quality assessment such as the Data Quality Self-Assessment (DQS)[2].

Objectives

The DQR is designed to assess the quality of data generated by information system(s) based in health facilities. The objectives of the DQR are:

- to institutionalize an annual system of data quality assessment including verification of data and periodic in-depth assessments;
- to identify weaknesses in the data management system and interventions for system strengthening; and
- to monitor data quality performance over time and the capacity to produce good-quality data.

Links to country planning and review processes and intended use

A comprehensive overview of the quality of routine data reported by health facilities should be conducted annually as part of the data consolidation process for annual statistical reports or health sector performance reports. This would allow wide dissemination of assessment results within the Ministry of Health, development partners and other stakeholders, and would show whether health facility data used in annual health sector reviews are of adequate quality for effectively monitoring progress and performance. Poor-quality data can call into question demonstrations of progress towards health sector objectives and are detrimental to annual planning processes since the plans are based on faulty data. It is therefore crucial to discuss any issues pertaining to data quality, to identify measures to improve data quality and to develop action plans to implement such measures. When determining funding levels for programmes and priority areas of the health system, health planners need to know what level of trust they can place in the data. Planners also need to know what investments they must make to strengthen data quality and reporting systems. This also applies to in-depth DQRs, with the results included in the programme review or annual health sector review.

Coordination

A national coordinating body or mechanism should be put in place to facilitate planning and marshal resources. This body should have the authority to make the decisions necessary to plan and implement the DQR. The team should comprise health sector stakeholders from government, the private sector, the donor community, and multinational organizations such as WHO, GAVI and the Global Fund. Monitoring and evaluation technical working groups or health information system governance boards already exist in many countries and can serve as the DQR coordinating mechanism. The DQR can be implemented using in-country resources (e.g. Ministry of Health, national statistics bureaux, public health schools), or countries may request technical assistance from external sources or from technical assistance partners such as WHO, UNAIDS and bilateral donors.

[2] World Health Organization. Immunization Data Quality Self-Assessment Tool. http://www.who.int/immunization/monitoring_surveillance/routine/coverage/en/index3.html (Accessed September 23, 2015)

The toolkit

The DQR toolkit includes guidelines and tools that lay the basis for a common understanding of data quality so that a regular mechanism for data quality assessments can be institutionalized in the country. The toolkit enables countries to conduct regular data quality assessments in accordance with the following structure:

- Part 1: Framework and metrics
- Part 2: Country planning and implementation
- Part 3 : Analysis, dissemination and use
- Annexes: Additional resource materials
  - Recommended core and additional indicators
  - Data quality metrics: definitions and requirements
  - Data and formatting requirements for the desk reviews
  - Recommended source documents and spot-checks for data verification
  - DQR data collection instruments
  - A tool developed in Excel that will automate analyses of data quality metrics
  - Electronic data-collection forms (in this version of the toolkit, electronic forms have been developed for use with the free software CSPro[3]).

[3] For information about the Census and Survey Processing System (CSPro), including free download, see: http://www.census.gov/population/international/software/cspro/ (accessed 29 May 2015).

1.3 | DQR methodology

The DQR methodology comprises two separate but interrelated processes – a desk review of the data that have been reported to national level and a sample facility assessment to undertake a data verification exercise. A system assessment can also be carried out at the time of the facility assessment visit.

The DQR methodology includes:

- a desk review component whereby the quality of aggregate reported data for recommended programme indicators is examined using standardized data quality metrics (most of the data quality metrics for the DQR are compiled and analysed during this component);
- a health facility assessment consisting of two data collection modules.

Desk review

The desk review is an evaluation of data quality dimensions (completeness, internal consistency, etc.). Normally, the desk review requires monthly or quarterly data by subnational administrative area for the most recent reporting year and annual aggregated data for the last three reporting years for the selected indicators.

Through analysis of these programme indicators, the process quantifies problems of data completeness, accuracy and consistency, and thus provides valuable information on the adequacy of health facility data to support planning and annual monitoring. WHO recommends that the desk review component of the DQR be conducted annually.

The desk review has two levels of data quality assessment:

- an assessment of each indicator aggregated to the national level;
- the performance of subnational units (e.g. districts or provinces/regions) for the selected indicators.

Facility assessment (site visit to sampled facilities)

Data verification

Data verification is an internal consistency measure of the DQR and requires primary data collection at the facility level.

The objective of data verification is to measure the extent to which the information in the source documents has been transmitted correctly to the next level of reporting; the verification applies to each level of the reporting hierarchy (from the health facility level to the national level). This allows systematic errors that occur in the reporting of data to be identified and gives an estimate of the degree of over-reporting or under-reporting in the system at national level for specific indicators.

For data verification, data from source documents (registers and tally sheets) are compared to data reported through the HMIS in order to determine the proportion of the reported numbers that can be verified from the source documents. The values for selected indicators and reporting periods are recounted using the relevant source document at the facility. The recounted (or verified) value of the

indicator (for the selected reporting period) is then compared to the value reported by the facility for the same reporting period. A standardized data collection instrument is available in both paper and electronic format.

It is essential to use a sound probability-based sampling methodology so that the results of the data verification are representative of all the health facilities. A nationally representative health facility assessment usually has a sample of more than 100 health facilities, which constitutes a sufficient sample for data verifications. The primary data collection can be conducted as part of a larger health facility assessment, such as a Service Availability and Readiness Assessment (SARA), or as a discrete event.

It is recommended that data verification be conducted annually along with the desk review, if possible as part of a harmonized health facility assessment plan.

System assessment

The system assessment is contained in an additional module of a health facility survey and can be conducted at the same time as the verification of data at health facility and district levels. The system assessment measures the capacity of the system to produce good-quality data. It evaluates the extent to which critical elements of the reporting system adhere to a set of minimum acceptable standards. The elements of the reporting system that are evaluated in the system assessment are as follows:

- monitoring and evaluation (M&E) structure and function;
- indicator definitions and reporting guidelines;
- data collection tools and reporting forms;
- data quality monitoring and supervision;
- data maintenance and confidentiality;
- demographic information (district level only);
- use of data for decision-making (district level only).

The Global Fund and GAVI have used a system assessment as part of their data quality assessment methodologies for a number of years, and the system assessment within the DQR is harmonized from the tools developed by these organizations.

While the system assessment is not a measure of data quality, it is included in this toolkit because it provides information that will potentially enable managers to determine the causes of data quality problems. Consequently, it is recommended that the system assessment should be periodically implemented with the DQR, with the data verification module. It is not necessary to conduct a system assessment annually; it should be conducted every 3–5 years or as needed. While the system assessment can be a conducted as a discrete activity, it is recommended that it be part of a larger health facility assessment.

1.4 | Data quality metrics

The DQR examines a set of standard indicators that are routinely reported through facility information systems and quantifies any problems of data completeness, timeliness, consistency and reliability in order to ascertain to what extent the health facility data are fit for purpose. For example:

- Quality data should be complete and timely – i.e. there is sufficient information available when required to make decisions about the health of the population and to target resources to improve health system coverage, efficiency and quality.
- Quality data should be consistent and reliable – i.e. there is sufficient consistency to compare results confidently from different parts of the country or from different periods in time.
- Quality data should be accurate – i.e. they have sufficient detail to answer the questions we have about the performance of the health system for vulnerable subpopulations.

The DQR examines the quality of data of a selected number of indicators covering the different programme areas that are reported through routine facility information systems. The DQR may be implemented as a holistic review across several programme areas or as an in-depth assessment of a particular programme area.

Core indicators

The proposed core indicators were selected on the basis of their importance for programme monitoring and evaluation. They include core tracer indicators on antenatal care, immunization, HIV, tuberculosis (TB) and malaria. Table 1.1 lists the recommended core or "tracer" indicators recommended for regular DQR.

Table 1.1 Recommended core indicators for the DQR

Recommended DQR indicators		
Programme area	**Indicator name**	**Full indicator**
Maternal health	Antenatal care 1st visit (ANC1)	Number (%) of pregnant women who attended at least once during their pregnancy
Immunization	DTP3/Penta3	Number (%) of children < 1 year receiving three doses of DTP/Penta vaccine
HIV/AIDS	ART coverage	Number and % of people living with HIV who are currently receiving ART
TB	Notified cases of all forms of TB	Number (%) of all forms of TB cases (i.e. bacteriologically confirmed plus clinically diagnosed) reported to the national health authority in the past year (new and relapse)
Malaria	Confirmed malaria cases[4]	Number (%) of all suspected malaria cases that were confirmed by microscopy or RDT

Note: ANC = antenatal care; ART = antiretroviral therapy; DTP3 = diphtheria-tetanus-pertussis three-dose vaccine; Penta = pentavalent vaccine; RDT = rapid diagnostic test.

While it is recommended that countries should select indicators from the core list, they may select other indicators or expand the set of indicators on the basis of their needs and the resources available. A full set of core and additional indicators is available in Annex 1**.** It is important to note, however, that the more indicators that are selected, the more time-consuming and expensive the exercise will be. This is

[4] If the number of confirmed malaria cases is not collected, total malaria cases can be substituted.

particularly relevant to the selection of indicators for the data verification component. A guiding principle is that a team of data collectors should not spend more than one day in each facility. Thus it is recommended that no more than 4–5 indicators should be included at any one time for the record review exercise.

If other priority indicators are not included in the core or additional lists, they can be selected (cautiously[5]) to replace one or more of the core indicators.

Dimensions of data quality

This DQR framework examines each of the selected indicators from four perspectives, or dimensions, namely:

- Dimension 1: completeness and timeliness of data;
- Dimension 2: internal consistency of reported data;
- Dimension 3: external consistency – i.e. agreement with other sources of data such as surveys; and
- Dimension 4: external comparisons of population data (a review of denominator data used to measure performance indicators).

Completeness and timeliness

The completeness of the data is assessed by measuring whether all the entities which are supposed to report actually do so. This applies to health facility reporting to districts and to district reporting to the regional or provincial level. Timeliness of data is assessed by measuring whether the entities which submitted reports did so before a predefined deadline.

Internal consistency of reported data

Three measures of internal consistency are included in the DQR. These are:

- ***Consistency over time:*** The plausibility of reported results for selected programme indicators are examined in terms of the history of reporting of the indicators. Trends are evaluated to determine whether reported values are extreme in relation to other values reported during the year or over several years.

- ***Consistency between indicators:*** Programme indicators which have a predictable relationship are examined to determine whether, in fact, the expected relationship exists between those indicators. In other words, this process examines whether the observed relationship between the indicators, as depicted in the reported data, is that which is expected.

- ***Consistency of reported data and original records:*** This involves an assessment of the reporting accuracy for selected indicators through the review of source documents in health facilities. This element of internal consistency is measured through a data verification exercise which requires a record review to be conducted in a sample of health facilities. It is the only dimension of data

[5] Not all data quality metrics apply to all indicators. In addition, it might be difficult to get denominators at the sub-national level or make comparisons with measures of the indicator from other sources for some of the core and additional indicators (e.g. in HIV). In this case, the data quality checks can be performed only on the numerator data (the metrics for which are included in the DQR dimensions 1 and 2).

quality that requires additional collection of primary data. Data verification examines the agreement between the total number of service outputs recorded in source documents at the health facility and the total number of service outputs reported to the reporting system (either HMIS or programme-specific reporting system) for selected indicators. Values of selected indicators for a given reporting period are recalculated, using the primary sources of data for the indicators. The recalculated value is then compared to the value that was initially reported through the system for the given reporting period. The ratio of the recounted value to the reported value is called the "verification ratio" and constitutes a measure of accuracy for the indicator. This exercise should be conducted at the facility level, and again at the district and provincial levels, and a verification ratio should be calculated for each level.

External consistency with other data sources

The level of agreement between two sources of data measuring the same health indicator is assessed. The two sources of data usually compared are data flowing through the HMIS or the programme-specific information system and a periodic population-based survey. The HMIS can also be compared to pharmacy records or other types of data to ensure that the two sources fall within a similar range

External comparisons of population data

This involves determining the adequacy of the population data used in the calculation of health indicators. Population data serve as the denominator in the calculation of a rate or proportion and provide important information on coverage. This data quality measurement compares two different sources of population estimates (for which the values are calculated differently) in order to ascertain the level of congruence between the two. If the two population estimates are discrepant, the coverage estimates for a given indicator can be very different even though the programmatic result (i.e. the number of events) is the same. The higher the level of consistency between denominators from different sources, the more likely it is that the values represent the true population value.

Definitions and benchmarks

It is also useful to establish a benchmark that reflects the desired/acceptable level for each of the metrics for each of the core indicators. Benchmarks for quality will depend on the country implementing the DQR. For instance, a reporting rate of 80% might be acceptable in one country with historically low reporting performance but not in another country which has more mature systems and current reporting rates closer to 100%. Benchmarks for quality can vary across programme areas for certain data quality metrics. For instance, the recommended threshold of quality for completeness of indicator data in maternal health might be 90%, but for immunization it could be 67% since there is often variability in immunization service delivery from month to month and it is not unusual to find zero values (or missing values). Similarly, the threshold for TB might be 75% since TB is a relatively rare event in the population, particularly in sparsely populated subnational administrative areas.

Countries with mature information systems, with standardized indicators and tool, and a well-trained workforce, should expect to have more stringent thresholds for quality than countries without.

Table 1.2 shows the different metrics that are included in each of the four dimensions of data quality. The quality of data of recommended core indicators is examined against these standard metrics. The benchmarks for measuring quality are also shown. These recommended benchmarks should be tailored to

the country context. More detailed descriptions of data quality definitions, requirements and calculations are contained in Annex 2.

Table 1.2 Data quality dimension, metrics and standard benchmarks

DIMENSION 1: COMPLETENESS OF REPORTING			
An assessment of each dimension should be conducted for each of the recommended core indicators: antenatal care (ANC), immunization, HIV, TB and malaria. Additional indicators can be selected according to the priority and focus of the data quality assessment.			
Data quality metric	**Definition**		
	National level		**Subnational level**
Completeness of district reporting	% of expected district monthly reports (previous 1 year) that are actually received		Number and % of districts that submitted: 1) at least 9 out of 12 expected monthly reports; 2) 100% of expected monthly reports
Timeliness of district reporting	% of submitted district monthly reports (previous 1 year) that are received on time (i.e. by the deadline for reporting)		Number and % of districts that submitted on time at least 75% of the monthly reports received at national level from the district[6]
Completeness of facility reporting	% of expected facility monthly reports (previous 1 year) that are actually received		Number and % of districts with at least 9 out of 12 monthly facility reports received
			Number and % of facilities that submitted 100% of expected monthly reports
Timeliness of facility reporting	% of submitted facility monthly reports (previous 1 year) that are received on time (i.e. by the deadline for reporting)		Number and % of districts that received on time at least 75% of monthly facility reports that were submitted
Completeness of indicator data (% of data elements that are non-zero values, % of data elements that are non-missing values – do each analysis separately)	ANC 1st visit		Number and % of districts with < 90% 1) non-zero values; 2) non-missing values
	3rd dose DPT-containing vaccine[7]		Number and % of districts with < 67% 1) non-zero values; 2) non-missing values
	ART coverage		Number and % of districts with < 90% 1) non-zero values; 2) non-missing values
	Notified cases of all forms of TB[8]		Number and % of districts with < 75% 1) non-zero values; 2) non-missing values
	Confirmed malaria cases		Number and % of districts with < 90% 1) non-zero values; 2) non-missing values
Consistency of reporting completeness	Each information system	Evaluate the trend in completeness of reporting from district to national level over the past 3 years	Evaluate the trend in completeness from facility to district level over the past 3 years

Note: ART = antiretroviral therapy.

[6] Denominator is reports received (not expected).
[7] Immunization programmes expect some months will have zero values for vaccination indicators.
[8] TB reporting generally takes place quarterly.

