Data Verification and Improvement Guide

VERSION 1.0

AUGUST 2021

fhi360

THE SCIENCE OF IMPROVING LIVES

Suggested citation: FHI 360. FHI 360. Data Verification and Improvement Guide Version 1.0. Durham (NC): FHI 360; 2021.

Acknowledgments

This guide was developed by Caximo Caximo, Pradeep K. Thakur, and Linda Muyumbu with advice from staff members within FHI 360. We are grateful for valuable support, technical leadership, and contributions from Navindra Persaud and Gina Etheredge. We extend special thanks to the bilateral and the Meeting Targets and Maintaining Epidemic Control (EpiC) programs and projects that shared data verification tools and guides, which informed this guide's development. We also thank the FHI 360 team that developed the Participatory Data Verification and Improvement Tools v. 2.0 (2019) and the FHI 360 teams that piloted the accompanying tools and provided invaluable feedback.

Table of Contents

Glossary of Terms

This glossary, mainly informed by definitions of data quality by MEASURE Evaluation, provides a shared understanding of the multiple and often conflicting interpretations of data quality dimensions and related terms.

Term	Meaning
Data Quality	Data is generally considered high quality if it is "fit for [its] intended uses in operations, decision-making, and planning."
Data Quality Assurance	A process for defining the appropriate dimensions and criteria of data quality, and procedures to ensure that data quality criteria are met over time. It involves a process of data profiling to unearth inconsistencies, outliers, missing data interpolation, and other anomalies.
Data Quality Assessment	A review of program or project monitoring and evaluation (M&E)/Information Management (IM) systems and data to ensure that the quality of data captured by the M&E/IM systems is acceptable.
Data Verification	A process of assessing the data fitness of FHI 360 programs and projects to determine if it meets U.S. Government (USG) and other funders standards for reporting, decision-making, and planning. It involves comparing reported data with the original records.
Accuracy	The degree to which data accurately reflects an event or object described. For example, if a beneficiary's age is 32, but the system says she/he is 34, that information is inaccurate.
Reliability	Data is considered reliable when repeated measurements using the same procedures get the same results at all places and all times.
Consistency	Data are consistent when the value of any given data element is the same across forms, applications, and systems.
Completeness	Data is considered "complete" when it fulfills expectations of comprehensiveness, that is, all desired data elements are collected. For example, you require a beneficiary to supply his or her full name for HIV testing but then only record their first name, then the data is incomplete.
Timeliness	The degree to which data are current and available for use as specified and in the time frame in which they are expected.
Integrity	Data have integrity when the system used to generate it is protected from deliberate bias or manipulation for political or personal reasons. In other words, the data is not made up or changed.
Uniqueness	"Unique" means there is only one instance of a particular data point appearing in a database.
Confidentiality	Clients' personal data are not disclosed inappropriately, and data in hard copy and electronic form are treated with appropriate levels of security, e.g., kept in locked cabinets and in password-protected files.
Beneficiary	Any verifiable individual who is targeted by a program/project for specific services and who accesses the specific services.
Beneficiary Verification	Steps taken to confirm that an individual has received a specific service; may be accomplished through a physical visit or telephone call.

1.0 Introduction

FHI 360 recognizes that processes for data quality assurance are an integral part of an effective program/project monitoring system. Data verification (DV) and improvement is an integral component of data quality assurance. Data quality assurance (DQA) is an activity in which routinely collected and reported data is assessed to determine whether it is fit for purpose, identify potential quality issues and their causes, and develop action plans to address any identified problems. This process serves to confirm the degree of alignment of the reported data with the services provided and strengthen confidence when using routine service delivery data to make decisions. Implementing systematic and simple procedures for conducting routine DV helps to ensure that the six dimensions of data quality, namely accuracy (validity), completeness, integrity, precision, reliability, and timeliness, are upheld.

1.1 Rationale for Data Verification Guide

Donors in health and development programs are increasingly demanding high-quality data to support adaptive planning and programming, account for results, and measure program performance (outputs, outcomes, and impact). Some donors periodically undertake independent data quality audits on the programs/projects they fund.

The decision to develop this simple and user-friendly guide along with the accompanying *Data Verification and Improvement Tools* for assessing data quality came from the recognition of the increasingly complex data quality challenges and several limitations with existing DV guides/approaches. Currently available tools examine many factors, all of which may not be directly linked to data quality. Moreover, these approaches require extensive skills, time, and resources, and do not lend themselves to rapid, routine, and/or widespread use. Second, they are quite complicated and require special training and skills to implement. Third, data management systems have become increasingly electronic, with little or no paper trail to audit. Fourth, they all require a site visit to conduct the verification and do not make provisions for those unique situations such as the ongoing COVID-19 outbreak or security challenges that may hinder travel and physical access to the sites.

Building on extensive experiences in different settings and situations, FHI 360 developed this *Data Verification and Improvement Guide* and tools to assist projects and partners to meet these challenges and employ a more streamlined and targeted approach to improving data quality. The guide emphasizes the need for the various actors involved in data collection and reporting to work together to identify and address root causes affecting data quality across all levels. In addition to the commonly used record verification, it also includes processes for physical verifications of program/project beneficiaries, guidelines for sample size calculations that are needed in special circumstances, and steps for conducting DV in contexts of limited/restricted access to program/project sites and physical interactions with partners/beneficiaries. It also provides recommendations on frequency for conducting data quality assessment in different contexts.

