

Participatory Data Verification & Improvement Tool

2013

ACKNOWLEDGEMENTS

This tool was developed with input from a number of people within FHI 360. We are grateful for the support, technical leadership, and input from Sam Wambugu, Joshua Volle, Juliana Conjera, Onalenna Serufho, Johannes van Dam, and Tim Mastro. A special thank you to the FHI 360 M&E teams from programs and projects that participated in the piloting of this tool and provided invaluable feedback, specifically: FHI 360 Botswana, FHI 360 Ghana, and FHI 360 Mozambique.

Participatory Data Verification and Improvement Tool

FRAMEWORK AND OPERATIONAL GUIDE FOR IMPLEMENTATION

2013

BY:

Francis Okello
Boniface Kitungulu
Inoussa Kabore
Rajatashuvra Adhikary
Mike Merrigan
Katherine Lew
Gina D. Etheredge



Introduction

FHI 360 recognizes that processes for data quality assurance are an integral part of effective program and project management and comprehensive monitoring system. Data quality assessment (DQA) is an integral component of data quality assurance. It is a snapshot activity in which the quality of routinely collected and reported data is assessed to identify problems or the absence of problems with the data as well as create specific action plans to address any identified or envisioned problems. Implementing systematic and simple procedures for conducting routine DQAs helps to ensure that the six dimensions of data quality, namely accuracy (validity), completeness, integrity, precision, reliability, and timeliness, are upheld.

Beyond FHI 360, donors are increasingly demanding quality data to account for results, measure outputs, outcomes and impact, and to justify funding decisions. Some of them periodically undertake independent data quality audits on the program/projects they fund. In recent years, there is an increased global emphasis on the importance of high-quality data for measuring and documenting achievements as well as continually informing program planning and policy decisions.

The decision to develop this simple and user-friendly tool for assessing data quality came from the recognition of a number of limitations with existing DQA tools, specifically: 1) currently available tools require extensive time and resources, and thus do not lend themselves to rapid, routine and/or widespread use; 2) currently available tools tend to examine an exhaustive number of factors which may or may not be directly linked to data quality; and 3) currently available tools are quite complicated and require special training to implement detailed guidelines and analysis formats.

FHI 360 field and HQ staff collaborated to develop the Participatory Data Verification and Improvement Tool (DVT) to assist countries to employ a more streamlined and targeted approach to improving data quality. The tool is termed 'participatory' to

highlight the need for the various level of actors involved in data collection and reporting systems to work together in a collaborative manner to identify and address root causes affecting data quality across all levels. The DVT facilitates a problem-solving process focused on the accuracy (validity) of routinely reported data and recommends regular use so that stakeholders can confidently use routinely collected data for program and policy planning and strengthening.

OBJECTIVES OF THE DVT

The overarching aim of this tool is to strengthen the quality of data reported by programs and projects.

Specifically, the tool has been designed to:

- Determine the accuracy of historical data reported;
- Facilitate identification of specific root causes of inaccurate data;
- Build capacity of IP and site staff in reporting high quality data;
- Enable rapid, routine, and widespread use.

HOW TO USE THIS TOOL

The DVT is a generic tool developed as a diagnostic exercise for programs and projects to implement at all levels to determine the accuracy of reported data,

conduct root cause analysis, and develop a quality improvement plan to address factors directly affecting data quality. Repeated implementation of the DVT can provide evidence of changes in data quality over time.

This tool can be used as a standalone activity to assess data quality, or can be used as part of FHI 360's M&E System Assessment Tool which provides a comprehensive overview of the strengths and weaknesses of an M&E system. Furthermore, while the tool was designed to improve the data quality in programs and projects supported by FHI 360, either directly or indirectly, it can also be easily applied for use by other partners, projects, and national or sub-national governments.

FRAMEWORK AND METHODOLOGY

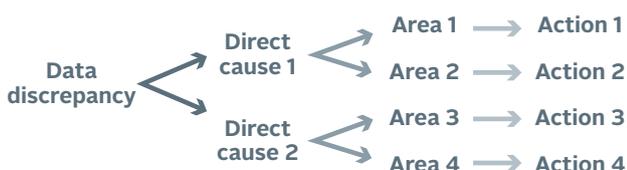
This tool is intended to be used in a participatory process, involving the program team, implementing partner, and CBO/site staff in coordination, planning, and assessment.