DIMENSION 2: INTERNAL CONSISTENCY OF REPORTED DATA		
Data quality measure	**Definition**	
	National level	**Subnational level**
Outliers[9] Complete for each of 5 indicators: - ANC 1st visit - 3rd dose DPT-containing vaccine - ART coverage - notified cases of all forms of TB - confirmed malaria cases	**Extreme:** % of monthly subnational unit values that are extreme outliers (at least 3 SD from the mean)	Number and % of subnational units in which 1 or more of the monthly subnational unit values over the course of 1 year is an extreme outlier
	Moderate: % of subnational unit values that are moderate outliers (±2–3 SD from the mean or > 3.5 on modified Z-score method).	Number and % of subnational units in which 2 or more of the monthly subnational unit values for the indicator over the course of one year are moderate outliers
Consistency over time Complete for each of 5 indicators: - ANC 1st visit - 3rd dose DPT-containing vaccine - ART coverage - notified cases of all forms of TB - confirmed malaria cases tested	Conduct one of the following based on the expected trend of the indicator: • comparison of current year to the value predicted from the trend in the 3 preceding years (for indicators or programmes with expected growth), or • comparison of current year to the average of the 3 preceding years (for indicators or programmes expected to remain constant)	Number and % of districts whose current year-to-predicted value ratio (or current year to the average of the preceding three years) is at least ±33% different from the national ratio
	Graphic depiction of trend to determine plausibility based on programmatic knowledge	
Consistency between related indicators	**Maternal health:** ANC1 – IPT1 or TT1 (should be roughly equal)	Number and % of subnational units where there is an extreme difference (≥ ±10%)
	Immunization: DTP3 dropout rate: (DTP1–DTP3)/DTP1 - should not be negative	Number and % of subnational units with the number of DTP3 immunizations higher than DTP1 immunizations (negative dropout)
	HIV/AIDS: ART coverage – HIV care coverage (ratio should be less than 1)[10]	Number and % of subnational units where there is an extreme difference (≥ ±10%)
	TB: TB cases notified – TB cases put on treatment (in the past year) (should be roughly equal)	Number and % of subnational units where there is an extreme difference (≥ ±10%)
	Malaria: Number of confirmed malaria cases reported – cases testing positive (should be roughly equal)	Number and % of subnational units where there is an extreme difference (≥ ±10%)
Verification of reporting consistency through facility survey	% agreement between verified counts for selected indicators in sampled facility records, and reported values for the same facilities	**Maternal health:** ANC 1st visit
		Immunization: Penta/DTP 1–3 in children < 1 year
		HIV/AIDS: HIV coverage
		TB[11]**:** Notified cases of all forms of TB
		Malaria: Suspected malaria cases tested

Note: ANC = antenatal care; ART = antiretroviral therapy; SD = standard deviations; DTP3 = diphtheria-tetanus-pertussis vaccine third dose; IPT = intermittent preventive therapy; TT = tetanus toxoid vaccine; Penta = pentavalent vaccine.

[9] For programmes with inconsistent levels of service delivery and for which outliers are common (e.g. immunization), a customized threshold can be set on the basis of programmatic knowledge. Data that have high variability month to month can also be evaluated for outliers using the modified Z-score method (see section 3.1) which is based on the median and has higher tolerance for extreme values than the standard deviation method.

[10] The extent of difference between the two indicators depends on the national treatment guidelines and when people living with HIV/AIDS are eligible for antiretroviral therapy.

[11] Sampling of health facilities requires stratification by facility type in order to ensure an adequate number of facilities providing TB services.

DIMENSION 3: EXTERNAL COMPARISON (Comparison of routine data with population-based survey values from the same period)[12]		
Indicator	**Definition**	
	National level	**Subnational level**
ANC 1st visit	Ratio of facility ANC1 coverage rates to survey ANC1 coverage rates	Number and % of aggregation units used for the most recent population-based survey (such as a province/state/region) whose ANC1 facility-based coverage rates and survey coverage rates show at least 33% difference.
3rd dose DPT-containing vaccine	Ratio of DTP3 coverage rates from routine data to survey DTP3 coverage rates	Number and % of aggregation units used for the most recent population-based survey (such as a province/state/region) whose DTP3 facility-based coverage rates and survey coverage rates show at least 33% difference.
HIV	---	---
TB[13]	---	---
Malaria IPT		
Comparison between programme and HMIS values	For selected indicators, compare the value aggregated for 12 months from the HMIS with the programme data	For selected indicators, compare the subnational unit values aggregated over 12 months for number and % of districts with > 10% difference in annual values between the HMIS and the programme data.

Note: ANC = antenatal care; ART = antiretroviral therapy; DTP3 = diphtheria-tetanus-pertussis three-dose vaccine; IPT = intermittent protective therapy.

DIMENSION 4: EXTERNAL CONSISTENCY OF POPULATION DATA (Evaluation of adequacy of denominators used for calculating performance indicators)		
Indicator	**Definition**	
	National level	**Subnational level**
Consistency of population projections	Ratio of population projection of live births from the Country Census Bureau/Bureau of Statistics to a United Nations projection of live births for the country	NA
Consistency of denominator between programme data and official government population statistics	Ratio of population projection for selected indicator(s) from the census to values used by programmes	Number and % of subnational units where there is an extreme difference (e.g. ± 10%) between the two denominators.
Consistency of population trend	Ratio of population values for selected indicator(s) from the current year to the predicted value from the trend in population values up to 3 preceding years	Number and % of subnational units where there is an extreme difference(e.g. ± 10%) between the two denominators.

[12] Complete for each programme area (if sufficient recent survey data are available). Administrative data should preferably be from the same year as the survey value. Denominators used for coverage estimates from administrative data may need adjustment to make them comparable to survey values (e.g. women attending ANC at public facilities).
[13] No viable survey indicator for TB.

Part 2:

Country planning and implementation

2.1 | Planning the DQR

Step 1. Establish a DQR coordinating group at national level

Bringing country stakeholders together is a critical first step towards successful implementation of DQR. One of the first activities is to identify and establish a group of core stakeholders at country level to oversee, coordinate and facilitate the planning and implementation of the DQR and the dissemination and use of the DQR findings.

The group should comprise technical focal points among health-sector stakeholders from government (including the different programme stakeholders), development partners and multinational organizations such as WHO, GAVI and the Global Fund. Monitoring and evaluation technical working groups or health information system governance boards, which already exist in many countries, can serve as the DQR coordinating team. Development and technical partners can greatly contribute to the success of efforts to improve data quality and should agree on a standardized set of data quality indicators.

The role of the DQR coordinating group is to:

- develop a harmonized plan for data quality assessments;
- identify technical support requirements for implementation and quality assurance;
- identify funding sources;
- oversee the selection of core indicators and the establishment of benchmarks;
- monitor implementation of the DQR;
- ensure promotion and dissemination of the findings.

Step 2. Develop a harmonized plan for data quality assessments

The DQR coordinating team creates a schedule for the DQR linked to the annual planning cycles of the Ministry of Health. The results of the DQR should be available in advance of the planning so that stakeholders will understand the strengths and limitations of the data used for planning.

A harmonized plan for data quality assessments should ideally include:

- annual independent DQRs of the core indicators, including an annual verification of data quality on a sample of facilities, timed so that the results can be used to prepare the annual statistical reports/analytical performance reviews;
- periodic independent assessments of programme-specific data from health facilities (every 3–5 years to support programme reviews);
- development and monitoring of data quality improvement plans.

Step 3. Develop an implementation plan and budget

In order for the results to be available for the Health Sector Review, the DQR should be conducted well in advance to allow time to correct data or fill gaps if necessary. Depending on whether the DQR is conducted with a health facility survey or as a stand-alone exercise, planning and implementation may require up to 6 months, or up to 3 months, respectively. If a country undertakes only the desk review and does not conduct primary data collection for data verification and the M&E system assessment, the DQR can be completed in around 1 month.

An implementation plan should be based on the purpose and components of the DQR being considered. The DQR coordination committee should decide on the mechanisms for implementation. The DQR has two components (a national-level desk review and a health facility survey), so implementation mechanisms should be considered for both. As data quality is important to many donors, the DQR coordination team should explore whether in-country partners can support the process. Partners often have funds allocated to data quality assurance mechanisms and may be willing to assist with implementation. For instance, if funds have been allocated in the GAVI health system strengthening support for a health facility survey or for data quality assessment, the DQR team should explore whether these funds can be used for a DQR. The identification of potential funding mechanisms is made easier if the DQR coordination team is a multistakeholder committee comprising persons from the Ministry of Health, donors, multilateral agencies, etc. The cost of the DQR ultimately depends on the level of effort required.

Resource implications for the desk review

It is recommended that the desk review be conducted with the support of an independent entity such as a national institute or consultant to help ensure unbiased evaluation of data quality. The desk review component of the DQR requires compilation of HMIS data for the relevant indicators in a specified format. This means obtaining data from the HMIS and/or programmes for the selected indicators. It is recommended that a national consultant or national institute should work with the Ministry of Health focal points to prepare the data for the selected core indicators.

In general, for the preparation of the data, a timeframe of about 1.5–2 weeks (8–10 person-days) is necessary, in addition to a further 1–1.5 weeks for the analysis and reporting. In total, about 20 person-days are required for the DQR. The level of effort may be more or less, depending on the number of indicators the country chooses to include in the assessment.

Resource implications for the facility assessment (data verification and system assessment)

The level of effort required for data verification depends on the number of facilities included within the health facility assessment or sample size, the number of indicators included in the data verification exercise, the volume and organization of the data at the health facilities, and the complexity of the reporting system. Given this complexity, it is recommended that data verifiers work in pairs.

Data verification and the system assessment at small facilities generally requires 3–4 hours for an assessment of 4–5 indicators. Larger facilities or hospitals will require more time as the volume of service provision and of records of service provision is greater. In general, for a sample of 100 health facilities, 10 data collection teams (with two persons in each) will take 8–10 working days, depending on the factors noted above. This amounts to 160–200 person-days. Depending on whether the data collection is

conducted using paper or electronic versions of the questionnaire (or both), several days may be required for data entry and checking prior to analysis.

It is recommended that the health facility survey component of the DQR should be conducted in conjunction with a larger health facility survey.[14] (This component is currently administered as one module of the Service Availability and Readiness Assessment, or SARA). Combination with an existing survey will greatly minimize the need to identify separate funds for the data verification. However, the health survey component of the DQR may also be administered as a stand-alone survey.

Step 4. Select core indicators and establish benchmarks

The DQR can be implemented as a holistic review across several programme areas or as an in-depth assessment of a particular programme area. The indicators selected should align with the purpose of the assessment and the intended use of the results.

The DQR coordinating group should oversee the selection of indicators and benchmarks. As a general rule, the recommended core indicators (antenatal care, immunization, HIV, TB, malaria) should be examined on an annual basis.

It is important to note that the same indicators that are selected for the desk review should also be selected for the data verification. Because of the time involved in data verification, it is recommended that no more than 4–5 indicators are selected for the data verification exercise.

Variations often exist between countries in the naming and definition of indicators, as well as in the services available. Some indicators may not be relevant or appropriate in some countries. Ultimately, the DQR coordinating team should determine what is appropriate, worthwhile and manageable in the country context.

Step 5. Identify the implementing agency and quality assurance

The DQR coordination committee needs to decide on the mechanisms for implementation of the DQR. As there are two components of the DQR (national-level desk review and health facility survey), the mechanisms for implementation should be considered for both.

In order to build technical capacities and ensure objectivity for a DQR, links should be forged with national statistics agencies, academic institutions and technical/development partners. A selection of an external agency or institution to support MOH in the implementation of the DQR or providing quality assurance will also ensure objectivity. The DQR should be conducted in a spirit of openness and transparency and should include regular feedback to data producers at the health-facility and district levels.

[14] Planning and budgeting for the SARA are provided in SARA reference documents at the following website: http://www.who.int/healthinfo/systems/sara_introduction/en/ (accessed 8 June 2015).

Step 6. Training requirements

The DQR requires advanced planning not only for the implementation of the desk review and the health facility-based data verification but also for training the various personnel who will take part in the process.

A training plan should be developed and budgeted as part of the overall DQR planning and budgeting. All personnel should be identified, recruited and trained well before the start of the DQR.

Training needs will differ according to the type of personnel and the tasks performed. These needs are outlined below along with the estimated number of training days required.

2.2 | Implementation

Preparation and implementation of the desk review

Data requirements

For the desk review, the data required for the selected indicators are monthly or quarterly data by subnational administrative area for the most recent reporting year and annual aggregated data for the last three reporting years.

Information on submitted aggregate reports and when they were received will be required to evaluate completeness and timeliness of reporting. Data on submitted reports for the three years prior to the year of analysis are required in order to evaluate the trend in reporting completeness. If information for all selected primary indicators is reported on the same forms (e.g. the HMIS form) these data will suffice for all selected indicators. If indicator values are reported on separate forms (e.g. programme-specific reporting forms), a separate analysis will be required for each set of reporting forms used for the selected indicators.

Other data needs include denominator data for calculating coverage rates for the selected indicators and survey results (and their standard errors) from the most recent population-based survey – such as the Multiple Indicator Cluster Survey (MICS), the Demographic and Health Surveys (DHS) and immunization coverage surveys. See Annex 3 for a more detailed description of the data and formatting requirements for the desk review.

Data collection

Data for the indicators are collected from the HMIS or health programmes, depending on which source is used most frequently for programme planning, monitoring and evaluation. As the purpose of the DQR is to evaluate the quality of routine health facility data , it is important to note that support from the leadership of the health ministry is essential to acquire and evaluate the quality of HMIS or programme data. It is necessary to work closely with programme, M&E and HMIS managers to analyse the quality of facility-reported data.

Data formatting and compilation

Once data are acquired they must be formatted to facilitate the analysis, typically by creating a "flat file" in which data for facilities and/or districts are contained in rows and where indicator values by month or year are in columns (e.g. Excel, CSV). Detailed guidance for formatting data for the DQR is available in Annex 3. The DQR analyses can be programmed into most software applications, and they can also be conducted on paper.

After the data are formatted appropriately, the analysis can proceed. Data quality metrics should be programmed into the software selected by the DQR coordinating team, with output to graphs (where appropriate), tables and lists. Metrics that identify administrative units with potential data quality issues should generate lists of those administrative units so that attempts can be made to understand the anomalies and, if possible, to correct the data. Indicators that compare trends should produce graphs showing the trend lines side-by-side for ease of comparison. Comparisons of administrative units should produce bar charts of performance by unit to facilitate understanding of the relative performance of these units.

However, for countries that are not inclined to invest in modifying existing software to accommodate the DQR methodology, an MS Excel-based tool is available from WHO for compiling and formatting data in layouts which facilitate data analysis. Once the data are inputted into the standardized tool, the data quality metrics are automatically calculated, as are graphic depictions of performance.[15]

Results from the health facility survey component of the DQR (data verification and system assessment results) should be integrated into the desk review analysis. Information on the accuracy of reporting for selected indicators will inform the confidence of policy-makers in the reported data. Information on weaknesses or gaps in the capacity of the reporting system can point to system strengthening activities.

Preparation and implementation of the health facility survey component

Requirements for data verification and system assessment

Lists of recommended source documents and cross-checks for data verification are available in Annex 4 and Annex 5.

Sampling health facilities

A representative sample of health facilities should be drawn for data verification and for administering the system assessment module. A "master facility list" – or a list of health facilities with attribute data (e.g. management authority, facility type, and location in terms of region and district) – is required and constitutes a prerequisite for implementation of the DQR. Once the objectives of the DQR are determined, the sampling methodology can be developed. For instance, health facility assessments such as the SARA typically employ list and/or area sampling, while data quality assessments have used a modified two-stage cluster sampling methodology. If regional estimates of data accuracy, or estimates specific to certain types of health facilities (e.g. management authority or type of facility) are required, the sampling methodology must take account of these requirements. Specialty services (e.g. TB diagnosis and treatment, HIV testing and treatment) are not offered at all facilities so the sample might need to be

[15] In addition, work is underway to incorporate many of these data quality metrics into the DHIS software and to ensure that data can be output easily in the required format for analysis in Excel for DHIS users.

adjusted if indicators from these programme areas are to be assessed. The technical requirements of drawing the sample and deriving estimates from the resulting data are not trivial. Care should be taken when developing the sampling methodology according to individual country requirements. A statistician should be consulted to ensure that the sample is drawn appropriately. Annex 6 provides more information on sampling of health facilities for the DQR.

Identify, adapt and reproduce survey tools (paper and/or electronic)

Standardized tools have been developed for data verification and the system assessment to assist countries in implementing the DQR at health facility and district levels. The tools were developed as modules of the SARA toolkit but can be employed as stand-alone tools when data quality assessment is the primary purpose.

The tools should be adapted to the country context prior to implementation (e.g. by specifying programme areas, indicators and source documents). If data are to be captured electronically (e.g. on a tablet computer) a database should be developed to facilitate data entry. Sampled health facilities should be prepopulated in the database, and facility database records should be made available on the tablets used in the field. Data verification and system assessment modules have been developed in CSPro 5.0 and can be obtained from WHO. These modules are designed to work in conjunction with the CSPro SARA database but can be made to function independently with additional programming. As with the paper version of the survey tools, the database modules should be adapted to the country context prior to implementation of the DQR. Paper data collection tools are available in Annex 7.

Organize training of fieldworkers (enumerators)

Fieldworkers conducting the health facility survey should be trained in the methods of data verification and in administration of the system assessment. Data verification across programme areas requires familiarity with a variety of data collection tools (registers, patient records, tally sheets, etc.) according to the indicators and programme areas. Enumerators should ideally have experience of recording public health data and exposure to the data collection tools used in the field. Training of enumerators should include practice in compiling indicators for each programme area using the tools they are likely to encounter in the field.

Notify sites and subnational authorities

Several weeks prior to implementation, the health facilities sampled for the DQR should be notified of the impending visit by the enumerators. The relevant data management staff at the selected health facilities will need to be present at the facility on the day of the assessment in order to facilitate access to the appropriate records, provide responses for the system assessment, and assist with the completion of the survey at the facility. These staff and their supervisors should be informed of the survey and the date of the visit to ensure they are present at the facility on the day of the visit. Similarly, subnational HMIS management authorities, such as HMIS managers at the district and/or regional levels, should also be informed both to satisfy potential administrative protocols and to enlist their support/cooperation in completing the survey.