1.2 Objectives of the Guide

The overarching aim of this guide is to strengthen the quality of data collected, analyzed, reported, and used by programs and projects globally.

Specifically, guidance is provided on how program/projects should plan and implement DV to meet the following objectives:

- Determine the accuracy of reported data
- Facilitate identification of specific root causes of inaccurate data
- Collect information to redesign data collection and management systems and define capacity-building interventions for partners and site staff to report high quality data
- Generate metrics from repeat assessments for tracking trends and making comparisons across different sites
- Support partner profiling that will help prioritize technical assistance (TA) based on performance scores
- Guide development of specific quality improvement plans to improve data quality

1.3 How to Use this Guide

This *Data Verification and Improvement Guide* and tools may be used by programs/projects and other implementing partners to maintain the accuracy of reported data, conduct root cause analysis for any problems identified, and develop a quality improvement plan. Advice on frequency of DV is provided. Repeated DV can provide evidence of changes in data quality and improvements in information systems over time. The tools are recommended for use during standalone DV and may be complemented with FHI 360's *Monitoring and Evaluation Systems Assessment Tool* (M&ESAT), which provides a comprehensive overview of the strengths and weaknesses of a program or project's M&E system. This guide may also be used by other partners, projects, subnational, and national governments.

Programs and projects are implemented in contexts that vary by geographic area. In addition, the specific goals and objectives, strategies, and project dynamics may vary over time. While the overall DV framework should be followed by all projects and programs, adjustments may be needed to suit the specific objectives and context of each context.

2.0 Methods

This section provides guidance on the methods for planning, implementing and reporting on DV activities.

2.1 Verification Schedule

Routine verification by project staff: For all new programs/projects, DV should be implemented at all reporting sites in the first quarter after start-up to establish a baseline. This initial assessment will also enable the team to detect any issues with information systems that need to be addressed and identify threats to data quality. Subsequently, DV should be conducted at least quarterly to ensure that data reported to the donor is of good

quality. For projects that collect data monthly, DV should be done at each site at the end of each month.

External data verification: Each project should receive a comprehensive external DV in the first year and at periodic intervals thereafter based on the findings. This could be conducted by the program/project backstops or an external consultant and will help to validate the results of the DQA conducted by the project staff. Follow-up external assessment should be planned based on the findings of the initial assessment. Table 1 below provides a recommended follow-up schedule.

Table 1. Frequency of follow-up data verification by HQ/regional backstops

Type of sites/partners/projects (DQA score)	Follow-up external DV
Low-risk site/partner (95%–100%)	1 year after last DV
Medium-risk site/partner (80%–94%)	6 months after last DV
High-risk site/partner (less than 80%)	3 months after last DV

Existing projects: Existing programs/project that have no baseline verification results and have at least two years to go should implement a comprehensive external DQA within the next six months or at least once before closeout.

2.2 Preparation

Data verification is a participatory process, involving the program team, implementing partner(s), and site staff, and thus requires careful planning and coordination.

Team formation and size: Ideally, the verification team is comprised of the project M&E lead, program/technical staff, and representatives from implementing partners and sites with data collection and reporting responsibilities. This is critical to ensure ownership and buy-in of the end results of the verification from the implementing partners. If multiple sites will be visited, leaders should be selected for individual teams conducting DV at each site. These team leads should report to the person leading the DV exercise, generally the program M&E lead.

Based on the number of sites to be covered, the sample size, and the indicators to be covered, the team needs to be precisely defined, and team members need to be allocated to the sites to finalize the work within the agreed-upon time.

Definition of intended objectives and expected results: The team should define the objectives of the exercise and clearly articulate the expected output. This could be DV under normal circumstances or verification of beneficiaries or both. Clear objectives and expected outcomes will influence the entire process and are key for successful completion of the activity.

Selection of indicators to verify: All donor-required indicators should be selected for verification. This should be followed by prioritization of country-specific, national-level

indicators, and critical program/project-specific custom indicators and problematic indicators such as those newly introduced or where definition is not clear/well understood generally and observed to have data quality issues. The rationale for selection of indicators should be specified in the protocol. For verification of beneficiaries, the team should identify which service will form the basis of the verification. A precise definition of each of the services should be created to guide the beneficiary verification.

Definition of verification period: The team should define the reporting period for which data will be verified. When beneficiary verification is being done it should cover the services accessed within one month of the reporting to minimize recall bias.

Selection of sites to visit: Select the implementing partners and sites that will be visited based on agreed-upon criteria. For projects with a small number of reporting sites, data from all sites should be verified before it is reported. Projects/programs with a large number of sites or have conditions in which it is impossible to conduct verification at all sites before reporting can consider conducting the routine verification at those sites that contribute 80 percent of program/project results for the reporting period. This provides a reasonable coverage of DV.