Part A. Data Verification involves recalculating results from source documents and uses a simple calculation to determine the accuracy of reported data. The level of accuracy is measured by percentage variance, which is defined as the variance between the recalculated value compared to what was reported. The formula used to calculate variance is as follows:

$$\% \text{ Variance} = \left[\frac{\text{Reported value} - \text{Verified value}}{\text{Verified value}} \right] \times 100$$

Where a positive variance (+) reflects over-reporting and negative (-) variance reflect under-reporting. A threshold of +/- 5% variance has been suggested for high quality data.

Part B. Data Improvement is intended as a guided process to identify root causes affecting data quality and to develop immediate- and long-term plans for addressing threats to data quality.



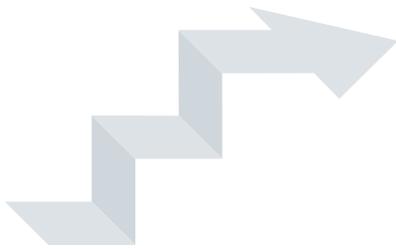
It is important to note that while the foundation for assessing data quality lies in determining the accuracy (validity) of the reported data values, it is expected that threats to other dimensions of data quality, specifically completeness, integrity, precision, reliability, and timeliness, will be uncovered during this process. For example: It may not be possible to verify certain indicators for a given month due to the incompleteness or untimeliness of the report.

DURATION

The process for completing data verification varies depending on the size of the program and number of sites being assessed. The team should anticipate spending between one half-day to one full day with each site, with the variation in time depending on the number of indicators being assessed and period of time being assessed.

FREQUENCY OF IMPLEMENTATION

This tool should be implemented during the first year of project start-up, within 6 – 12 weeks after beginning data collection. Repeated implementation would ideally occur once each quarter per site throughout the life of the project (LOP). The frequency of implementation can be reduced once 90% or more of indicators assessed have variance less than +/-5% and this has been sustained for 6 or more months. Implementation at the site level can be conducted through a mixture of internal self-assessments combined with less-frequent external assessments by the IP or program team.



Guidance on Data Verification and Improvement Steps

STEP 1: PREPARATION

- a. Team formation** – Ideally, the verification team comprises the program M&E lead, program/technical staff, and representatives from IPs and sites with reporting duties. Where relevant, team leaders should be selected to lead individual teams conducting data verification at sites. These will report to the person leading the data verification exercise, generally the program M&E lead. A timeline for implementation should be developed and agreed upon.
- b. Identification of relevant source and secondary documents** – Documents may include the project program/proposal, M&E plan/PMP, M&E data collection tools, service delivery guidance documents, donor reports, and national and global indicator guidelines (e.g. PEPFAR, UNGASS, etc.).
- c. Selection of indicators to verify** – All required global indicators should be selected for verification, followed by prioritization of country-specific, national-level indicators, and critical program/project-specific indicators. A rationale for selection of indicators should be specified.
- d. Selection of sites to visit** – Select the IPs and sites that will be visited based on agreed upon criteria. Priority or higher volume sites may need more regular visits.
- e. Notification of selected sites** – Sites should be provided in advance a copy of the tool and advised on what source documents to prepare and which staff members should be involved (e.g. program/project coordinators, outreach workers, and/or data clerks).

Factors to consider in selecting sites to implement data verification:

- Previously identified issues with submitted data
- Sites that have not been assessed in the preceding 6-12 months
- Geographic coverage
- Scope of site services

STEP 2: IMPLEMENTATION

- a. Introductory briefing with IP/site staff** – The purpose of the briefing is to orient the IP/site staff on the purpose of the data verification, methods, and tool used. Furthermore, staff from the visited IP/site should participate in the verification; after one or two applications of the tool, IPs/sites should be in a position to lead the assessment to promote sustainability.
- b. Review of source and secondary documents** – Relevant data collection tools (e.g. registers), reporting or summarizing formats, and other identified secondary documents should be reviewed at each site.
- c. Re-calculation of indicators** – Using reports submitted by the site as the source of the denominator (reported value), variables should be recalculated for each indicator verified from source documents (e.g. registers, client intake forms) and then compared to what was reported.

Example

Using the formula below, the site underreported indicator 'A' by 10%:

$$\% \text{ Variance} = \left[\frac{\text{Reported value} - \text{Verified value}}{\text{Verified value}} \right] \times 100$$

Indicator (Include any required disaggregation, e.g. age, sex)	Source Document (Include the row, column, etc. of the source, if appropriate)	Month 1: ____ / ____		
		Verified value	Reported value	Variance (%)
Indicator 'A'		110	100	-9.1

For purposes of this tool, we recommend that data from the previous three months be used for verification (this is linked to the recommendation for quarterly assessments at each site). Variance (%) is generated for each indicator and overall. For each indicator where variance was greater than +/- 5%, further analysis is warranted.

d. Corrections of the error and reporting the correction – Errors of any magnitude should be corrected, and this may entail several steps depending on national protocols. Typically this stage starts with making corrections to the reported value, submitting a revised report, and ensuring there is a sufficient paper trail outlining when the correction was made, by who and where.