Conducting the survey at the health facility

Survey teams should work in pairs to maximize efficiency and to control quality during visits to health facilities. The teams should plan to spend one complete day on each health facility assessment with data verification/system assessment, and at least half a day for a stand-alone DQR. Up to five indicators (one per programme area) are recommended for data verification; it can take considerable time to complete the survey, depending on the volume of service for the indicators (the number or records to recount), and the quality and organization of the data (ease of retrieval and recount). The system assessment should require no more than one hour at the health facility. The ideal respondent for the system assessment is the facility data manager (or the person responsible for compiling and reporting the data).

Conducting the survey at the district level

The DQR is also implemented at district HMIS management units in the districts involved in the data flow from sampled health facilities. At the district level the survey team will re-aggregate the district value of the selected indicators using the values submitted on the monthly reporting forms from all facilities in the district (not just the facilities in the sample). The team will also determine the completeness and timeliness of reporting at this level. The district-level system assessment module should be completed in an interview with the data manager or programme manager. Survey teams should plan to spend about half a day at the district HMIS management unit.

Oversight and quality control of the survey

Survey teams should be supervised in the field by dedicated staff. Supervisors should cover a predetermined geographical area and a specified number of survey teams. The supervisor's role is to assist the teams in the completion of the surveys (where necessary), to collect and review the completed questionnaires and to troubleshoot problems as they arise. Supervisors should revisit health facilities and verify the survey results for a small sample of facilities (e.g. 10%) to ensure accurate recording of results. If possible, independent monitors from national stakeholders (e.g. donors) can also play a role in monitoring implementation of the survey.

Compiling results

Survey team supervisors should bring the completed surveys to the designated DQR data management staff at national level. A small team should be assembled from available staff at the Ministry of Health and/or at stakeholder organizations to review submitted surveys, correct errors and enter the data into the computer program (e.g. CSPro 5.0) to facilitate analysis of the data. Depending on the number of facilities sampled and the number of indicators verified, it may take up to one week for team of 4–5 data managers to clean and input all the data.

Part 3: Analysis, dissemination and use

3.1 | Data quality analysis and interpretation

Data quality analysis is typically conducted on a computer using the standard DQR Excel-based tool (Annex 8). However, the analysis can also be carried out with a local application customized to conduct the DQR analyses.

Results should be presented in tables and graphs with ample space to add interpretation of results. Staff who understand the dynamics of service delivery in the year of analysis (i.e. programme managers) should participate in the interpretation of DQR results.

Desk review analytical outputs

Illustrations of output for the DQR analysis are presented for each dimension of data quality.

Completeness of reporting

Figure 3.1 shows the national-level results of analyses for completeness of district data (Indicator 1a), as well as the number and percentage of districts failing to meet the standard.

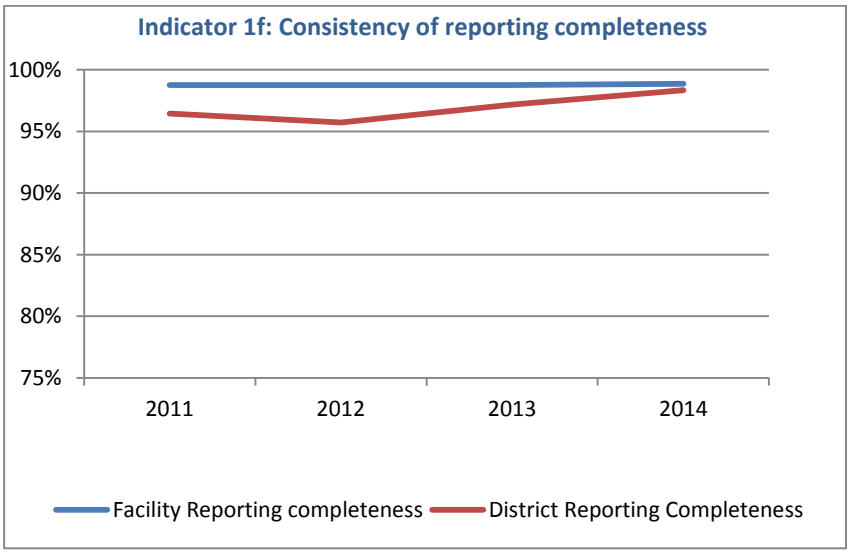
Figure 3.1 Example of dashboard results from the DQR for completeness of reporting

Indicator 1a: National district reporting completeness rate and districts with poor completeness rate

	2014
National district reporting completeness rate	98.3%
Number of districts with completeness rate below 75%	4
Percent of districts with completeness rate below 75%	5.6%
Districts with reporting completeness rate below 75%	District 1, District 3, District 7, District 10

Interpretation of results: Indicator 1a

- Good reporting completness continuing a trend upwards from recent years.
- Investigate districts with < 75% completeness for the year.
- Districts 1 and 3 had stock out of reporting forms during 2nd quarter of last year



Interpretation of results: Indicator 1f

- Overall consistent high levels of reporting completeness at facility and district levels
- District reporting completeness was lower than facility reporting completeness in previous years but has equalized in 2014

Internal consistency of reporting data

Outliers

Figure 3.2 displays results for the identification of outliers (extreme) relative to the mean of reported values for six indicators at national and subnational levels. Outliers relative to the median of reported values are also identified in the DQR.

Figure 3.2 Example of dashboard results from the DQR for identification of outliers

Indicator 2a: Identification of outliers

Indicator 2a.1: Extreme outliers (>3 SD from the mean)				2014
	National score	Districts with extreme outliers relative to the mean		
Program Area and indicator	%	No.	%	Name
Maternal Health - ANC 1st Visit	0.2%	2	2.5%	District 2, District 7
Immunization - 3rd dose DPT-containing vaccine	0.1%	1	1.3%	District 3
HIV/AIDS - Number of HIV+ persons currently on ART	0.0%			-
TB - Number of Notified TB cases (all forms of TB)	0.0%			-
Malaria - Number of confirmed malaria cases reported	0.0%			-
General service statistics (OPD total visits)	0.4%	4	5.0%	District 2, District 7, District 9, District 10
Total (all indicators combined)	0.1%			

Interpretation of results - Indicator 2a1

- Good results given the volume of data
- ANC outlier in district 2 looks like a data entry error - value is 10 x greater than other monthly values reported by the district last year - call district health information officer to investigate
- Values in OPD could be the result of social marketing campaign conducted last year to improve health care utilization - call

Outliers are indicative of problems in data quality or changes in service delivery patterns, or both. Some indicators (e.g. immunization) are expected to show variability, while others are not. The overall percentage of values that are identified as outliers is shown, as well as the number and percentage of districts with extreme values. Space is provided to record interpretation of the results.

Consistency over time

Within the context of internal consistency, consistency over time is evaluated in order to examine the current year's value against values reported in previous years. Depending on the expected trend of the indicator (constant, or increasing/deceasing) the current year's value is compared to the average of the values of the three previous years (constant trend) or the value forecast[16] from the values from the preceding years (nonconstant). The resulting ratio is calculated for each subnational administrative area and is compared to the ratio at national level. Subnational administrative areas that exceed a user-defined threshold for quality are identified for further investigation.

In Figure 3.3, the graph (which is an example) shows service output for outpatient days/visits (OPD) in the current year for each subnational administrative area compared to the mean for OPD in the three preceding years for the same administrative area. Dotted lines represent the recommended or user-defined quality threshold while the solid line indicates the national-level relationship between OPD in the current year and the average for the preceding years. Values for subnational administrative areas that exceed the threshold of quality would appear above or below the dotted lines. These areas are identified and are investigated to identify potential data quality problems.

[16] The forecasted value calculates what the current year should be based on the slope of the line of 3 previous years (the trend) and compares the calculated value to the real value.

Figure 3.3a shows a comparison of the current year value for OPD to the value forecasted from the preceding three years of values. (The graph at bottom-left indicates the actual trend in the indicator.) Subnational units are compared to the expected value (it is expected that, if the trend in reporting holds, the current year value will be the same, or similar, to the forecasted value for each subnational unit). Five subnational units have ratios of the current year to forecasted value ≥ 1 + 0.2 (or 20%, the specified quality threshold) or ≤ 1–0.2.

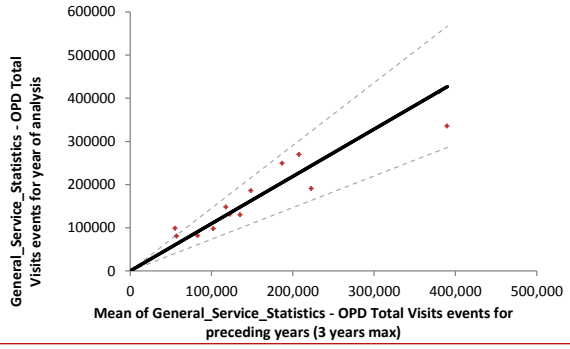
Figure 3.3 Example of dashboard results from the DQR for consistency over time – constant trend for the indicator

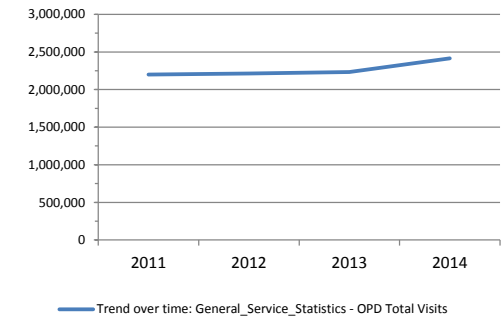
2b2: Consistency of 'General Service Statistics - OPD Total Visits' over time

Year	2014
Expected trend	Constant
Compare districts to:	national result
Quality threshold	33%
National score (%)	109%
Number of districts with divergent scores	1
Percent of districts with divergent scores	8%

Names of districts with divergent scores:

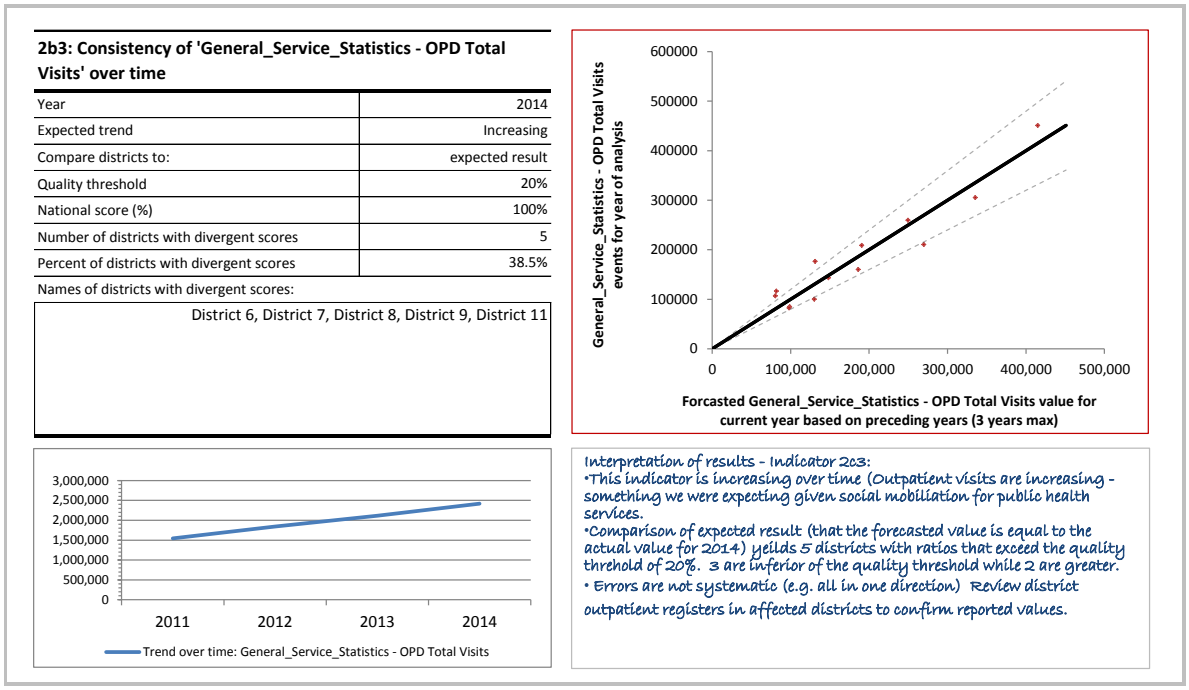
District 3





Interpretation of results - Indicator 2c2:
•Overall the national ratio was 1.09 which means that the OPD visits for the current year is 9% greater than the mean of OPD visits for the past three years.
•There are 6 districts which have seem to have the mean OPD visits of the past three years that are higher than the OPD visits for the current year. Need to determine why OPD visits are lower in the current year.
• Except for 1 district, all districts had a ratio similar to the national ratio (within 33% of the national ratio).
• The district that was outside of the quality threshold had a ratio where the mean OPD visits of the past three years were higher than the OPD visit of the current year. Need to determine if this is a data quality issue or a program issue. For data quality, need to check if all the OPD data for this district has been entered. Are the visits for the past years correct? See if previous data quality checks have been done. If this is not a data quality issue need to understand potential programmatic issues that could be causing this discrepancy.

Figure 3.3a Example of dashboard results from the DQR for consistency over time – increasing trend for the indicator



2b3: Consistency of 'General_Service_Statistics - OPD Total Visits' over time

Year	2014
Expected trend	Increasing
Compare districts to:	expected result
Quality threshold	20%
National score (%)	100%
Number of districts with divergent scores	5
Percent of districts with divergent scores	38.5%

Names of districts with divergent scores:

District 6, District 7, District 8, District 9, District 11

Interpretation of results - Indicator 2c3:
•This indicator is increasing over time (Outpatient visits are increasing - something we were expecting given social mobiliation for public health services.
•Comparison of expected result (that the forecasted value is equal to the actual value for 2014) yeilds 5 districts with ratios that exceed the quality threhold of 20%. 3 are inferior of the quality threshold while 2 are greater.
• Errors are not systematic (e.g. all in one direction) Review district outpatient registers in affected districts to confirm reported values.

Consistency between indicators

Additionally within Dimension 2, the consistency between related indicators is evaluated. In the example in Figure 3.4, the first antenatal care visit (ANC1) is compared to the first dose of intermittent preventive therapy (IPT1). In malaria endemic countries, IPT should be given to all pregnant women as a part of ANC. In theory, the number of women given the first dose of IPT should be equal to the number of women attending ANC for the first time. The ratio of ANC1 to IPT1 is calculated both for all subnational administrative areas and for the national level. In the example in Figure 3.4 the value at national level is 114%, which means that more women began ANC than received IPT. The subnational units with ratios over 10% greater (or 10% less) than 1 (i.e. ANC1 and IPT1 are equal) are flagged for investigation.

Figure 3.4 Example of dashboard results from the DQR for internal consistency between related indicators

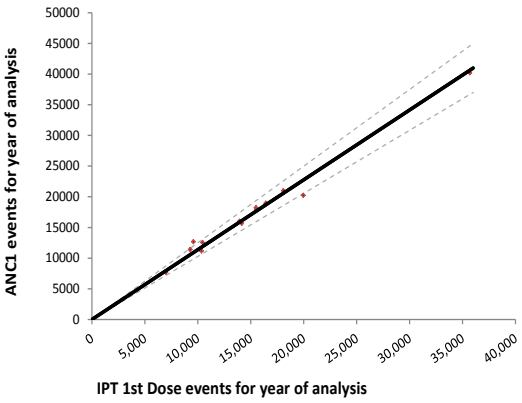
Indicator 2c: Internal consistency - consistency between related indicators

Consistency between related indicators - Ratio of two related indicators and districts with ratios significantly different from the national ratio *

2c1: Maternal Health Comparison: ANC 1st Visit : IPT 1st Dose	
Year	2014
Expected relationship	equal
Compare districts with:	national rate
Quality threshold	10%
National score (%)	114%
Number of districts with divergent scores	2
Percent of districts with divergent scores	15.4%

Names of districts with divergent scores:

District 5, District 6



Interpretation of results - Indicator 2c1:

- Data seem pretty good - only district 5 has a largely discrepant value
- IPT seens consistently lower than ANC1 - more pregnant women should be receiving IPT
- Stock out of fansidar in Region 2 could explain low number of IPT in Districts 5 . Call DHIO in these districts to investigate
- National rate is 114% - most districts are close to this value. District 6 is performing well relative to the other districts but is 'discrepant' relative to the national rate. - no follow up needed.

External comparison with other data sources

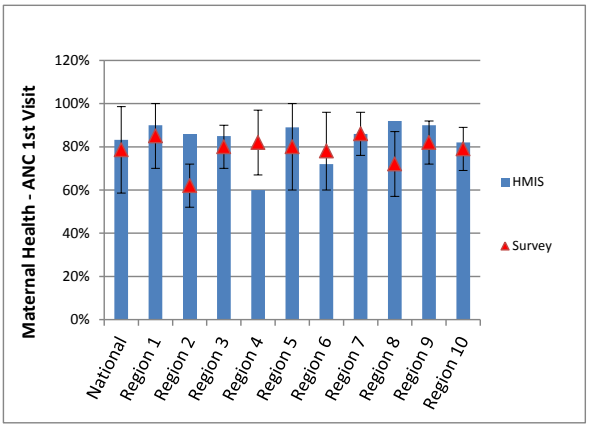
Figure 3.5 depicts results from an external comparison of HMIS data for ANC1 and the survey value of ANC coverage for the appropriate year. Vertical bars represent the ANC coverage from HMIS (annual ANC values aggregated across the appropriate subnational administrative areas over the appropriate target population for the area). The triangles represent the analogous survey indicator for the same subnational administrative areas, with error bars based on the standard error of the estimate. In Figure 3.5, Region 2 and Region 8 have survey values (and ranges of error) below the HMIS ANC coverage, indicating potential overreporting of ANC by HMIS. In Region 4, the HMIS estimate is less than the survey estimate, indicating potential underreporting of ANC by HMIS.