In instances where external verification is conducted, projects may also want to consider selecting a statistically representative sample of sites and then make inferences from the results. Please consult with your strategic information (SI) backstop or support team for guidance on how to determine the sample size for the DV.

Identification of relevant data sources and secondary documents: Determine the data sources needed to verify the data for selected indicators. This may include standard national tools such as registers, program data collection tools, and electronic databases. Beneficiaries as well as service providers are also relevant sources of data.

Data collection tools and user guidelines: Relevant standard tools and checklists are provided here, however, teams are encouraged to adapt these or include additional tools they may find useful and relevant to their context and objectives of the DV. It is important for the team to have a common understanding of the tools as part of preparation to ensure the verification is done well; plan to provide an orientation on the tools.

Define the sample size (where applicable): For external assessments in which visiting each site may be too expensive or impractical, it is important for the team to define, during the preparation phase, the sample size they need to cover. This should be developed based on the guidelines provided and the objectives of the verification. Please consult with your SI backstop or support team for guidance on how to determine the sample size.

Define all ethical issues involved: The team should outline how they will ensure ethical issues are addressed during all the steps of DV. Addressing ethical issues is especially important when beneficiaries are being contacted.

Prepare logistics for the exercise: Prepare materials, bookings, and transportation of other resources in advance to avoid compromising implementation of the field activities.

Notification of selected sites: Notify sites about the proposed DV in advance of the visit. Notification should ideally be two weeks in advance of the visit and should include the dates, duration of the visit, what source documents will be needed, and which staff members should be involved (e.g., program/project coordinators, health records officers, and/or data clerks).

Training for the verification team: Training is crucial for better understanding of the methodology as well as all ethical aspects. Projects/programs should decide the depth and length of the training. It may be a brief orientation session for a few hours or a more structured training.

2.3 Implementation at Site

Introductory briefing with partner/site staff: When the team arrives on site, an introductory briefing should be conducted for the site/facility staff. The objective is to discuss the purpose of the assessment, methods, tools to be used, benefits to the site/partner, and staff members who should participate. Staff participation is critical to ensure ownership of the process, the result, and the improvement plan that will be developed at the end of the exercise. The briefing meeting should cover, at minimum:

- Objective of the exercise
- Scope of the exercise
- Target group and indicators
- Sample size
- Timeline of the exercise
- Expected participants
- Responsibility of the parties
- Individual steps of the process

Participation of site/partner staff also serves to build capacity. After one or two applications of the tool, implementing partner and site staff should be able to lead the assessment, promoting sustainability.

Conduct data verification: Discuss how data is aggregated and reported at each implementing partner/site and the sources used to report data for the defined period of verification. Review source and secondary documents, discuss with implementing partner/site service providers and fill in the data collection forms. Relevant primary data collection tools (e.g., registers, enrollment forms), reporting or summarizing forms, and other identified secondary documents should be reviewed at each site. ***See details for conducting the verification in section 2.5–2.9 below.***

2.4 Data Analysis and Feedback at the Site

Before leaving the site, analyze the information collected to show the accuracy of reported data as verified from source documents. The *Verification Factor* is used to judge the program/project or site as either passing or failing in data quality. It is important for the team to review the findings with the staff while on site. The dashboard provides graphs to facilitate quick analysis of results that are useful for providing feedback and allow for deeper discussion before the team leaves the site. Errors during data collection can be addressed, data archiving issues can be corrected, and adjustments to the data can be made to define more realistic improvement actions.

2.5 Assessing the Quality of Reported Aggregate Data

Data is re-counted from source documents and compared with data reported for the period under review and entered into the ***Part A: Data Recounting Checklist*** **of the MS Excel DVI tool a screen shot of which is shown in *Appendix* I.** A simple calculation using the formula below determines how close the verified data is to the reported value.

$$Verification\ Factor = \frac{Verified\ value}{Reported\ Value} * 100$$

For each indicator, the reported value will be the sum of the values reported for the three-month period (column E of *Part A: Data Recounting Checklist* Appendix I), while the verified value will be the total of the re-counted values for the three-month period (Column F of *Part A. Data Recounting Checklist* Appendix I). The *Data Recounting Checklist* is configured in MS Excel to auto-calculate the verification factor, and the results can be found in column G.

Interpretation of verification factor results:

- Score > 100% = over reporting
- Score < 100% = under reporting

Scores between 95%–105% are within the acceptable threshold of data quality.

Based on the results from all the verifications, the sites/partners will be categorized according to one of three levels to indicate how much technical support will be needed to address any data quality gaps identified.

1. **Low-risk partners/sites—**Have an average verification score between 95% and 100%; considered lowest priority for project/program interventions/headquarters (HQ) technical assistance (TA).
2. **Medium-risk partners/sites—**Have an average verification score between 80% and 94%; considered second priority for interventions, HQ TA, and other program support.
3. **High-risk partners/sites—**Have an average verification score of less than 80%; considered top priority for interventions, HQ TA, and other support.

Trends in the verification factor should be tracked over time to determine progress in improving data quality and performance of the M&E systems over the life of the project.