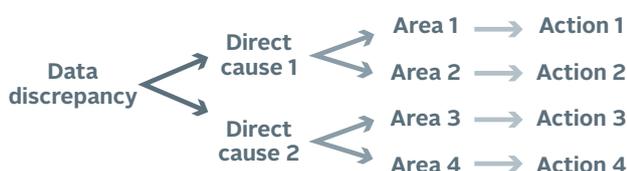
e. Root cause analysis – For indicators with more than +/- 5% variance, root cause analysis can be conducted to identify the specific stage or process where the data error was introduced.

Examples may include calculation error in totaling the results from the source document, transcription error when entering data into database, etc. The data flow and reporting system may need to be reviewed with site staff to understand where the error was introduced.

f. Identification of areas for intervention – These corrective measures are the immediate efforts necessary to rectify the inaccurate data (e.g. correcting source document or register, followed by correcting the program database and reporting revised data to donor).

g. Action planning and follow-up – Follow-up steps refer to longer term measures taken to prevent the specified error(s) from recurring. Examples may include: mentoring, technical assistance, training, improved supervision, distribution of guidelines, etc. These measures should be specific, responsive, and time-bound; furthermore, responsibilities should be indicated.

h. Debrief with IP/site staff – The purpose of the debrief is to present findings and immediate recommendations to IP/site managers. The debrief provides an opportunity for IPs/sites to provide valuable input as well as to obtain IP/site support for the important next steps.



Example

Indicator(s) with variance +/-5%	Identified data issue	Corrective measure taken	Follow-up steps
Indicator 'A'	Transcription error from the session attendance forms to the daily prevention registers	We updated the data based on daily prevention registers based on the daily session attendance forms	Mentoring for program manager responsible for compiling daily registers
	On the session attendance forms there is no provision for individual sessions (these were not verified)	Reports for individual sessions reduced to what can be verified from source documents	Revise the session attendance form to include the following: <ul style="list-style-type: none"> • One on one sessions • Type of session Age/ Gender Venue • Date of session • Name of the implementer Distribute revised forms and orientate field implementers during the monthly meeting

STEP 3: ANALYSIS AND FOLLOW UP

a. Prioritization of recommendations and work planning

– Findings from all sites should be collated, and the verification team should agree on recommendations, prioritize the most important, and prepare a work plan with sub-activities, responsibilities, and timelines for each key recommendation. It is important during this step not to lose any site-specific issues that may not be represented amongst the majority of other sites.

b. Communication of findings – Each participating site should receive a site-specific, completed version of the tool with findings and agreed upon next steps.

c. Follow-up with sites – All sites visited should be followed-up within three months of the verification to support the necessary corrections to data or data systems.

d. Data quality tracking – Results from the verification should be entered into a data quality monitoring database and analyzed in a time-series to illustrate improvement over time or persistent gaps, or compared between different IPs/sites to highlight relative strengths and weaknesses in data quality. Furthermore, the program team and IPs should maintain a calendar to track the dates of data verification activities at the site level to enable snapshot analysis of coverage and frequency.



FHI 360 HEADQUARTERS

2224 E NC Hwy 54
Durham, NC 27713 USA
T 1.919.544.7040
F 1.919.544.7261

WASHINGTON DC OFFICE

1825 Connecticut Ave, NW
Washington, DC 20009 USA
T 1.202.884.8000
F 1.202.884.8400

ASIA-PACIFIC REGIONAL OFFICE

19th Floor, Tower 3
Sindhorn Building
130–132 Wireless Road
Kwaeng Lumpini, Khet Phatumwan
Bangkok 10330 Thailand
T 66.2.263.2300
F 66.2.263.2114

**SOUTHERN AFRICA REGIONAL
SUPPORT OFFICE**

Ground Floor, Block E
Hatfield Gardens
333 Grosvenor Street
Hatfield, 0083 Pretoria
T 27.12.423.8600
F 27.12.342.7005

REGIONAL TECHNICAL HUB

Valley Road
Chancery Building, 2nd Floor
Nairobi, Kenya
T 254.20.271.3911/5
F 254.20.282.4170

www.FHI360.org