Comparison of different data sources

Figure 3.5 External comparison of ANC1 derived from HMIS with survey values

Indicator 3a: Comparison of routine data with population-based survey values from the same period

3a1: 'ANC 1st Visit' consistency ratio (ratio between the facility rates and survey rates)	
Year	2014
Quality Threshold	33%
National Score (%)	106%
Number of Regions with divergent scores	3
Percent of Regions with divergent scores	30.0%
Names of Regions with divergent scores:	
Region 2, Region 4, Region 8	



Interpretation of results - Indicator 3a1:
- ANC HMIS value in region 4 looks too low - could result from missing source documents or a failure to record service delivery. Review report forms from districts in the region to verify the reported values.
- ANC HMIS value in regions 2 and 8 seems too high - could be double counting or duplicate reporting. Call District Health Information Officers to investigate.
-

Quality of population data

Figure 3.6a shows an example of the dashboard for results of comparisons of population data. Indicator 4a shows the comparison of National Statistics Office values for live births with the United Nations population estimate for live births. The value of 0.98 indicates that the National Statistics Office value is lower than the United Nations estimate, but only slightly.

Figure 3.6a Example of dashboard results for the quality of population data

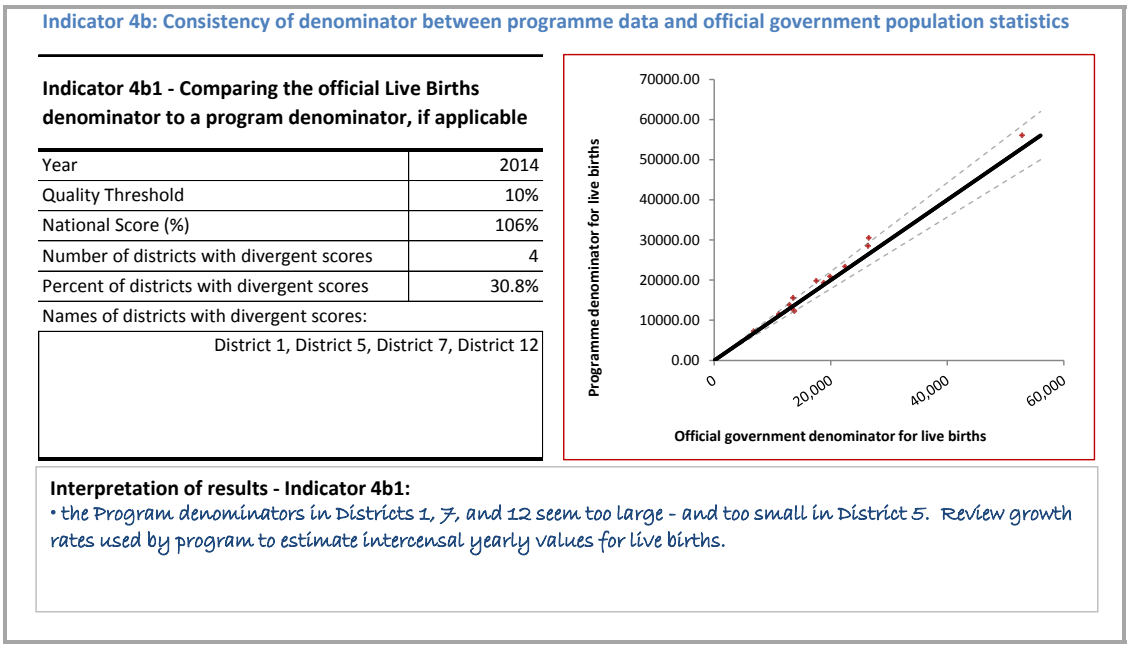
Indicator 4a: Consistency with UN population projection

	2014
Ratio of population projection of live births from the Bureau of Statistics to a UN live births projection	0.98

Interpretation of results - Indicator 4a:
- Good agreement between official government estimate of live births to the UN estimate. Discrepancy could be related to growth rate used to calculate intercensal years.

Indicator 4b (below) shows the dashboard for the comparison between a health programme estimate of live births and the official government value (of the National Statistics Office). Subnational administrative areas are also evaluated (depending on the availability of the data) and discrepant subnational units are identified. In the example below, District 1, District 7 and District 12 have have programme values for live births greater than the official government values for the same regions. District 5 has programme value for live births below the official government value for it.

Figure 3.6b Example of dashboard results for comparison of estimates of live births between two sources of data



Data verification analysis

The analysis of the data verification has several components:

1. *Determination of the number of facilities included in the subsample for the analysis of each indicator (since not all services are offered at all health facilities)*. The higher the number of facilities included in the sample, the higher the confidence in the data verification factor for the specific indicator.
2. *Calculation of the extent of missing documents (registers, tally sheets, monthly reports)*. A large percentage of missing documents will reduce the value of the verification factor. The extent to which the verification factor is affected by missing documents should be quantified if possible. To estimate the accuracy of reporting adequately it is necessary to have most, if not all, source documents and monthly reports available for review at health facilities.
3. *Calculation of the percentage of data points that are an exact match between the source document and the monthly reported value*. The total number of service outputs in the source document should ideally be the same as the number recorded in the monthly report. However, there are often mismatches between the two. It is important to determine the extent of matches or mismatches between source documents and monthly reports and the reasons behind discrepancies – such as data entry error, arithmetical error, partial inclusion of source documents, etc. Large mismatches will, of course, affect the calculation of the verification factor, but it is important to identify smaller mismatches to determine why there is discordance between source documents and monthly reports.

   When analysing (mis)matches between source documents and monthly reports, it is also important to

identify the percentage of facilities where source documents indicate a higher number of service outputs than are recorded in monthly reports, and vice-versa.

4. *Calculation of the verification factor*. Below are some examples of tables that should be included in the report. In additional to the verification factor, the information included in these tables will allow for a more nuanced interpretation of the data verification results.

Availability of services

The percentage of facilities in the sample providing the specific health services should be included in the presentation of results. This will provide information on the number of facilities on which the subsequent data verification results are based. Table 3.1 shows one example of how the data may be presented.

Table 3.1 Percentage of facilities in the sample providing each health service, by stratum

		Facility type		Stratum 2		
	Overall	Hospitals	Health centres			
ANC	%	%	%	%	%	%
EPI	%	%	%	%	%	%
Malaria	%	%	%	%	%	%
TB	%	%	%	%	%	%
HIV	%	%	%	%	%	%

Note: ANC = antenatal care; EPI = Expanded Programme on Immunization; TB = tuberculosis; HIV = human immunodeficiency virus.

Availability of documents for review

If a facility offers a particular service, it must also have the source documents (registers, tally sheets, etc.) of the three-month verification period available for review on the day of the data verification survey. The selected programme indicators (and their related services) should have standard Ministry of Health registers, tally sheets or other documents which health facilities are supposed to use to record daily activities. While it is possible that health facilities may use multiple documents to record the services provided, it is important to identify if there is a main source document that is used for monthly reporting. The analysis should include information on which documents are used as the main source for reporting service outputs and which of these documents are available. Table 3.2 shows the distribution of the main source document used by a facility for its monthly reports.

Table 3.2 Distribution of main document used by a facility to report service data for monthly reports, by programme area

	ANC	EPI	Malaria	TB	HIV
Register	N (%)	N (%)	N (%)	N (%)	N (%)
Patient cards	N (%)	N (%)	N (%)	N (%)	N (%)
Tally sheet	N (%)	N (%)	N (%)	N (%)	N (%)
Improvised document	N (%)	N (%)	N (%)	N (%)	N (%)
Others	N (%)	N (%)	N (%)	N (%)	N (%)
Total	N (100%)	N (100%)	N (100%)	N (100%)	N (100%)

Note: ANC = antenatal care; EPI = Expanded Programme on Immunization; TB = tuberculosis; HIV = human immunodeficiency virus.

Table 3.3 shows the percentage availability of these documents for all the three months.

% availability of source documents for each facility[17] = $$\frac{\sum_{i=1}^{n} Available\ month\ 1_i + Available\ month\ 2_i + Available\ month\ 3_i}{3n}\ X\ 100$$

where *n* is the total number of facilities using that particular source document

Table 3.3 Availability of main document used by a facility to report service data for monthly reports, by programme area

	ANC	EPI	Malaria	TB	HIV
Register	%	%	%	%	%
Patient cards	%	%	%	%	%
Tally sheet	%	%	%	%	%
Improvised document	%	%	%	%	%
Others	%	%	%	%	%

Note: ANC = antenatal care; EPI = Expanded Programme on Immunization; TB = tuberculosis; HIV = human immunodeficiency virus.

Match between source documents and monthly reports

The number of events recounted from the register should exactly match the number reported in the monthly reporting form. Table 3.4 shows the percentage match between the reported documents and recounted documents for all the three months.

% match between reported and recounted service outputs =

$$\frac{\sum_{i=1}^{n} (Report\ month\ 1_i = Recount\ month\ 1_i) + (Report\ month\ 2_i = Recount\ month\ 2_i) + (Report\ month\ 3_i = Recount\ month\ 3_i)}{3n}\ X\ 100$$

where *n* is the total number of facilities using that particular source document

[17] Except in the case of TB which has quarterly reporting rather than monthly reporting. For TB the available source documents will be for one quarter-year instead of for three months.

Table 3.4 Percentage of facilities that have an exact match between recounted numbers of main source documents and monthly reports, by facility type

		Stratum 1	
	Overall	Hospitals	Health centres
ANC2	%	%	%
Deliveries	%	%	%
Measles immunization	%	%	%
Outpatient visits	%	%	%
Malaria cases (outpatient)	%	%	%
Average	%	%	%

Note: ANC = antenatal care.

Verification factor

Even if the reported and recounted numbers do not match exactly , it is useful to take into account the degree of disparity between the two.

For a given indicator, the verification ratio at a facility is computed as the recounted number of events from source documents divided by the reported number of events from the HMIS.

$$\textit{Verification factor} = \frac{\text{Recounted number of events from source documents}}{\text{Reported number of events from the HMIS}}$$

A verification ratio higher than 1 implies that there is underreporting of events in the HMIS for the verification period. If the verification ratio is less than 1, this implies that there is overreporting of events in the HMIS for the period chosen for the analyses (alternatively, this may be due to missing documents as the services may have been rendered but cannot be verified). For each health facility in the sample, a verification ratio is computed for each of the core data verification indicators. If a particular service is not offered at the facility, the verification ratio cannot be calculated for the corresponding indicator. To obtain a national verification factor for an indicator, a weighted mean is calculated from individual health facility verification ratios.

The verification factor is averaged across the sample domains of estimation and is then weighted using previously reported “population” values (facility values from the HMIS aggregated across domains of estimation from all facilities). Frequently the facility values used to derive the weights are not available at the national level since facility values typically are not reported beyond the district in paper-based systems (still a majority of developing countries). Alternative methods of deriving weights for domain-specific estimates of accuracy are presented in Annex 6. However, these methods may be deemed impractical or impossible to derive, in which case the unweighted estimates may be presented with the understanding that precision may be lost in the estimate. Ultimately, the DQR coordinating team must determine what is possible and acceptable given the objectives of the DQR and the country situation.

The weighted estimates of accuracy for the assessed indicators should be compared to previous DQR (or other data quality assessments) results to determine trends in accuracy. Comparisons should also be made between subnational units to determine where resources should be targeted for system strengthening.

Verification factor for the district

The verification factor at the district level is calculated by re-aggregating the value of the selected indicators from the health facilities reporting to the district on monthly summary report forms. The re-aggregated value is divided by the value reported by the district for the reporting period in question to derive a district verification factor. The district verification factor is an independent assessment of the accuracy of reporting for the district HMIS or programme office. The district verification factor is not factored into the composite verification factor derived from the full sample of health facilities.

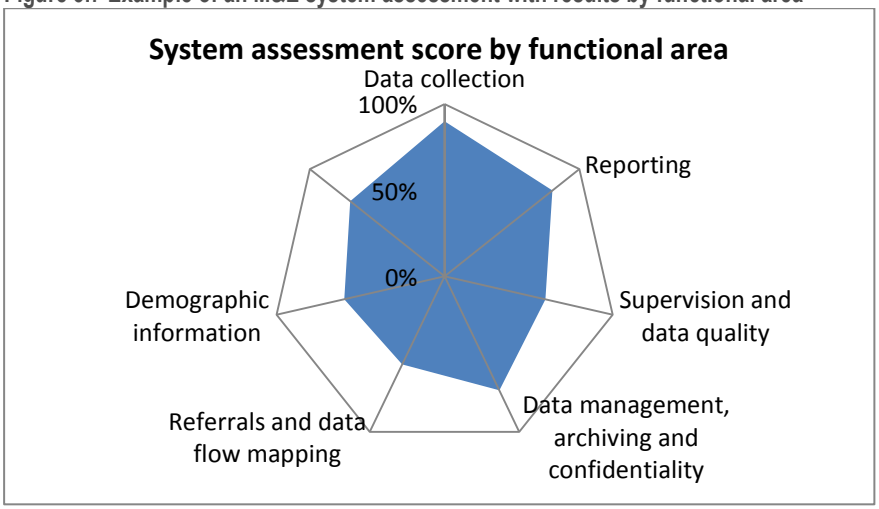
If a two-stage cluster sampling method is employed to select health facilities, the cluster (usually districts) specific verification factor is weighted on the volume of service in the cluster. An adjustment factor is applied to each cluster – i.e. the ratio of the district value found in the district office and the value for the district found at national level. A weighted average of the adjusted cluster-specific verification factors is then calculated to obtain the national-level estimate of accuracy based on the sample.

System assessment analysis

The analysis of the system assessment indicators involves calculating the percentage of sampled health facilities in each response category (e.g. “yes, completely”, “partly”, etc.) for each of the metrics of the system assessment (disaggregated by stratification variables used in the sampling) – e.g. the percentage of health facilities that were scored “yes, completely” for “Staff responsible for data collection and compilation of reports have received the appropriate training”. Additionally, scores are aggregated by functional area to determine an overall score for each functional area. Again, functional area scores are aggregated across health facilities within the strata defined by the sampling methodology such that each stratum has specific scores for each functional area and metric. Results can be displayed graphically using bar charts or radar graphs (as shown in Figure 3.7).

As the metrics used in the system assessment are proportions, much like those of other indicators in a health facility assessment (e.g. proportion of facilities that have a clean water source), system assessment metrics should be analysed in the same way as indicators in the larger health facility assessment (when the system assessment is conducted as a component of a health facility assessment).

Figure 3.7 Example of an M&E system assessment with results by functional area



In addition, responses can be colour-coded and plotted by reporting facility or district in order to determine “hotspots” – i.e. areas of concern for the reporting system as characterized by the colour assigned to poor scores in the system assessment. In Figure 3.8 below presents an example of the system assessment results for two functional areas. Metrics are arranged in columns while facilities by district are shown in rows. The responses indicating weakness of system assessment metrics are coloured red. It is evident from the example that weaknesses are systematic for certain functional areas and metrics. See Annex 5 for a detailed list of system assessment indicators by functional area.

Each of the data quality dimensions described above should be analysed and interpreted by data managers and programme managers familiar with the health programmes and indicators reviewed. What looks at first glance like a data quality problem may actually be an unusual pattern of service delivery. Only staff with a good knowledge of programme activities and output throughout the year can make such distinctions and should be involved in the analysis and interpretation of the DQR results.

Interpretations of results should be recorded in the final report, together with recommended actions for remediation. Using these results, a data quality improvement plan can then be prepared.