2.6 Assessing the Quality of Data on Recording Forms

The second step in DV includes a review of data recording forms/tools such as patient charts, case files, intake forms, and registers. Review of the data recording forms, which is conducted as part of verification, assesses the five dimensions of data quality namely validity, completeness, timeliness, integrity, and consistency. The assessment team selects a pre-determined sample of primary data recording forms to check and records "Yes" or "No" for each of the five dimensions for each form reviewed.

Selecting recording forms/tools for review: For verification of the five dimensions of data quality, the team should review 10 percent of the data recording forms for the period covered during the DQV. For instance, if 350 HIV tests were reported then 35 of the testing forms should be randomly selected for review.

The results for this verification will be recorded in the ***Part B: Recording Tools Verification Checklist*** **as illustrated by screenshot in *Appendix II*.** It is Part B of the MS Excel DVI tool**.** Each row represents each form verified, while the columns C to F represent each data quality dimension being verified. For each of the dimensions, if the particular form verified complies with the quality requirements, enter "Yes" and, if not, select "No." To compute the percentage for each dimension, the system sums the number of forms with "Yes" among the forms reviewed and this will be the numerator for "# forms with valid responses." The denominator will be the total number of forms reviewed for each of the dimensions. The percentage is calculated for each of these five dimensions of data quality. The calculations of verification factor will be done using the following expression:

$$Verification\ (for\ each\ dimension) = \frac{\#\ of\ forms\ with\ valid\ responses}{\#\ of\ forms\ reviewed} * 100$$

Interpretation of data quality dimension verification factor:

- 95%–100% = acceptable (high data quality)
- < 95% = low data quality

2.7 Assessing the Quality of Data from Beneficiaries

Traditional verification using re-counting of physical forms exclusively can be misleading as it does not always detect fraud, in which providers intentionally report services that were not truly provided to beneficiaries. In addition, discussions with beneficiaries can provide insights on the quality of services provided beyond just a count of the number of units of services provided. Projects are encouraged to include beneficiary verification in their data verification processes. All programs/projects with a community-based service delivery component are encouraged to integrate beneficiary verification in their DV activities. This includes "physical or virtual" verification of a random sample of beneficiaries.

The objectives of the beneficiary verification are usually threefold:

- To confirm the existence of program/project beneficiaries
- To confirm whether the reported services were received by the beneficiaries
- To collect qualitative information on the quality of reported services

Physical verification of beneficiaries through home visits

If verification of beneficiaries is planned through visiting them either at home or other sites, then select 5 percent of the individuals who were reported during the period covered by the DQA. Once the list is finalized, then all selected clients should be visited either at home or another convenient place to verify whether they received the services indicated in the facility records. When contacted the person could be asked to verify:

- Sex
- Dates when they were last seen by the service provider
- Place/physical location where services were received
- Information or services they received

Information provided by the client should be entered into the MS Excel DVI tool **Part C-1 which is the *Physical Verification of Beneficiaries Checklist* illustrated with a screen shot in *Appendix IV*.** Some probing may be required if beneficiaries do not recall the name of the service provider. In this case, a description of the individual might help them remember. Also, be prepared to give a more detailed description of the service for the same reason.

Data from this verification will be aggregated and analyzed using the following expression:

$$Beneficiary\ Verification\ Factor = \frac{\#\ of\ beneficiaries\ confirmed}{\#\ of\ beneficiaries\ selected} * 100$$

Interpretation of beneficiary verification factor:

- 100% = Accurate data
- < 100 % = Inaccurate data

To compute the beneficiary factor, use Part C-1 of the MS Excel DVI (Screenshot Appendix IV). Each row will represent one beneficiary verified. The team will fill in columns E to H, with background information from what was reported by the assessed partner, and will check if all the reported background information is consistent with what is found during the verification. If all variables from E to H are confirmed, team will select “Yes” in column I of the checklist. If at least one information detail cannot be verified, the team will select “No." Therefore, the numerator of the beneficiary verification factor calculation expression above will be the total number of beneficiaries with “Yes” in column I, while the denominator will be the sum of beneficiaries selected.

Telephonic checks with beneficiaries/clients

When it is not possible to visit beneficiaries, then virtual verification via telephone or other means could be attempted. If remote verification will be conducted via telephone or other means of communication, then select 10 percent of the individuals who were reported during the period covered by the DQ verification. A higher sample size is recommended for virtual verification to compensate for missing and invalid contact information. Once the list is finalized, then the DV lead and site supervisor should call each beneficiary and introduce themselves. The client should then be informed that this is a program quality call, and the program is interested in understanding the services which they recently received. If the beneficiary would prefer to discuss in person, then provide this option and set up an appointment for the visit. It is important to assure clients about confidentiality, and ensure that data verification is being conducted with the right person. As such, verify that you are talking to the right person, and ask if they agree to answer a few short questions in a five-minute call. The sample script below may be adapted for use.

Hello! My name is (caller's/visitor's name), *and I work for FHI 360* (indicate project name and work location). *I am working with* (supervisor name, partner name) *as part of our internal program for strengthening data quality assurance. For this call/visit, I am going to ask questions to verify data on services you recently received through our program* (indicate project name and site location), *which we extracted from the program database. Please note that this verification activity is voluntary and includes a discussion with you* (which we are conducting now). *The information is confidential and will only be used to strengthen data quality and service provision. Is this a good time to discuss? Let me know if you have any questions before we proceed.*

Once the individual is contacted, understands the purpose of the call, and agrees to participate, then ask for the following information:

- Sex
- Date when individual received service
- Place where services were offered
- What information or services they received

It is important that open-ended questions like this are used. Information provided by the client should be entered into the **Part C-2 *Electronic Verification of Beneficiaries Checklist* of the MS Excel DVI (Screenshot *Appendix III*).** Note there may be some probing required if beneficiaries do not recall the name of the service provider. In this case, a description of the individual might help them remember. Also be prepared to give a more detailed description of the service for the same reason. The proportion of persons verified is calculated using the formula above.

2.8 Data Verification for Projects with Electronic Systems

This section outlines procedures for verifying routine data collected exclusively through electronic systems. Increasingly, data collection is moving to electronic platforms including

smartphones, tablets, and laptops/personal computers. Electronic data capture has several advantages including increased efficiency, improved aggregation, and reduction of errors in data entry. On the other hand, in the absence of paper data reporting forms, verification can be challenging. Some quality-related advantages of electronic versus paper systems include:

1. **Built-in validation checks for devices and databases:** Certain validation checks including range and consistency checks can be programmed into data entry devices and the program database to guard against data entry errors and falsification of information.

2. **Backup data collection systems and data security:** Electronic data is easy to back up, and is not as vulnerable to environmental degradation or outside manipulation as hard copy files and folders. Teams can ensure data is regularly backed up, password protected, and off-site storage is in place for additional security.

3. **Automatic aggregation of data:** Electronic data does not need to be manually transcribed or aggregated and is also free of legibility issues.

4. **Savings with paper costs:** In certain contexts, cost of paper forms consumes a significant part of project budgets; this is aggravated by the fact that these costs are ongoing as forms need to continually be reproduced. Electronic systems tend to have stabler costs, related to initial development and updates when required, which makes the costs much lower, depending on the specific project.

Even with these advantages in mind, the shift to electronic program data poses special challenges for validating data and requires new and strengthened procedures to prevent data fraud and ensure quality. Without a paper trail, verification of electronic data poses additional challenges due to the absence of physical records to review. The following measures should be used to review and validate the quality of data in electronic systems that do not have a paper trail. This guidance assumes the program/project is supporting the delivery of services to beneficiaries and it is possible to verify the receipt of these services. All steps outlined should be conducted together with the implementing partner supervisor or program manager, who can introduce themselves but otherwise remain a silent observer. Projects are encouraged to include partner staff in this exercise to improve acceptability of the results. They may also help with local terms or language related to the program that will help validate the reported service episode.

Data verification at service provider level: It is critical to initially verify with a sample of service providers whether activities reported took place. To minimize recall bias, attempts should be made to verify data not later than a month after the service was reportedly provided. From the electronic register/system, the DV team leader should **select a random sample** of service delivery personnel per site (or per geographic location) and, for each individual, extract electronic records from their **last five reported client interactions**. The next step is to summarize the services recorded for the selected individuals. The service provider is then either visited at their place or contacted virtually via telephone, WhatsApp,

Skype, etc. along with their supervisor. After contact is made with the service provider, they should be informed that the purpose of the visit/call is to conduct a data quality meeting. The supervisor and DV lead should also explain that individuals were randomly selected and reassure the individual that honesty is important because the results will be used for system strengthening, not punitive measures. Below is some sample script that could be used for the introduction.

Hello! My name is (caller's name), *and I work for FHI 360* (indicate project name and work location). *I am working with* (supervisor name, partner name) *as part of our internal program for strengthening data quality assurance. For this call/meeting, I am going to ask questions to verify data on services you recently provided to your most recent five clients, which we extracted from the program database. Please note that this verification activity includes a discussion with you* (which we are conducting now); *and may include phone calls or visits with clients who received services. Let me know if you have any questions before we proceed.*

After the introduction, focus on the most recent five clients extracted from the electronic register and ask the individual who provided the service for the following details for each of the clients:

- Demographic characteristics
- When and where the client was last seen
- What services were provided on that day

Information given should be entered into the ***Electronic Verification of Beneficiaries Checklist*** **in screenshot** ***Appendix III*****.**

If any of the information provided by the health care worker for any of the clients does not match what is in the database then proceed to conduct a beneficiary verification. Even if the information does match, the team may decide to verify a percentage of interactions by contacting beneficiaries directly. Please consult ***Section 2.7*** above for details regarding the steps for contacting clients. If virtual verification is attempted and the phone numbers on record are not working for one or more clients, then continue adding random clients until the sample of five is reached.