Figure 3.8 Visual representation of system assessment results – "hotspots"

	I - M&E Structure, Functions and Capabilities						II- Indicator Definitions and Reporting Guidelines					
Service Delivery Site	Responsibility for recording service delivery on source document is clearly assigned to relevant staff	There is process in place to ensure data compilation and reporting is completed even when responsible staff is not available to do the job	There are designated staff responsible for reviewing reports prior to submission	The health facility receives regular feedback on the quality of submitted reports according to the guidelines.	The health facility receives regular supervisory visits from district and/or national staff according to guidelines.	…If yes, the last visit was within the past three months.	The National M&E Office has provided written guidelines to each sub-reporting level on what should be recorded in source document.	,,, what should be included on the monthly report.	… how (e.g., in what specific format) reports are to be submitted.	… to whom the reports should be submitted.	… when the reports are due.	The written instructions provided by the Program are adequate to ensure standardized recording and reporting of program data.
Service Point 1	Not at all	Not at all	Partly	Not at all	Yes - completely	Not at all	Yes - completely	Not at all	Not at all	Partly	Yes - completely	Not at all
Service Point 2	Partly	Not at all	Partly	Not at all	Yes - completely	Not at all	Yes - completely	Not at all	Not at all	Partly	Yes - completely	Partly
Service Point 3	Partly	Partly	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Partly	Partly	Yes - completely	Partly
Service Point 4	Partly	Not at all	Not at all	Not at all	Yes - completely	Not at all	Yes - completely	Not at all	Partly	Not at all	Yes - completely	Not at all
Service Point 5	Partly	Not at all	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Not at all	Partly	Yes - completely	Partly
Service Point 6	Partly	Yes - completely	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Yes - completely	Partly	Yes - completely	Partly
Service Point 7	Partly	Partly	Partly	Not at all	Yes - completely	Partly	Yes - completely	Partly	Partly	Partly	Partly	Partly
Service Point 8	Partly	Yes - completely	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Yes - completely	Partly	Yes - completely	Not at all
Service Point 9	Not at all	Not at all	Partly	Not at all	Yes - completely	Not at all	Yes - completely	Not at all	Not at all	Partly	Yes - completely	Not at all
Service Point 10	Partly	Not at all	Partly	Not at all	Yes - completely	Not at all	Yes - completely	Not at all	Not at all	Partly	Yes - completely	Partly
Service Point 11	Partly	Partly	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Partly	Partly	Yes - completely	Partly
Service Point 12	Partly	Not at all	Not at all	Not at all	Yes - completely	Not at all	Yes - completely	Not at all	Partly	Not at all	Yes - completely	Not at all
Service Point 13	Partly	Not at all	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Not at all	Partly	Yes - completely	Partly
Service Point 14	Partly	Yes - completely	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Yes - completely	Partly	Yes - completely	Partly
Service Point 15	Partly	Partly	Partly	Not at all	Yes - completely	Partly	Yes - completely	Partly	Partly	Partly	Partly	Partly
Service Point 16	Partly	Yes - completely	Partly	Not at all	Yes - completely	Yes - completely	Yes - completely	Partly	Yes - completely	Partly	Yes - completely	Not at all
yes -completely	0	4	0	0	16	8	16	0	4	0	14	0
Partly	14	4	14	0	0	2	0	10	6	14	2	10
Not at all	2	8	2	16	0	6	0	6	6	2	0	6
N/A	0	0	0	0	0	0	0	0	0	0	0	0
Total	16	16	16	16	16	16	16	16	16	16	16	16
% yes	**0%**	**25%**	**0%**	**0%**	**100%**	**50%**	**100%**	**0%**	**25%**	**0%**	**88%**	**0%**
% partly	**88%**	**25%**	**88%**	**0%**	**0%**	**13%**	**0%**	**63%**	**38%**	**88%**	**13%**	**63%**
% no	**13%**	**50%**	**13%**	**100%**	**0%**	**38%**	**0%**	**38%**	**38%**	**13%**	**0%**	**38%**
%N/A	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**	**0%**

3.2 | Dissemination and use of the DQR results

A report presenting the findings of the DQR should be prepared along with interpretation by programme managers and recommendations for system strengthening. The report should be disseminated to all staff expected to participate in health-sector planning initiatives (e.g. health sector review) several weeks prior to the planning event. Other stakeholders – such as donors, technical assistance organizations, relevant national and international NGOs, private-sector bodies (e.g. universities, civil society organizations), and concerned ministries – should receive copies of the report.[18]

On the basis of the findings of the DQR, the coordination team should lead the development of the Data Quality Improvement Plan and should ensure that all relevant internal stakeholders (both public, such as the Ministry of Health and the National Statistics Office, and private, such as public health institutes) and external stakeholders (bilateral and multilateral donors and technical assistance agencies, such as WHO) are involved.

A separate document on remedial measures to improve the quality of data should also be prepared. The Data Quality Improvement Plan should identify the data quality concern and the measure needed to strengthen the system and resolve the problem. The plan should include a responsible organization with appropriate staff, a timeline for implementation, and identified resources to ensure completion of the necessary measures. If resources for system strengthening are not available through the current budget, the DQR coordinating team should carry out advocacy among the donor community to raise the necessary resources. Measures for system strengthening should be prioritized so that measures with the highest likelihood of success, and those making the greatest impact on overall data quality, should be implemented first. See Table 3.5 for a sample Data Quality Improvement Plan.

The Data Quality Improvement Plan should seek to identify and address the root causes of data quality problems revealed by the DQR. The actions outlined in the plan should be specific, time-bound and costed. The agency or entity responsible for implementation should be identified. The DQR coordination team is responsible for regular monitoring of the improvement plan to ensure its implementation.

Data quality concerns should be categorized by functional area of the reporting system (e.g. data collection tools/reporting forms, use of data for decision-making, demographic information, etc.) and prioritized for resolution. Simple, low-cost solutions may be available for some problems, while others may be more costly and/or time-consuming. Remedial measures that have a large effect on improving data quality but are relatively less costly should be prioritized. Adding a data quality check to supervisory visits to health facilities is an example of a low-cost intervention that could produce large gains in improved data quality. Upgrading computers at the district level is an example of a high-cost measure.

Sometimes the solution to data quality problems is simple but prohibitively costly. Ensuring a stable supply of updated blank source documents goes a long way towards improving data quality but may be beyond the available budget of the country. For instance, facility registers are costly to produce so they are printed only every five years or so. Economies of scale are gained by printing several years' worth of registers at one time, so the country's budget may not extend to printing the registers more often. However, indicators collected on these registers often change more frequently than every five years. The point here is that recommendations for system strengthening should be manageable within the constraints that exist in every country.

[18] It is recommended that the report includes the explanation of statistical methods used for the calculation of the verification factor.

Table 3.5 Example of a Data Quality Improvement Plan

Data quality finding	Evidence of finding (interpretation)	Remedial measures	Scope	Timeline	Responsible	Resources
Domain: Indicator definitions and reporting guidelines						
Lack of understanding of indicator compilation techniques at health-facility level for PMTCT/HCT - Pregnant women are not disaggregated from HCT results	Systematic over-counting of HCT indicator values in some districts (as revealed by data verification)	Improved supervision and mentoring in affected districts Emphasis on indicator compilation during pre-service and in-service training -Ensure that printed copies of indicator definitions and compilation procedures are available in health facilities	Regions 2, 7, 10	One year (2015), then re-evaluate	-District health information officers or their designates (whoever is conducting supervision at the facility) -Pre-service, in-service curriculum design team (HMIS unit at national level)	District health information budgets -HMIS training budget (2015 allocation) -MOH nurse training (2015 budget) Global Fund Round 9 HSS grant
Domain: Data maintenance and confidentiality						
Source documents are not available for data verification	A significant proportion of service delivery for malaria could not be verified because of the non-availability of source documents -poor record-keeping/archiving of reported results	-Districts should work with affected health facilities to develop sound storage areas (closet or cabinet with locking mechanism in a cool, dry place) -shelves should be built using locally-available materials	Identified health facilities in Region 2 (districts 4 and 6) and Region 9 (districts 27 and 34).	2015, then re-evaluate	District health management teams; facility in charge; Regional Health Authority (facilities management unit)	2015 Facilities Management Budget - Global Fund Round 9 HSS grant

Note: HCT = HIV counseling and testing; HMIS = health management information system; HSS = health system strengthening; MOH = Ministry of Health; PMTCT = Prevention of mother-to-child transmission .

Annexes

Annex 1 | Recommended indicators

Core indicators

Recommended core of "tracer" indicators for DQR		
Programme area	**Indicator name**	**Full indicator**
Maternal health	Antenatal care 1st visit (ANC1)	Number (%) of pregnant women attended, at least once during their pregnancy
Immunization	DTP3/Penta3	Number (%) of children < 1 year receiving three doses of DTP/Penta vaccine
HIV/AIDS	ART coverage	Number and % of people living with HIV who are currently receiving ART
TB	Notified cases of all forms of TB	Number (%) of all forms of TB cases (i.e. bacteriologically confirmed plus clinically diagnosed) reported to the national health authority in the past year (new and relapse)
Malaria	Confirmed malaria cases[1]	Number (%) of all suspected malaria cases that were confirmed by microscopy or RDT

Note: ANC = antenatal care; ART = antiretroviral therapy; DTP3 = diphtheria-tetanus-pertussis three-dose vaccine; Penta = pentavalent vaccine; RDT = rapid diagnostic test.

Additional indicators

Additional DQR indicators		
Programme area	**Indicator name**	**Full indicator**
General	Total outpatient visits (OPD)	Number of outpatient visits per 10 000 population
Maternal Health	Antenatal care 4th visit (ANC4)	Number (%) of women aged 15–49 years with a live birth in a given time period who received antenatal care four or more times
	Institutional delivery	Number (%) of women who delivered in a health facility
	Postpartum care coverage	Number (%) of mothers and babies who received postpartum care within two days of childbirth (regardless of place of delivery)
	Tetanus toxoid 1st dose	Number (%) of pregnant women who received the 1st dose of tetanus-toxoid vaccine
Immunization	DTP1-3/Penta1-3	Number (%) of children < 1 year receiving 1 dose, 2 dose, 3 dose of DTP/Penta vaccines
	MCV1	Number (%) of infants who have received at least one dose of measles containing vaccine (MCV) by age 1 year
	PCV 1-3[2]	Number (%) of children < 1 year receiving 1 dose, 2 dose, 3 dose of pneumococcal vaccines
HIV/AIDS	People living with HIV diagnosed	Number of people living with HIV who have been diagnosed and received their results
	HIV care coverage	Number and % of people living with HIV who are receiving HIV care (including pre-ART and ART services)
	PMTCT ART coverage	Number and % of HIV-positive pregnant women who received ART during pregnancy
	ART retention	% of people living with HIV and on ART who are retained on ART 12 months after initiation
	Viral suppression	% of people living with HIV and on ART who are virologically suppressed

[1] If the number of confirmed malaria cases is not available, use all malaria cases,
[2] If this vaccine is not used in country, substitute with another vaccine used in the national programme.

Additional DQR indicators		
Programme area	**Indicator name**	**Full indicator**
TB	Notified cases of all forms of TB	Number (%) of all forms of TB cases (i.e. bacteriologically confirmed plus clinically diagnosed) reported to the national health authority in the past year (new and relapse) – *Assess if quarterly case notification report blocks 1 and 2[1] are correct as per standards and benchmarks (B1.4) for paper-based systems*[2]
	Treatment success rate	Percentage of TB cases successfully treated (cured plus treatment completed) among TB cases notified to the national health authorities in a specified time period – *Assess if quarterly treatment outcome report block 1 is correct as per standards and benchmarks (B.14) for paper-based systems.*
	Treatment success MDR-TB	Number of bacteriologically-confirmed RR and/or MDR-TB cases enrolled on second-line anti-TB treatment during the year of assessment who are successfully treated (cured plus completed treatment)
TB-HIV	HIV test results for registered new and relapse TB patients	Number (%) of new and relapse TB patients who had an HIV test recorded in the TB register expressed as a percentage of the number registered in a specified time period
	Proportion of HIV-positive new and relapse TB patients on ART during TB treatment	Number (%) of HIV-positive new and relapse TB patients who received ART during TB treatment expressed as a percentage of those registered in a specified time period
Malaria	Suspected malaria cases tested	Number of all suspected malaria cases that received a parasitological test
	Confirmed malaria cases receiving treatment	Number (%) of confirmed malaria cases treated that received first-line antimalarial treatment according to national policy at public sector facilities
	Malaria cases (suspected and confirmed) receiving treatment	Number (%) of malaria cases (presumed and confirmed) that received first line antimalarial treatment
	IPTp3	Number of pregnant women attending antenatal clinics who received three or more doses of intermittent preventive treatment for malaria

[1] Definitions and reporting framework for tuberculosis – 2013 revision. Geneva: World Health Organization; 2013 (WHO/HTM/TB/2013.2; http://apps.who.int/iris/bitstream/10665/79199/1/9789241505345_eng.pdf?ua=1, accessed 11 June 2015).

[2] Standards and benchmarks for tuberculosis surveillance and vital registration systems: checklist and user guide. Geneva: World Health Organization; 2014 (WHO/HTM/TB/2014.02; http://apps.who.int/iris/bitstream/10665/112673/1/9789241506724_eng.pdf?ua=1, accessed 11 June 2015).

Annex 2 | Data quality metrics: definitions and requirements

Dimension 1. Completeness and timeliness of data

This dimension measures the extent to which data that are reported through the system used for planning, monitoring and evaluation are available and adequate for these purposes. Are the data complete enough to determine whether the health programme is effective and is achieving the desired results? Are the data sufficiently recent for achievements (or gaps) indicated by the data to reflect the current level of achievement of health indicators? The DQR methodology measures completeness of the data by examining whether all entities that are supposed to report are actually reporting. The indicators in this dimension include completeness of reporting at the health-facility level (usually the level of the first administrative unit), completeness of reporting at levels higher than the health facility (e.g. the district), and the completeness of data elements in submitted reports (i.e. identification of missing data) for programme indicators across the selected programme areas.

Data quality metric: completeness and timeliness of administrative unit reporting

Definition

Completeness of administrative unit reporting (e.g. district or provincial reporting) is defined as the number of administrative unit monthly reports received divided by the total number of reports expected for a specified time period (usually one year). A completeness rate of 100% indicates that all units reported.

Timeliness of administrative unit reporting: It is recommended that the timeliness of reporting should also be evaluated (see Annex 1). Timeliness is defined as the number of reports from subnational administrative units submitted to the national level by the deadline of reporting over the number of reports actually received.

Data requirements

National:
Number of reports received at the national level from the immediately preceding subnational level (e.g. district or region) for the selected period.
Number of reports expected for the period.

Subnational:
Number of reports received from health facilities by month and by district for the selected period
Number of reports expected by month by district.

Calculation

National:
The number of administrative unit monthly reports received divided by the total number of reports expected for a specified time period. A completeness rate of 100% indicates that all units reported.

Subnational:
At the subnational level, a completeness rate is computed for each administrative unit over the specified time period. Units that have a completeness rate less than or equal to 75% are considered to have poor reporting (three or more missing reports for the year).

Box A2.1 Example: completeness of administrative unit reporting

At the national level, if the country has 10 districts, the expected number of reports would be 120 reports (10 reports per month x 12 months). The actual number of reports received was 97 (shown in Table A2.1a). Therefore, the completeness rate would be 97/120 = 81%.

At the subnational level, suppose there are 10 districts which are expected to report monthly. Table A2.1a shows an example of monthly reporting by 10 districts over the span of 12 months. Five out of the 10 districts (50%) have district completeness reporting rates of 75% or less. And, 3 out of 10 districts have 100% reporting completeness rate.

Table A2.1a District reporting example.
District health offices submitting monthly reports on time are indicated with tick marks. Districts with poor reporting (i.e. completeness rate ≤ 75%) are shown in red.

	Month												Total	Completeness rate
	1	2	3	4	5	6	7	8	9	10	11	12		
District 1	✓	✓		✓		✓	✓	✓	✓		✓	✓	9	75%
District 2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	12	**100%**
District 3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	12	**100%**
District 4	✓	✓	✓	✓	✓	✓	✓		✓	✓	✓		10	83%
District 5	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	11	92%
District 6	✓	✓		✓		✓	✓	✓	✓	✓	✓		9	75%
District 7	✓	✓			✓	✓		✓	✓	✓			7	58%
District 8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	12	**100%**
District 9	✓		✓	✓		✓		✓		✓	✓		7	58%
District 10	✓			✓	✓	✓	✓		✓	✓		✓	8	67%
National	**10**	**8**	**6**	**8**	**7**	**10**	**8**	**8**	**9**	**9**	**8**	**6**	**97**	**81%**

Table A2.1b Example of summary results

Metric	Results
National district monthly reporting completeness rate	81%
Number (%) of districts with completeness rate below 75%	5 (50%)
Districts with completeness rate below 75%	District 1, District 6, District 7, District 9, District 10
Number (%) of districts with 100% of expected reports	3 (30%)
Districts with 100% of expected reports	District 2, District 3, District 8

Data quality metric: completeness and timeliness of facility reporting

Definition

The number of reports received from all health facilities nationally, divided by the total number of expected reports from all facilities that are supposed to report to the HMIS for a specified time period (usually one year). The numerator is the actual number of facilities that submit a report and the denominator is the total number of health facilities that are expected to submit a report.

Timeliness of facility reporting is defined similarly: the proportion of reports received from health facilities by subnational administrative units by the deadline for reporting.

Data requirements

Total number of reports received in the fiscal year of analysis from health facilities in the administrative level of analysis (i.e. districts). For example, the total number of health facilities' monthly reports received for January–December 2012 by the administrative level of analysis. For timeliness: the number of reports received by the deadline for reporting.

The total number of health facilities by administrative level of analysis. Please include only facilities that are expected to report to the HMIS system (or any other programme reporting system). If private facilities in a district are not expected to report to any system, then they should not be included in this total count. For timeliness: the number of reports received by the deadline for reporting.

Calculation

National:
The number of reports received from all health facilities nationally, divided by the total expected reports from all facilities that are supposed to report to the HMIS for a specified time period (usually one year)

Subnational:
The facility reporting completeness rate is computed for each administrative unit over the specified time period (usually one year). The number of facilities that submit a report is divided by the number of health facilities expected to submit a report for each administrative unit. Administrative units with reporting rates of 75% or less for facilities within their administrative boundaries are considered to have poor reporting completeness.

Box A2.2 Example: completeness of facility reporting

At the national level, if a country has 1000 facilities reporting to the HMIS, the total number of expected reports for one year would be 1000 x 12 = 12 000 reports. At the end of the year only 10 164 reports have been received (shown in Table A2.2a below). Completeness of facility reporting rate = 10 164/12 000 or 85%.

At the subnational level, facility reporting rates within each of the 10 districts are examined. Districts that have less than 80% facility reporting completeness are shown in red. Three out of 10 districts (30%) have facility reporting rates of less than 80%. A summary of the results is shown in Table A2.2b.

Table A2.2a Facility reporting rate within districts
Districts with facility reporting rates of less than 80% are shown in red.

	Total number of facilities	Expected reports (total facilities x 12 months)	Actual number of reports received in 12 months	Facility completeness rate (%)
District 1	100	1200	1200	100%
District 2	150	1800	1140	63%
District 3	50	600	554	92%
District 4	80	960	960	100%
District 5	120	1440	1080	75%
District 6	170	2040	1920	94%
District 7	130	1560	1270	81%
District 8	100	1200	1200	100%
District 9	40	480	240	50%
District 10	60	720	600	83%
National	**1000**	**12 000**	**10 164**	**85%**

Table A2.2b Example of summary results

Metric	Results
National facility reporting completeness rate	85%
Number (%) of districts with facility reporting completeness rate below 80%	3 (30%)
Districts with completeness rate below 80%	District 2, District 5, District 9

Data quality metric: completeness of indicator data

Definition

Completeness of indicator data is measured by examining the proportion of non-zero values for specific indicators. This is achieved in two ways: 1) by measuring the proportion of blank cells on reporting forms (in the cell where a specific indicator value should be), and 2) by measuring the proportion of cells with a zero recorded as the value.

Missing data should be clearly differentiated from true zero values in district and facility reports. A true zero value indicates that no reportable events occurred in the specified reporting period; a missing value indicates that reportable events occurred but were not actually reported. In many HMIS reports, missing entries are assigned a value of zero, making it impossible to distinguish between a true zero value (no events occurred) and a missing value (events occurred but were not reported). Since it is difficult to differentiate between a true zero value and a true missing value, both these criteria are assessed here. The results of these indicators must be interpreted by data managers and programme managers to ascertain whether zero values represent true zeros.

Data requirements

National:
Number of missing values for selected indicators on administrative unit reports.
Number of zero values for selected indicators on administrative unit reports.
Number of reports received from administrative units at the national level.

Subnational:
Number of health facility reports that contain a missing value for selected indicators in place of an expected indicator value.
Number of health facility reports that contain a zero value for selected indicators in place of an expected indicator value.
Number of health facility reports received for the specified reporting period.[1]

Calculation

National:

a) Completeness of indicator data (zero) (%) is defined as the average percentage of monthly values for selected indicators combined that are not zero for the specified time period (usually one year). Thus the indicator is calculated by subtracting the percentage of values that are zeros from 100%.

b) Completeness of indicator data (missing) (%) is defined as the average percentage of monthly values for selected indicators combined that are non-missing for the specified time period (usually one year). Thus the indicator is calculated by subtracting the percentage of values that are missing from 100%.

[1] These data may not be available at the national level.

Subnational:

At the subnational level (e.g. district or province), this indicator is defined as the percentage of administrative units in which < 90% of the monthly values are non-zero values. This percentage is calculated by summing all the zero values within an administrative unit for each selected indicator for a specified period of time, and dividing by the total number of expected values for the administrative unit for the same specified period of time.

The percentage of administrative units in which non-missing values account for < 90% of monthly values on submitted reports is calculated as above.

Note that the threshold for quality for this metric will vary by health programme (and possibly by country).

Box A2.3 Example: completeness of indicator data - missing values at national and subnational levels

(Use the same procedure to identify the extent of zero values in reported data.)

The example below shows the percentage of missing values for ANC1. Each '✓'means that the district had a non-missing value for the month in question. When examining monthly district-level data for ANC1 over a period of one year, it is seen that, nationally, district-level data shows missing values on 21 occasions. (Follow the same procedure to calculate the percentage of zero values.)

The numerator, 21, is the national total of district-level missing values for ANC1. The denominator is the total expected number of values. With 10 districts and 12 expected monthly values for each district for ANC1, the total expected values nationally are 120. The total % of missing values nationally for ANC1 is 17.5% (21/120). However, since we are calculating values that are not missing, the indicator is 100%-17.5% = 82.5%.

At the subnational level, Table 1 shows that 5 out of 10 districts (50%) have more than 10% missing values for ANC1 within their districts.

Table A2.3a: Missing values by district for ANC1

Districts are marked in red if 10% or more of their values are missing values.

	Month												Total missing	Completeness
	1	2	3	4	5	6	7	8	9	10	11	12		
District 1	✓	✓		✓		✓	✓	✓	✓	✓	✓	✓	2	
District 2	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	0	
District 3	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	0	
District 4	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓		1	
District 5	✓	✓	✓		✓	✓	✓	✓	✓	✓	✓	✓	1	
District 6	✓	✓		✓		✓	✓	✓	✓	✓	✓		3	
District 7	✓	✓			✓	✓		✓	✓	✓			5	
District 8	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	✓	0	
District 9	✓		✓	✓		✓		✓		✓	✓		5	
District 10	✓			✓	✓	✓	✓		✓	✓		✓	4	
National	**0**	**2**	**4**	**2**	**3**	**0**	**2**	**1**	**1**	**0**	**2**	**4**	**21**	**17.5%**

Table A2.3b Example of summary results

Metric	Results
National district monthly reporting completeness rate	100%-17.5% = 82.5%
Number (%) of districts with completeness rate below 90%	5 (50%)
Districts with completeness rate below 90%	District 1, District 6, District 7, District 9, District 10

Dimension 2. Internal consistency of reported data

This dimension examines the plausibility of reported results for selected programme indicators based on the history of reporting for those indicators. Trends in reported data are evaluated over time (one year) to assess whether specific reported values (e.g. monthly) within the selected period are extreme in relation to the other values reported and if they are potentially indicative of data quality problems. Trends in reporting over multiple years are also evaluated to identify extreme or implausible values year-to-year.

Within this dimension, the results of programme indicators are compared to other indicators with which they have a predicable relationship in order to determine whether the expected relationship exists between the two indicators. In other words, is the observed relationship between the indicators, as depicted in the reported data, that which we would expect on the basis of our knowledge of the indicators, the health programme and the country?

This dimension also seeks to determine the accuracy of reporting for selected indicators through the review of source documents (i.e. the documents in which reported events are first recorded) to compare reported values to a validated value. This aspect of the DQR is conducted through a health facility assessment linked to the implementation of the DQR. (See "Data verification" below.)

Data quality metric: outliers in the current year

Definition

An outlier is defined as a value in a series of values that is extreme in relation to the other values in the series. Outliers can be the result of changes in programmatic activities (such as an intensified campaign) or can be the result of data quality problems. Extreme values should be identified and investigated to determine whether they are valid or if they are the result of insufficiencies in data quality.

Outliers can be identified by various methods, though it is recommended that one of the following two methods be used:

1. *Multiples of the standard deviation of the mean:* Values in a series greater than multiples of the standard deviation of the mean of the series of values (i.e. ± 2SD, ± 3SD, etc.) are identified as potential outliers and are evaluated for data quality problems. Outliers identified as greater than 2SD from the mean are considered "moderate" outliers, while those identified as greater than 3SD from the mean are considered "extreme".
2. *Modified Z-score:* The Z-score of an observation refers to the number of standard deviations from the mean. A "modified Z-score" applies the median computation technique to measure the deviation and in many cases provides more robust statistical detection of outliers (than use of the mean). This method is useful for small samples and is more tolerant than the Z-score to extreme values. Mathematically the modified Z-score can be written as:

$M_i = 0.6745 * (X_i - \text{Median}(X_i)) / \text{MAD}$,

where MAD is defined as the median absolute deviation. MAD = $\textit{median}(|X_i - \tilde{X}|)$, where $\tilde{X}$ is the median of the series. Any number in a dataset with the absolute value of modified Z-score exceeding 3.5 is considered an outlier.[1]

Data requirements

National:

[1] Iglewicz B, Hoaglin D. The ASQC basic references in quality control: statistical techniques. Volume 16: How to detect and handle outliers. Milwaukee (WI): American Society for Quality; 1993.

Monthly indicator values for selected indicators from administrative units reporting to the national level from HMIS reports (or database) for the selected period. Identify extreme values (outliers) according to selected method from above.

Subnational:
Monthly indicator values for selected indicators from health facilities from HMIS reports (or database) for the selected period. Identify extreme values (outliers) according to selected method from above.

Calculation

National:
Moderate outliers for monthly values of a selected indicator are identified from values reported for a given period using the methods described above. The total number is divided by the expected number of values for the indicator. If the time period of analysis is one year and reporting is monthly, the total number of expected values for one indicator equals the total number of administrative units of analysis multiplied by 12. A similar calculation is performed for extreme outliers.

Subnational:
Moderate outliers: At the subnational level (e.g. district or province), the aim is to calculate the percentage of administrative units in which two or more of the monthly values of the selected indicator are moderate outliers (between ± 2–3 SD from the administrative unit mean, or a value of > 3.5 on the modified Z-score). This percentage is calculated by identifying and counting all the moderate outliers within an administrative unit for the selected indicator for a specified period of time and dividing the result by the total number of expected values for the indicator in the administrative unit for the same period of time.

Extreme outliers: At the subnational level (e.g. district or province), the percentage of administrative units in which one or more of the monthly administrative unit values for the selected indicator is an extreme outlier (± 3SD from the administrative unit mean) is calculated by dividing the total number of administrative units with extreme outliers for the specified time period with the total number of administrative units.

Box A2.4 Example: outliers in the current year

First, we will examine outliers for ANC1. Table A2.4a below shows moderate outliers for ANC1. There are 8 moderate outliers for ANC1, and these are highlighted in red. Eight of the districts have at least one occurrence of a monthly ANC1 value that is a moderate outlier.

Table A2.4a Monthly ANC1 values by district

Values in red are moderate outliers.

	Month												Total	% of values that are outliers
	1	2	3	4	5	6	7	8	9	10	11	12		
District 1	2543	2482	2492	2574	3012	2709	3019	2750	3127	2841	2725	2103	1	8.3%
District 2	1547	1340	1403	1593	2161	1729	1646	1642	1355	1581	1412	1410	1	8.3%
District 3	776	541	515	527	857	782	735	694	687	628	596	543	0	0.0%
District 4	1184	1118	1195	1228	1601	1324	1322	711	1160	1178	1084	1112	2	16.7%
District 5	1956	1773	1768	2062	2997	2056	1839	1842	2028	2002	2032	1904	1	8.3%
District 6	819	788	832	802	999	596	672	792	933	1134	810	789	1	8.3%
District 7	781	1199	981	963	818	897	853	736	2208	2734	1323	1229	1	8.3%
District 8	1382	1379	1134	1378	1417	1302	1415	1169	1369	1184	1207	1079	0	0.0%
District 9	1992	1751	1658	1823	3306	2692	2300	2218	2026	2003	1752	1753	1	8.3%
District 10	3114	2931	2956	4637	6288	4340	3788	3939	3708	4035	3738	3606	1	8.3%
National	**0**	**0**	**0**	**0**	**5**	**0**	**0**	**1**	**0**	**2**	**0**	**1**	**9**	**6.7%**

Nationally, this indicator is a percentage of values that are moderate outliers for the indicator. The numerator is the number of outliers across all administrative units (9). The denominator is the total number of expected reported values for the indicator for all the administrative units. It is calculated by multiplying the total number of units (at the selected administrative unit level) with the expected number of reported values for one indicator for one administrative unit. The denominator is then calculated as follows: 10 districts x 12 expected monthly reported values per district for one indicator = 120 total expected reported values. The average percentage of reported values that are moderate outliers equals ((9/120)*100 ≈ 7.5%).

Subnationally, the number of outliers is calculated for each district. This is done by counting the districts where there are two or more outliers (for moderate outliers) among the monthly values for the district. This is then divided by the total number of administrative units: 1/10 = 0.1 * 100% = 10%.

Table A2.4b Example of summary results

Metric	Results
% of district monthly values that are moderate outliers (between ± 2–3 SD from the district mean)	7.5%
Number and % of districts in which two or more of the monthly district values for the indicator are moderate outliers (between ± 2–3 SD from the district mean)	1, 10.0%

Data quality metric: consistency over time

Definition

Consistency over time (%) is defined as the average ratio of events/service outputs for the current year of analysis to the mean events/service outputs of up to three preceding years for selected indicators. Consistency over time is also measured as a comparison of the current year to the value predicted from the trend over the three preceding years for indicators or programmes with expected growth or decline. Current year values are compared to forecasted values (the value predicted by the slope of the values of three previous years) for indicators with non-constant trend (i.e. increasing or decreasing).

This indicator shows the consistency of the values for key indicators in the most recent year compared with the mean value of the same indicator for the previous three years combined (or the forecasted value for indicators with non-constant trend). Differences in values are expected from one year to the next; however, if the differences are very large, they warrant further scrutiny. While large differences usually suggest some type of reporting error, it is also possible that the introduction of a new intervention may have contributed to a large percentage increase in indicator values from one year to the next. Hence, interpretation of the results with programme managers is critical.

Data requirements

Annual totals by subnational unit for selected indicators for the year of analysis plus the preceding three years.

Calculation

National:
At the national level this indicator is as defined above – the ratio of the current year total to the average of the preceding three years., or

The current year value compared to the value forecasted from the three previous years of values for indicators with non-constant trend (i.e. increasing or decreasing)
Subnational:
Subnationally, this indicator looks at the percentage of administrative units at the selected administrative level of analysis with at least 33% difference between their ratio and the national ratio for selected indicators.

Alternatively, the subnational unit ratios can be compared to the 'expected value', i.e. equality between the current year value and the average of the three preceding years (or forecasted value). For this comparison the subnational unit value is compared to 1± the quality threshold. For example, if the quality threshold is set at 33% subnational units with ratios ≥ 133%, or ≤ 67% would be flagged as potential data quality problems.

Box A2.5 Example: consistency over time

First, we consistency over time is examined for institutional deliveries:

National total for institutional deliveries for 2010 = 211 194
National total for institutional deliveries for 2011 = 205 863
National total for institutional deliveries for 2012 = 199 344
National total for institutional deliveries for 2013 = 220 916.

The mean of 2010, 2011 and 2012 = ((211 194 + 205 863 + 199 344)/3) = 205 467.

The ratio of the current year to the mean of the past three years for ANC1 = 220 916/205 467) ≈ 1.08.

The average ratio of 1.08 shows that there is an overall 8% increase in the service outputs for institutional deliveries in 2013 when compared to the average service outputs for the preceding three years.

Table A2.5a Consistency trend: comparison of district ratios to national ratios
More than 33% difference between district and national ratio is highlighted in red

	2010	2011	2012	2013	Average of preceding 3 years	Ratio of current year to mean of preceding 3 years	% difference between national and district ratios
District 1	30242	29543	26848	32377	28878	1.12	0.04
District 2	19343	17322	16232	18819	17632	1.07	0.01
District 3	7512	7701	7403	7881	7539	1.05	0.03
District 4	15355	15047	14788	25123	15063	1.67	0.55
District 5	25998	23965	24023	24259	24662	0.98	0.09
District 6	10234	9458	9654	9966	9782	1.02	0.05
District 7	14011	13987	14355	14722	14118	1.04	0.03
District 8	15233	15974	14733	15415	15313	1.01	0.06
District 9	23033	24544	24433	25274	24003	1.05	0.02
District 10	50233	48322	46875	47080	48477	0.97	0.10
National	**211 194**	**205 863**	**199 344**	**220 916**	**205 467**	**1.08**	

Subnationally, evaluate each district by calculating, for institutional deliveries, the ratio of the current year (2013) to the average of the previous three years (2010–2012). For example, the ratio for District 1 is 32 377/28 878 = 1.12.

Next, calculate the % difference between the national and district ratios for each district. Again, for District 1:

$$\left|\frac{District\ 1\ ratio - National\ ratio}{National\ ratio}\right| = \left|\frac{1.12-1.08}{1.08}\right| = 0.04 = 4.0\%$$

The percentage difference between the district ratio and the national ratio for institutional deliveries in District 1 is less than 33%. However, there is a difference of approximately 55% between District 4's institutional deliveries ratio and the national ratio.

To calculate this indicator subnationally, all administrative units whose ratios are different from the country's ratio by ± 33% or more are counted. In this example, only District 4 has a difference greater than ± 33%. Therefore, 1 out of 10 districts (10%) has a ratio that is more than 33% different from the national ratio.

Table A25b Example of summary results

Metric	Results
Average ratio of events/service outputs for the current year to the mean events/service outputs for the three preceding years for institutional deliveries	8.0%
Number (%) of districts with at least 33% difference between the district and national ratio	1 (10%), District 4

Data quality metric: consistency between related indicators

Definition

This data quality metric examines the extent to which two related indicators follow a predictable pattern. If this pattern is not followed at the national level or for a particular subpopulation it may be indicative of data quality problems.