2.9 Virtual Data Verification

This section outlines procedures for verifying routine data during unique situations such as the ongoing COVID-19 outbreak or security challenges that may hinder travel and physical access to the sites and program beneficiaries. The steps for a virtual DQA will follow the same sequence described above in Section 2:

- Preparation
- Implementation at site
- Data analysis and feedback at the site
- Assessing the quality of reported aggregate data

- Assessing the quality of data on recording forms
- Assessing the quality of data from beneficiaries

For virtual data verification, all steps above should be done using Microsoft Teams or a meeting platform appropriate for the context. Prior to the virtual assessment meetings, the DQ teams will send a meeting invite to each partner stating the objectives, scope, and dates of the exercise; target group and indicators; necessary source documentation to present during the scheduled virtual meetings; timeline of the exercise; and expected participants. The length of meetings will vary depending on the volume of data and internet connectivity challenges; however, a maximum of two full working days is estimated to complete this task (see Figure 1).

The DV team will assess the quality of reported data by asking the local partner to share source documents (registers) by attaching them in the chat box or showing the same on video. The team will ask the program/project team to re-count values of selected indicators and provide the aggregate for confirmation against the report for the assessment period and recorded on an electronic copy of the ***Data Recounting Checklist*** **in screenshot in *Appendix I*.**

For the data recording forms, the verification team will read out preselected unique identifier codes and request the program/project team to find the forms. The program/project team will read out the values according to details in ***Appendix II: Recording Tools Verification Checklist*** and attach the forms (if electronic) in the chat or show the details on video.

All virtual DQA should be accompanied by a beneficiary check, which should be done following the guidance provided under “telephonic checks with beneficiaries/clients” in Section 2.7 above. Record all virtual DQ verification exercises conducted. All recordings should be filed in the shared drive accordingly, as per the *Digital Filing System SOP* of the program/project.

Figure 1: Virtual Data Quality Verification Steps

FHI 360 SI/M&E Unit Level	DQV Steps	DQV Tools	Local Implementing Partners Level
STEP 1: Preparation 1. Form a data verification (DV) team (M&E and program staff), with assigned team lead for the virtual DQV. 2. Select PEPFAR and custom indicators and determine the reporting period under assessment.  3. Determine the source documents that will be used for verification including electronic systems where applicable. Prepare data sources to ensure access during the assessment. 4. Randomly select sites that contributed to 80% of quarterly achievements for the period of assessment. 5. Send a Microsoft Teams meeting invite to respective partners, stating objectives, dates, length of meetings, and necessary source documentation for period under assessment to present.	**Before start of meeting:** Introduce the exercise and objectives and expectations. Required electronic (soft) copies/scanned of forms/registers can be attached in the chat as the session goes on or shown on screen. Ask for copies of sources to be sent via email after the DQV. Record all virtual DQV exercises and ensure all members taking part are notified. Videos will be required for the virtual data verification. **STEP 2: Data Verification** 1. Assess the quality of reported aggregated data. The program team will re-count data from source documents for selected PEPFAR/custom indicators. Record the results in the ***Data Recounting Verification Checklist.*** 2. Compare the re-counted value with the number reported in DATIM/summary report and calculate variance to determine how close the verified data is to the reported value. 3. Review 10% of a pre-determined sample of the data recording forms/tools such as patient charts, case files, intake forms, and registers for completeness, validity, timeliness, integrity, and consistency. The DV team will provide the unique identifiers for forms and record the results in the ***Recording Tools Verification Checklist.*** 4. Assess the quality of data from a pre-determined sample of beneficiaries by calling. Follow the standard calling script provided to introduce the team, explain the exercise, and obtain consent. Verify the data and record in ***Electronic Verification of Beneficiaries.*** 5. Analyze scores; discuss and agree on findings. **STEP 3: Data Quality Improvement Action Plan** 6. Identify indicators with issues, establish the root causes, discuss and agree on improvement actions. 7. Jointly develop Data Quality Improvement Plan and follow-up actions using the template. 8. Decide the next data verification date according to guidance of DQV SOPs.	 The DQ team displays the DQ verification tool on the screen and completes it according to the observations and findings from LIPs. **Action Plan**	Local implementing partner (LIP) staff being interviewed, presenting, and re-counting relevant source documents 

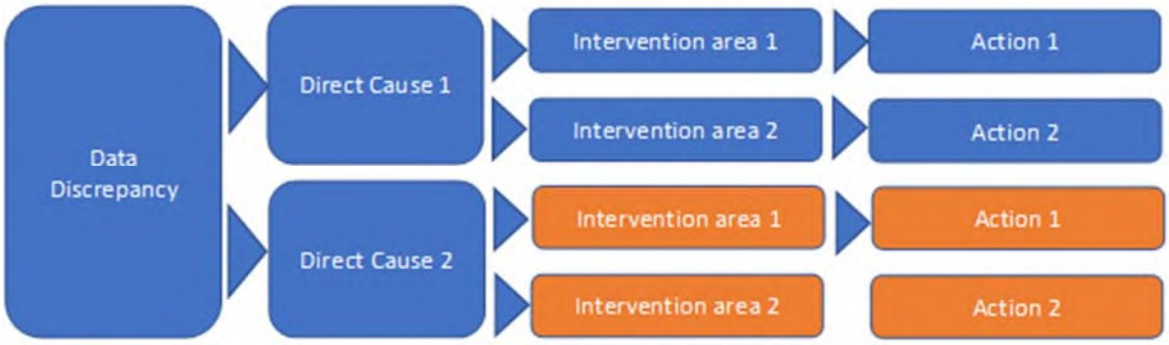
3.0 Steps for Data Quality Improvement Planning

This section outlines the process of developing the data quality improvement plan. This is a guided participatory process to identify root causes of poor data quality and to develop immediate- and long-term plans for addressing threats to data quality. After the results of the DV are computed, the entire team including implementing partner staff should review the results and develop a joint plan for addressing the issues identified. A participatory improvement planning process ensures ownership of the corrective measures to be defined as part of the work plan and contributes to lasting improvements in the identified challenge areas. It also ensures that issues found are analyzed, root causes are identified, and required actions are outlined. The process reflects the reality on the ground, and responsibilities for implementation of the identified actions are assigned in a realistic and clear way to be the most effective in producing the intended changes.

Step 1: Identify indicators with issues — Programs/projects should start by outlining or listing all the indicators that have a concurrency below 95% or above 105%. These should be prioritized by site, i.e., sites that have the above-mentioned levels of concurrency in the listed indicators and results from the beneficiary verification.

Step 2: Conduct root cause analysis — For indicators with more than 105% or less than 95% Verification Factor, conduct a root cause analysis using the tool below to identify the specific stage or process where the data error was introduced (see Figure 2).

Figure 2: Root Cause Analysis Tool



Step 3: Identify 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 the donor).

Step 4: Determine action needed — Follow-up steps refer to short- and long-term measures taken to strengthen the information system and prevent the detected error(s) from recurring. Examples may include mentoring, technical assistance, training, improved supervision, and distribution of guidelines. These measures should be specific and time-bound; furthermore, responsibilities should be indicated, together with any need for additional support or resources.

Step 5: Prepare the action plan —Summarize the steps and/or actions that will be followed to correct the identified data quality issues. ***Appendix V*** shows a screenshot of a template for a Data Quality Improvement Action Plan. Provide a copy of the plan for all parties involved for their signature.

4.0 After the Data Verification

The team will provide feedback on findings of the data verification to the program/project team including the management with recommendations, the developed action plan on how to address challenges identified, and timelines. The SI headquarters (HQ) backstop will provide regular follow-up to ensure plans are followed through to logical conclusion.

5.0 Report Writing

Following completion of the data verification, the DV team will provide a detailed written report of the exercise to program/project management within two weeks. The report will detail assessment methodologies, findings, recommendations, action plans, and limitations/challenges of the exercise. Relevant data summarization/visualization including infographics, charts, maps, and tables will be used to summarize the findings as appropriate.

Appendices

Appendix I. Part A: Data Recounting Checklist

Part A: Data Recounting Checklist

Program/Project Name: ________ **Report Period:** ________

Site Name: ________ **Date:** ________

Team Leader/Members: ________

Indicator description (include any required disaggregation, e.g. age/sex)	Indicator code	Source Document (Include the row, column, variable name of the source, if appropriate)	Month, Year	Reported value	Verified value	Verification factors	Status	All selected months summary reports are available?	10% Sampled recording tool of this month are available?	Summary scores of the reviewed forms: % recording tools have correct UIC (if applicable)/ registration number? (Integrity and consistency)	Summary scores of the reviewed forms: % of recording tools that have all data fields are filled out per logic/rule? (Completeness)	Summary scores of the reviewed forms: % of recording tools have correct date? (Timeliness)	Summary scores of the reviewed forms: % of recording tools have been signed by supervisor? (Validity)	Remarks
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											
			< Enter Month 1: name, Year>											
			<Enter Month 2: name, Year>											
			<Enter Month 3: name, Year>											

Part A. Data Recounting | Part B.Recording Tools | Part C-1. Physical Verification | Part C-2. Electronic Verifica | Part D. DQ Improvement Plan | Partner_Sites Profiling

Appendix II. Part B: Recording Tools Verification Checklist

	A	B	C	D	E	F
1	**Part B: Recording Tools Verification Checklist**					
2	**Indicator code:**				**Month:**	
3	**Data source tool name:**				**Year:**	
4	**Total Number of sampled tools:**				**Total "Yes":**	0
5	**Overall Score:**					
6						
7	**S. N.**	**UIC or registration number of tool**	**The recording tool has correct UIC (if applicable)/ registration number? (Integrity and consistency)**	**All data fields of recording tool are filled out per logic/rule? (Completeness)**	**The recording tool has correct date? (Timeliness)**	**The recording tool has been signed by supervisor? (Validity)**
26						
27						
28						
29						
30						
31						
32						
33						
34						
35						
36						
37						
38						
39						
40						