Consistency between two indicators is defined as the ratio between these two indicators. For some indicators, the ratio should be 1 or below; for other indicators the ratio is ≥ 1 (see Annex 1 for indicator-specific details).

Data requirements

Yearly values of selected indicators for national and subnational levels.

Calculation

National:
At the national level, this indicator is the ratio of the two selected indicators.

Subnational:
At the subnational level, for indicators which should be roughly equal, this indicator shows the percentage of subnational administrative units that have an extreme difference (e.g. ≥ ± 10%). For indicators which should be ≥ 1, districts with ratios of < 1 should be flagged. The number and percentage of subnational units with anomalous values is calculated (number of subnational units with anomalous values/total number of subnational administrative units).

The relationship between two indicators at subnational units can also be assessed by comparing their ratio with the ratio between the two indicators at the national level. In this instance the ratio percentage difference is calculated between the ratio at subnational level and the ratio at national level. Subnational units with a percentage difference greater than the specified quality threshold (e.g. ≥ 10%) are flagged for follow-up.

Box A2.6 Example: consistency between related indicators

The annual number of pregnant women started in antenatal care each year (ANC1) should be approximately equal to the number of pregnant women who receive intermittent preventive therapy for malaria (IPT1) in ANC since all pregnant women should receive this prophylaxis. First, the ratio of ANC1 to IPT1 is calculated for the national level and then for each district (Table A2.6a). At the national level the ratio of ANC1 to IPT1 is 154 285/134 341 = 1.15.

Table A2.6a % difference between ANC1 and IPT1 by district
Districts with % difference ≥ ± 20% are flagged in red.

	ANC1	IPT1	Ratio of ANC1/IPT1	District ratio/national ratio
District 1	20 995	18 080	1.16	1.01
District 2	18 923	16 422	1.15	1.00
District 3	7682	6978	1.10	0.96
District 4	15 669	14 151	1.11	0.97
District 5	12 663	9577	1.32	1.15
District 6	20 233	19 960	1.01	0.88
District 7	11 402	9291	1.23	1.07
District 8	12 520	10 461	1.20	1.04
District 9	15 984	13 930	1.15	1.00
District 10	18 214	15 491	1.18	1.03
National	**154 285**	**134 341**	1.15	

At the subnational level we calculate the comparison of the subnational unit ratio to the national ratio:

$$\frac{ANC1_{(subnational\ unit)}/IPT1_{(subnational\ unit)}}{ANC1_{(national)}/IPT1_{(national)}}$$

Any subnational unit with a value ≥ national ratio + specified quality threshold (e.g. 20%), or a value ≤ national ratio – specified quality threshold, is flagged as a potential data quality problem.

Comparison of subnational unit ratio to the expected ratio:

Since all pregnant women entering into ANC should receive IPT, the expected result is that the value of IPT1 should be roughly equal to the value of ANC1, or slightly less. The ratio of IPT1:ANC1 should be roughly equal to 1.

$$ANC1/IPT1 \approx 1$$

Any subnational unit with a value of ANC1/IPT1 ≥ 1 + specified quality threshold, or ≤ 1 – specified quality threshold, should be flagged for follow-up.
In the example above, we see that there are three districts with a ratio of ANC1 to IPT1 greater than 20% - districts 5, 7 and 8. When district ratios are compared to the national ratio, no districts surpass the quality threshold of 20%.

Table A2.6b Example of summary results

Metric	Results (comparison with expected result)	Results (comparison with national result)
National ratio of ANC1 to IPT1	1.15	1.15
Number (%) of districts with ratio of ANC1 to IPT1 of ≥ 20%	3 (30%)	0
Districts with ANC1:IPT1 ≥ 20%	District 5, District 7, District 8	None

Data quality metric: verification of reporting consistency through facility survey

Definition

% agreement between verified counts for selected indicators in sampled facility records and reported values for the same facilities (requires visit to health facilities).

(See section 3.1 on data quality analysis and interpretation for more detail)

Data requirements

District: Monthly district indicator values for a selected reporting period for priority indicators.
Recounted monthly district indicator values of the same indicators and reporting period.

Health facility: Monthly health facility indicator values for a selected reporting period for priority indicators.
Recounted monthly health facility indicator values for the same indicators and reporting period.

Calculation

Verification factor: $\frac{\text{Recounted number of events from the source document}}{\text{Reported number of events from the HMIS}}$
…aggregated over domains of estimation.

Dimension 3. External consistency

(Agreement with other sources of data such as surveys)
The purpose of this dimension is to examine the level of agreement (i.e. external consistency) between two sources of data measuring the same health indicator. The two sources of data are the routinely collected and reported data from the HMIS or programme-specific information system, and a periodic population-based survey. Surveys are generally considered to have reliable results since the methods for conducting surveys are highly standardized and great care and expense are used to ensure high-quality implementation and estimates of health indicators. Survey results are often considered the "gold standard" or true value of the indicator in the population.

The expense of surveys means that they cannot be conducted regularly, and there are limitations on the interpretation of survey results in smaller geographical areas. For these reasons surveys alone are not adequate for routine monitoring of health sector and programme results.

Data quality metric: external comparison with survey results

Definition

The purpose of this metric is to examine the external consistency of the indicators generated from the health facility data. This implies a comparison with an external source, such as coverage data from a recent household survey.

While survey results are often considered a "gold standard," surveys are also subject to data quality problems, and if these problems are systematic the survey-based coverage estimate can be well off the true population value. In addition, surveys are based on a sample and therefore have a range of possible values (i.e. confidence interval, limits). Confidence intervals are larger if the sample is smaller, and therefore much larger for subnational levels than for the national level. (Confidence intervals are often presented in the annexes of survey reports, e.g. DHS.) In a comparison with routine data, the survey confidence limits need to be taken into account. If the routine value lies within the range, it cannot be concluded that there is a significant difference from the survey value. Aditionally, survey results may reflect past performance (often three or five years before the survey), while coverage rates based on routine data are usually for the most recent year. Thus, this comparison should be made with caution.

External comparison of selected indicators is defined as the ratio of the coverage derived from routinely reported data (e.g. HMIS) and the coverage rate from household survey data.

Comparison of HMIS values to health programme values for selected indicators: this metric can be calculated using the same method as the comparison of routinely reported data to survey results.

Data requirements

- National and subnational administrative values for selected indicators and year.
- Appropriate denominators to derive coverage rates for routine data.
- Analogous survey value for the same year from a recent household survey with methodology that meets international standards for quality (e.g. MICS, DHS).

Calculation

National:
At the national level this indicator is defined as the ratio of the routine value to the survey value.

Subnational:

At the subnational level, the ratio of the coverage rates is calculated for each administrative unit. Any administrative ratio that has at least 33% difference between the two coverage rates is flagged. The number and percentage of administrative units with at least 33% difference is then calculated. This comparison is possible only if the survey coverage estimates are available for the indicator for the same administrative level. For example, if the administrative unit of analysis is a district but survey coverage rates for the indicator are not available at the district level, this subnational comparison will not be possible at the district level. However, if provincial or regional level survey data are available, the comparison can be made at the provincial level.

Box A2.7 Example: external comparison with survey results

If the HMIS is accurately detecting all ANC visits in the country (and not just those limited to the public sector) and the denominators are sound, the coverage rate for ANC1 derived from the HMIS should be very similar to the ANC1 coverage rate derived from population surveys. However, the coverage rates from HMIS often differ from survey coverage rates for the same indicator.

Table A2.7a Comparison of HMIS and survey coverage rates for ANC1
Discrepancies of more than 33% between the two are highlighted in red.

	Facility coverage rate	Survey coverage rate	Ratio of facility to survey rates	% difference between official and alternate denominator
District 1	1.05	0.95	1.1	**10%**
District 2	0.91	0.97	0.94	**6%**
District 3	1.39	0.9	1.54	**54%**
District 4	0.76	0.95	0.8	**20%**
District 5	0.96	0.8	1.2	**20%**
District 6	0.93	0.98	0.96	**4%**
District 7	0.84	0.86	0.98	**2%**
District 8	1.1	0.98	1.13	**13%**
District 9	1.38	0.92	1.5	**50%**
District 10	0.91	0.79	1.16	**16%**
National	**0.98**	**0.93**	**1.05**	**5%**

At the national level:

The coverage rate from the HMIS is 98%.
The coverage rate from the most recent population-based survey is 93%.
The ratio of the two coverage rates is: 98%/93% = **1.05**.
If the ratio is 1, the two coverage rates are exactly the same.
If the ratio is > 1, the HMIS coverage rate is higher than the survey coverage rate.
If the ratio is <1, the survey coverage rate is higher than the HMIS coverage rate.

The ratio of 1.05 shows that the two denominator values are fairly similar to each other, with approximately 5% difference between the two values.

At the subnational level, the ratio of denominators is calculated for each administrative unit. Districts with at least 33% difference between their two denominators are flagged. In Table A2.7a above, District 3 and District 9 have at least 33% difference between their two ratios.

Table A2.7b Example of summary results

Metric	Results
National ANC1 coverage rates consistency ratio	0.946
Districts with ANC1 consistency ratio below 0.67 (survey coverage rate is higher)	0
Districts with ANC1 consistency ratio above 1.33 (HMIS coverage rate is higher)	2 (10%)

Dimension 4. External comparisons of population data

(Review of denominator data used to measure performance indicators)

The use of population data in the calculation of health indicators allows comparisons of results within or across geographical areas, over time, and between population subgroups. The population data for a specific indicator and a specific geographical area or population subgroup (e.g. pregnant women) serve as the denominator in the calculation of a rate or a proportion and provide context to the numerator (e.g. the number of events, patients, commodities, etc. for the health process in question). The use of population data for the calculation of indicators is critical to the

effective monitoring and evaluation of health programmes. However, population data in many countries are known to be poor. The purpose of this dimension is to determine the adequacy of the population data used in the calculation of health indicators. It achieves this by comparing two different sources of population estimates (for which the values are calculated differently) to assess the level of congruence between the two sources. If the two population estimates are discrepant, the coverage estimates for a given indicator can be very different even though the programmatic result is the same (i.e. the number of events). The higher the level of consistency between denominators from different sources, the more confidence can be placed in the accuracy of the population projections.

Data quality metric: consistency with United Nations population projections

Definition

For this indicator, the denominator (total population of interest) used for one of the selected indicators included in the DQR is compared to United Nations population projections. Denominators that are used to calculate rates and ratios are usually derived from the census or civil registration system. Denominators from the census are usually population projections based on estimates of natural growth and migration.

Consistency with United Nations population projections is defined as the ratio between the official country projection for the number of live births or pregnant women divided by the official United Nations projection for the same population for the same year.

Consistency of denominators between programme data and official government statistics must be taken into account. If health programmes maintain their own population estimates, as is often the case in immunization programmes which conduct community microplanning, these programme-specific denominators can be compared with the same population estimates of the official government statistics office using the procedure outlined here (see Box A2.9).

Data requirements

Population estimates are used as denominators for calculating rates for selected indicators. The most common denominator used for calculating ANC rates and delivery rates is the total number of live births in a specified period of time. For immunization, the most commonly used denominator is the total number of surviving infants (total live births adjusted for infant mortality), and for outpatient visits it is the total population. Comparable denominators available from United Nations projections are births and total population.

Calculation

National:
At the national level this indicator is defined as the ratio between the official country projection (from the Census Office or National Statistics Office) and the United Nations population projection.

This quality metric is not calculated for the subnational level.

Box A2.8 Example: consistency with United Nations population projections

If the official live birth estimate for the year of analysis is 255 000 and the projected United Nations population is 200 000, the ratio of country population estimate to United Nations population projection is 255 000/200 000 ≈ 1.28.

This ratio shows that the country population estimate for live births is higher than the United Nations population projection for the same year.

Data quality metric: consistency of denominators – comparison of official government statistics and denominators used by health programmes

Definition

This metric measures the consistency of population estimates used for the calculation of health system performance indicators. Population data for common indicators from official government sources (e.g. National Statistics Office) are compared to values for the same populations used by health programmes (if applicable) to determine the level of agreement between the two sources. Recommended programme indicators (and their associated denominators) used for this comparison are noted in Annex 3.

Data requirements

National:
Official government population estimates for denominators used in the calculation of rates for selected indicators for the year of analysis (live births, expected pregnancies, children < 1 year, total population).

Data for analogous subpopulations used by health programmes.

Subnational:
Subnational administrative unit population estimates for denominators used in the calculation of rates for selected indicators for the year of analysis.

Calculation

National values from official government statistics for live births, expected pregnancies, children < 1 year and total population are divided by analogous health programme values to determine agreement. Values that differ significantly (recommended 10%, but can also be defined by users) are flagged for review.

At the subnational level this indicator is defined as the number and percentage of subnational units where there is a significant discrepancy (± 10%) between the two denominators.

Box A2.9 Example: consistency of population data

In this example using live births, the subnational administrative unit values from official government sources are compared to health programme sources in order to determine agreement.

Calculate the ratio of the number of live births from official government statistics nationally for the year of analysis to the value used by the selected health programme = 0.97

Calculate the ratio of subnational administrative unit live births in 2014 to the value used by the selected health programme (Table A2.9a).

Table A2.9a Consistency of population trend – national and subnational administrative unit ratios of official government live birth estimates for 2014 to the live births estimates for 2014 used by the health programme
Administrative units with a difference of ≥ 10% are indicated in red.

	Official government estimates for live births (2014)	Health programme estimates for live births (2014)	Ratio
District 1	29 855	29 351	1.02
District 2	23 398	23 032	1.02
District 3	6893	7420	0.93
District 4	18 832	19 938	0.94
District 5	15 032	14 844	1.01
District 6	25 023	30 141	0.83
District 7	14 983	15 004	1.00
District 8	14 556	14 960	0.97
District 9	12 973	13 054	0.99
District 10	25 233	25 283	1.00
National	**191 003**	**194 882**	**0.97**

District 6 has a difference of 0.17, or 17%.

Table A2.9b Example of summary results

Metric	Results
National ratio of official government estimate of live births to the value of live births used by the health programme	0.97
Number (%) of districts where the difference between official government live births to health programme live births is ≥ 10%	1 (10%)
Districts with a difference rate of ≥ 10%	District 6

Annex 3 | Data requirements and formatting for a DQR desk review

Table A3.1 shows the data requirements for a DQR.

Table A3.1 DQR data requirements

Programme	Data type		Indicator
General service statistics	Population		▪ Total population
	Routine		▪ Total outpatient visits
Maternal health	Population		▪ Estimated number of pregnant women ▪ Expected pregnancies ▪ Estimated number of live births
	Survey	Core	▪ ANC1 coverage ▪ Institutional deliveries
		In-depth	▪ Tetanus toxoid (TT) 1st dose
	Routine	Core	▪ ANC 1st visit
		In-depth	▪ ANC 4th visit ▪ Institutional deliveries ▪ ITP1 ▪ Tetanus toxoid 1st dose ▪ Postpartum care coverage
Immunization	Population		▪ Estimated number of children < 1 year
	Survey		▪ Estimated coverage with 3rd dose DTP-containing vaccine
	Routine	Core	▪ 3rd dose DTP-containing vaccine in children < 1 year
		In-depth	▪ 1st, 2nd, 3rd dose DTP-containing vaccine (DTP1-3/Penta1-3) ▪ Number of children vaccinated with 1st dose of measles-containing vaccine ▪ Doses PCV 1–3 in children < 1 year[1]
HIV/AIDS	Population		▪ Total population ▪ HIV prevalence to estimate population in need
	Survey	Core	▪
		In-depth	▪ HIV C&T during last 12 months ▪ Pregnant women HIV-tested in ANC
	Routine	Core	▪ Number and % of PLHIV who are receiving HIV care (including pre-ART and ART services) (HIV coverage)
		In-depth	▪ % of HIV+ persons on ART (or ART coverage)[2] ▪ PMTCT ART coverage ▪ ART retention at 12 months ▪ Viral suppression
TB	Population		Total population
	Routine	Core	▪ Number of notified TB cases (all forms of TB)
		In-depth	▪ Number of TB cases successfully treated (all forms of TB) ▪ Number of TB cases (new and relapse) tested for HIV ▪ Number of HIV+ TB patients initiated on ART ▪ Number of MDR-TB cases detected ▪ Number of MDR-TB cases successfully treated
Malaria	Population		▪ Total population
	Survey	Core	▪
		In-depth	▪ Proportion of pregnant women treated with 3 or more doses of IPTp ▪ Percentage of children with fever who took first-line antimalarial among those given any antimalarial treatment
	Routine	Core	▪ Number of all suspected malaria cases that received a parasitological test
		In-depth	▪ Number of confirmed malaria cases ▪ Number of confirmed malaria cases treated ▪ Total number of malaria cases (suspected and confirmed) treated ▪ Number of pregnant women attending antenatal clinics treated with 3 or more doses of IPTp

[1] If the country has implemented vaccination with PCV, note that some countries may use in a 2+1 schedule where the third dose may be given at or after 12 months.

[2] Depending on the country's policies on ARV coverage – e.g. adoption of WHO's 2013 ARV guidelines recommendation of 85% of HIV-infected persons on treatment.