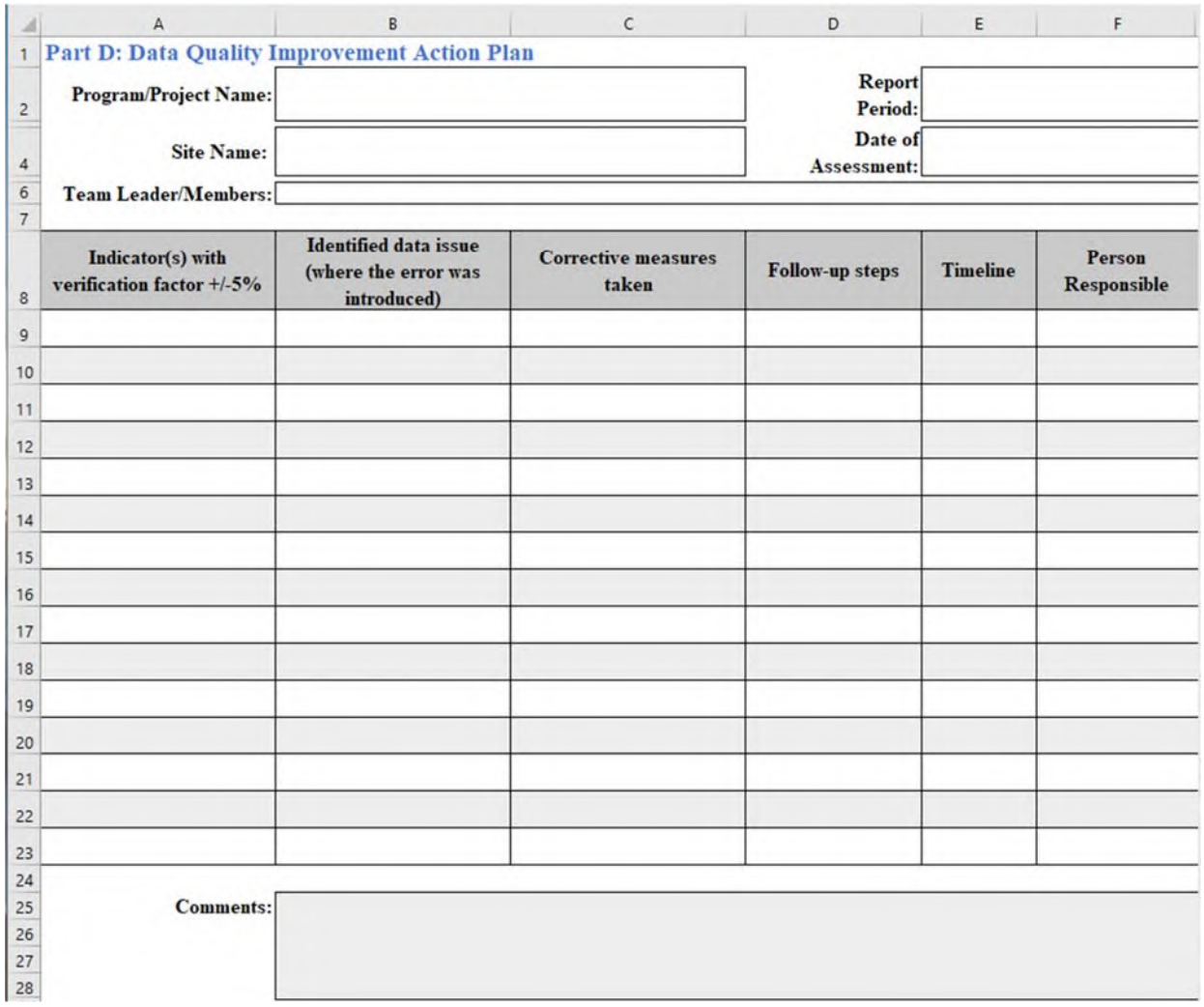
Appendix III. Part C-1: Physical Verification of Beneficiaries Checklist

	A	B	C	D	E	F	G	H	I	J
1	Part C-1: Physical Verification of Beneficiaries Checklist									
2	Total no. of beneficiaries to verify:					Total confirmed beneficiaries:	0		Score	0%
3	S.N.	UIC/Registration No/Service No.	Phone Number	Alternative Phone Number	Sex	Descriptive/physical Address	Date Service Received	Recieved Services	Confirmation of services after physical meeting	Remarks (Insert remarks for every failed verification)
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										
17										
18										
19										
20										

Appendix IV. Part C-2: Electronic Verification of Beneficiaries Checklist

	A	B	C	D	E	F
1	**Part C-2: Electronic Verification of Beneficiaries Checklist**					
2	**Total no. of beneficiaries to verify:**			**Total confirmed beneficiaries:**		**0**
3	**Score:**					
5	*Tick applicable level of verification (one or both)*					
6						
7	☐ **Level 1**: Implementation/Partner level					
8	☐ **Level 2**: National/Program Level					
9	**S.N.**	**UIC/Registration No/Service No.**	**Sex**	**Date of service**	**Place service was offered**	**Services provided**
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25						

Appendix V. Part D: Data Quality Improvement Action Plan



	A	B	C	D	E	F
1	Part D: Data Quality Improvement Action Plan					
2	Program/Project Name:				Report Period:	
4	Site Name:				Date of Assessment:	
6	Team Leader/Members:					
7						
8	Indicator(s) with verification factor +/-5%	Identified data issue (where the error was introduced)	Corrective measures taken	Follow-up steps	Timeline	Person Responsible
9						
10						
11						
12						
13						
14						
15						
16						
17						
18						
19						
20						
21						
22						
23						
24						
25	Comments:					
26						
27						
28						