Formatting the data

The data from the HMIS or the health programmes should, if possible, be monthly aggregate data for priority indicators from the health-facility level. If data are not available from the health-facility level, then – at a minimum – aggregate data should be obtained from the district level. The data should be formatted to facilitate the calculation of data quality metrics in the DQR – i.e. a flat file with one row per health facility (or district) and monthly indicator values in columns.

Monthly service data

Annual data (disaggregated by month) should be provided for the fiscal year of analysis (e.g. January–December, July–June, etc.) for the selected administrative level (the recommended level of analysis is the district level) for all the routine indicators listed above. All administrative units in the country should be included (for the selected level of analysis). The format of Table A3.2 can be used for each of the indicators.

Table A3.2 Format for monthly service data

Number	Administrative unit	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sept	Oct	Nov	Dec
1	District A												
2	District B												
3	District C												
4	District D												
5	District E												
6…	District F…												
…N	…District Z												

Service data trend information

Annual data for up to three years preceding the fiscal year of analysis should be provided for each administrative unit at the selected level of analysis for the selected programme indicators. For example, if the year of analysis is January–December 2013, annual data for 2010, 2011 and 2012 (if available) should be provided for each of the indicators. The format of Table A3.3 can be used for the trend data.

Table A3.3 Format for trend data

		Indicator 1			Indicator 2			Indicator 3			Indicator 4		
Number	Administrative unit	2011	2012	2013	2011	2012	2013	2011	2012	2013	2011	2012	2013
1	District A												
2	District B												
3	District C												
4	District D												
5	District E												
6…	District F…												
…N	…District Z												

Population data

Population data are required by the selected administrative level of analysis for specific population groups. For example, if the level of analysis is the district level, population data on the following will be required:

- total population
- number of pregnant women
- number of live births
- number of children under 1 year of age.
- number HIV positive

The format of Table A3.4 can be used for the data on specific population groups.

Table A3.4 shows the format for the comparison of official government denominators (e.g. from the National Statistics Bureau) to the same denominators used by health programmes (if applicable).

Table A3.5 shows the format for denominators used to calculate population rates for the program level indicators used in the DQR to assess data quality. These denominators are used in Domaine 3 – External Comparisons to compare routinely reported results to population-based survey results for the same indicators. Both the indicators in table A3.4 and table A3.5 are used in the DQR to evaluate data quality and the quality of denominator data.

Table A3.4 Format for data on population groups; domain four – external comparison of population data

		Denominators from Official Government Statistics Bureau				Denominators Used By Health Programmes			
Number	Administrative unit	Total population	Expected Pregnancies	Number of live births	Number of children under 1 year of age (surviving infants)	Total population	Expected Pregnancies	Number of live births	Number of children under 1 year of age (surviving infants)
1	District A								
2	District B								
3	District C								
4	District D								
5	District E								
6…	District F…								
…N	…District Z								

Table A3.5 Format for data on population groups; domain three – external consistency

		Denominators used to calculate rates for program level indicators selected for DQR				
Number	Administrative unit	Antenatal care 1st visit (ANC1) (Expected Pregnancies)	DTP3/Penta3 (Live Births)	ART Coverage (Number HIV+)	Notified cases of all forms of TB (Total Population)	Confirmed malaria cases (Total Population)
1	District A					
2	District B					
3	District C					
4	District D					
5	District E					
6…	District F…					
…N	…District Z					

Monthly HMIS or programme reports

Information on monthly reports submitted by health facilities to their reporting unit (usually the district) and from districts to their reporting unit is required for calculating the completeness of reporting. Information should be provided on the following items by administrative level of analysis for the full year.

- The total number of monthly reports received from the administrative level of analysis (i.e. districts) – e.g. if districts are expected to submit a report to their reporting unit each month, the actual number of reports submitted (if available) should be included.
- The total number of reports received from health facilities in the fiscal year of analysis at the administrative level of analysis (i.e. districts) – e.g. the total number of health facilities' monthly reports received for January–December 2012 by the administrative level of analysis.
- The total number of health facilities by administrative level of analysis. Only those facilities that are expected to report to the HMIS system (or any other programme reporting system) should be included. If private facilities in a district are not expected to report to any system, they should not be included in this total count.

The format of Table A3.6 for reporting on reports received.

Table A3.6 Format for reporting on reports received

Number	Administrative unit	Total of district reports received	Total of district reports received by the deadline	Total of health facilities reporting into the HMIS	Total of health facilities' reports received	Total of health facilities' reports received by the deadline
1	District A					
2	District B					
3	District C					
4	District D					
5	District E					
6…	District F…					
…N	…District Z					

Household survey data

Formatting the data can be difficult if the data originate from a wide variety of sources. Data managers should be allowed sufficient time to produce good-quality data for the analysis since hastily prepared data could hinder the calculation of data quality metrics.

Estimates from the most recent household survey for the selected indicators (Annex 1), with standard errors (where available) and by domain of estimation used in the survey (i.e. state/province/region), can be formatted as in Table A3.7.

Table A3.7 Format for reporting household survey data

Number	Administrative unit	Survey indicator 1		Survey indicator 2		Survey indicator 3	
		%	Standard error	%	Standard error	%	Standard error
1	Region A						
2	Region B						
3	Region C						
4	Region D						
5	Region E						
6…	Region F…						
:	**National**						

Administrative units and data flow information

Depending on the administrative level of detail (i.e. facility, district, region) that is permitted by the data available for the analysis, information will be required on the data flow from that level to the national level. If data are available for the district level and the chain of reporting is district→region→national, information should be provided on all districts in the country as well as on the regions to which they report. Table A3.8 demonstrates the desired format.

Table A3.8 Format for reporting on data flow

Number	Administrative level of analysis	Region to which the district reports
1	District A	Region UU
2	District B	Region UU
3	District C	Region VV
4	District D	Region VV
5	District E	Region XX
6…	District F…	Region XX
…N	…District Z	Region XX

Annex 4 | Recommended source documents and cross/spot-checks for data verification

Table A4.1 below shows the core and additional indicators with data sources and relevant cross-checks that can be implemented during data verification. However, it is recommended that these cross-checks by conducted during in-depth DQRs.

Table A4.1 Cross-checks and spot-checks for verification of data

Program	Indicator	Data source	Cross-checks and spot-checks
General service statistics	▪ Total outpatient visits (OPD)	▪ OPD register	
Maternal health	▪ ANC 1st visit ▪ ANC 4th visit ▪ Institutional deliveries ▪ PNC1 ▪ TT1	▪ Labour and delivery facility register ▪ ANC register ▪ PNC register	▪ ANC/PNC registers can be cross-checked with the patient cards if those are kept at the health facility. ▪ Speak with patients at the facility at the time of data verification and ask about the services they received. Check against the relevant register whether the services and treatments given have been captured correctly.
Immunization:	▪ DTP1–3 /Penta 1–3 ▪ MCV1 ▪ PCV 1-3[1]	▪ Tally sheets	▪ Immunization registers can be cross-checked with the number of doses of vaccine used (keeping in mind that some vaccines come in batches of 10-dose vials and one batch may be used for fewer than 10 children). ▪ Records of vaccination on a sample of child vaccination cards can be verified against the immunization register for children in the health facility on the day of the verification visit.
HIV/AIDS[2]	▪ ART coverage ▪ HIV coverage ▪ PMTCT ART coverage ▪ ART retention ▪ Viral suppression	▪ Programme records (ART register, ART patient cards) ▪ Facility-based ART registers ▪ Health facility data aggregated from patient monitoring system	▪ ART registers can be cross-checked against pharmacy records. ▪ Patient files can be cross-checked against the information in the patient database (if a database exists at the facility). ▪ Spot-checks: patients at the facility at the time of verification can be asked about the services they received. Confidentiality should be paramount; if the confidentiality of the patient cannot be guaranteed, the spot-check should not be conducted.
TB[3]	▪ Notified cases of **all** forms of TB ▪ Number of all forms of TB cases successfully treated ▪ Treatment success MDR-TB ▪ HIV test results for registered new and relapse TB patients ▪ Proportion of HIV-positive new and relapse TB patients on antiretroviral therapy (ART) during TB treatment	▪ TB unit registers	Cross-check: TB cases detected (from laboratory registers) checked against TB cases notified (initial defaulters) ▪ The TB unit register can be cross-checked against the TB treatment cards. ▪ The TB unit register can be cross-checked against the laboratory register to verify that those diagnosed are actually reported (if diagnosis is being conducted at the facility). ▪ The TB unit register can be cross-checked against the pharmacy records.
Malaria	▪ Confirmed malaria cases ▪ Suspected malaria cases tested ▪ Confirmed malaria cases receiving treatment ▪ Malaria cases (suspected and confirmed) receiving treatment ▪ IPTp3	▪ Facility register ▪ Facility laboratory register	▪ The facility register can be cross-checked against the laboratory register (for microscopy and RDT) for suspected cases receiving a parasitological test. ▪ The facility register can be cross-checked against the pharmacy records for treatments given. ▪ The ANC register can be cross-checked against patient cards for IPT if the patient cards are kept at the health facility. ▪ The HMIS report can be cross-checked against the malaria programme report if data are reported through these separate reports.

Annex 5 | System assessment of functional areas and indicators

Table A5.1 shows the indicators for assessment of each functional area.

Table A5.1 Indicators for system assessment of functional areas

[1] If this vaccine is not used in the country, substitute with another vaccine used in the national programme.
[2] Sampling of health facilities requires stratification by facility type in order to ensure an adequate number of facilities providing HIV/AIDS services.
[3] Sampling of health facilities requires stratification by facility type to ensure an adequate number of facilities providing TB services.

Functional area	Indicator	Administrative level	
		Facility	District
M&E structure and function	Recording the service delivery on source documents assigned to appropriate staff	X	
	Appropriate training of data collection staff	X	X
	Staff designated to review aggregated numbers prior to submission to the next level	X	X
	Staff designated to review quality of data received from facilities		X
Indicator definitions and reporting guidelines	Existence of reporting guidelines	X	X
	Standard indicator definitions (select up to five indicators for review)	X	X
Data collection tools (up to 5 tools)	Existence of instructions on completion of data collection tools	X	X
	Timely completion of data collection tools	X	X
	Availability of blank data collection tools	X	X
	Use of standardized data collection tools	X	
	Data collection tools up to date	X	
	Stock of data collection tools in last 12 months	X	
Reporting forms (up to 5 programme areas)	Use of standardized reporting forms	X	X
	Availability of reporting forms at the facility	X	
	Type of form in use (HMIS or programme)	X	
	Stock of reporting forms in last 12 months	X	
Data quality and supervision	Process within facility for checking the quality of compiled reports	X	X
	Routine accuracy checks	X	X
	Routine consistency checks	X	X
	Facility conducts timely entry and completeness checks	X	
	Written documentation on data quality checks	X	
	Written policy or guidance on data quality procedures	X	X
	Unit receives supervisory visits	X	X
	Unit received supervisory visit in last 12 months	X	X
	Supervisory visits are documented		X
	Monitoring of timeliness and completeness of reporting from facilities		X
	Feedback on reporting provided to facilities		X
Data maintenance and confidentiality	Existence of archived submitted reports	X	X
	Availability of filled/archived data collection tools	X	X
	Ease of retrieval of archived data	X	X
	Appropriate and adequate archive space	X	X
	Access to archive limited to appropriate staff	X	
	Database administration procedure for computerized systems	X	X
	Appropriate back-up of computerized systems	X	X
	Access controls for computerized systems	X	X
	Adherence to policies/standards on confidentiality	X	
Demographic information	Target populations exist for priority indicators		X
	Existence of district map showing catchment areas		X
	Existence of vital events data (births/deaths)		X
Data use	Coverage monitoring for priority indicators		X
	Tracking of progress towards targets		X
	Designated staff to analyse and use data		X
	Evidence-based decision-making		X
	Availability of technical support for data analysis and use		X

Annex 6 | Sampling methods and issues

Sample size calculation[1]

When conducted as part of a health facility assessment, the DQR uses the same sample size calculation as the health facility assessment. The sample size calculation for sampling from a list of health facilities utilized by the service availability and readiness assessment (SARA) methodology is included here as an example.

n = [[(z2 * p * q) + ME2] / [ME2 + z2 * p * q / N]]*d

where:
n = sample size
N = total number of facilities
z = Z value (e.g. 1.96 for 95% confidence interval)
ME = margin of error (e.g. ±15%, value = 0.15)
p = the anticipated proportion of facilities with the attribute of interest (e.g. 0.5)
q = 1-p
d = design effect (a measure of the variation of the sample from simple random sampling estimated from previous surveys).

As can be seen from the formula, the sample size will depend on the desired precision of the key estimates of interest of the health facility survey (including data accuracy) and the acceptable margin of error. Other considerations include the availability of resources and the desired level of application of the estimates (N.B. provincial-level estimates require a greater sample size than estimates for the national level). Note also that the verification factor is a ratio (recounted values over reported values) and will need to be converted to a proportion for this sample size calculation.

The DQR coordination team will need to work with the health facility survey organizers to determine the appropriate sample size for the health facility survey on the basis of the country's priorities with regard to level of application of the estimates, available resources and the precision desired for the estimates.

Weighting of data verification estimates

Data verification estimates based on the sample of health facilities must be weighted to adjust for discrepancies between the sample and the sample frame in the distribution of the number of health interventions of interest (e.g. births attended by skilled health personnel). If the sample is stratified, the stratum-specific estimates of data accuracy should be weighted. In general, the weights for each stratum for a given indicator are computed as the number of events in the stratum in the population divided by the number of events in the stratum in the sample. Since the number of events measured for the sample and in the population (i.e. HMIS) will be different for each indicator reviewed, the weighting of the estimates will need to be conducted separately for each indicator.

The data required for weighting the estimates may often not be available (e.g. the number of events in the population). The HMIS in many countries may not report facility-level data to the national level.

[1] Adapted from WHO's guide on service availability and readiness assessment (Chapter 2). Service availability and readiness assessment (SARA): an annual monitoring system for service delivery. Implementation guide. Geneva: World Health Organization; 2013 (WHO/HIS/HSI/RME/2013/2; http://www.who.int/healthinfo/systems/sara_introduction/en/, accessed 12 June 2015).

For example, to calculate weights for an estimate of accuracy of ANC1 derived from a facility survey in which the sample was stratified by facility type (large facilities and primary care facilities) one needs to know:

- The recounted value of the indicator from the sample aggregated over the domain of estimation of the sample (e.g. region) – for large facilities and for primary care facilities;
- The reported value of the indicator from the sample aggregated over the domain of estimation of the sample – for both large facilities and for primary care facilities;
- The total number of interventions in the population (the HMIS value) over the domain of estimation of the sample – for both large facilities and for primary care facilities (See Table A6.1).

Table A6.1 Example of weighting of stratum-specific estimates for data verification

		Total number of interventions - HMIS (A)	Total number of intervent-ions (REPORTED in sample) (B)	Total number of intervent-ions (RECOUNTED in sample) (C)	Verification factor (D = C/B)	Weight (E = A/B)	Weighted factor (F = D x E)
Strata 1	Primary Health Care Facility	8778	2432	2207	0.907	3.609	3.275
Strata 2	Larger facilities	30483	3597	3145	0.874	8.475	7.410
	Total	39261	6029	5352	0.888	12.084	10.685
			Unweigted verification ratio	0.888	Weigted factor	0.884	

These data are often not available at national level. In many countries the health facility data are not reported all the way to national level. It may not be possible from a given data source to know the type of facilities reporting the data.

If the data are not available they should be estimated. The following methods can be used to estimate weights that are required to derive the composite accuracy estimation for different strata:

- Matching: If the size of the health facility is not known, but other information about the health facility is available and size information about other health facilities is known, one can match the health facility against other similar health facilities where the size is known. This approach assumes that (a) one knows the size (or whatever the stratifier is) of enough health facilities to facilitate matching, and (b) the health facilities for which size information is known are roughly comparable with the health facilities without this information.
- Proxy measures: Other information may be available from health facilities that will allow estimates of the facility weights. For instance, when weighting on the distribution of the indicator of interest in the population one is weighting on the volume of service. Other indicators – such as number of outpatient visits – may be available that may allow for approximation of the volume of service. This could work for high-volume indicators such as immunization but may be less applicable if the indicator relates to something with uneven distribution – such as malaria cases which depend on transmission patterns that are unrelated to service volume.
- Finding the values at subnational level: During the course of the health facility survey it may be possible to acquire the data needed for weighting at the subnational offices of the HMIS – such as regional or district offices to which data are reported from health facilities. These offices will probably require visits during the health facility survey.

Information on stratification variables is also required and should be available from the master facility list. used to conduct the sampling (e.g. information on facility type or management authority).

Annex 7 | Data collection instruments and Excel tool for automated data quality analysis

The data collection tools include the data verification component and the system assessment tool at facility and district levels. Current work is underway to incorporate the DQR in the DHIS 2.0 software, which will benefit countries that are using this software. An Excel tool has been developed to facilitate the annual data quality analysis for those countries using another software system or a paper-based system. The data collection instruments and the Excel tool are not included in this document. They will accompany this guidance document as separate attachments.