JOYFUL HEARTS

Data Quality Management Plan

Version 01

Name of Program:- MALINDZA TB/HIV PROJECT

Date of Original Plan: June 27, 2014

Date updated: June 27, 2014

Table of Contents

Section 1- Background

1.1 Introduction

- All projects and programs that involve the collection, collation, manipulation and reporting of data must ensure that such data meets minimum criteria for data quality. These criteria are far more than a statistical exercise and involve the design, implementation and management of a Data Quality Management System (DQMS) documented in this Data Quality Plan (DQP). This means that persons managing the data must know and understand data quality criteria, be able to design a plan to manage data quality and be able to internally audit the quality of the data. In essence, the end user or recipient of data has a right to information that is valid, reliable and accurate. The M&E officer has the obligation to ensure that this is the case.

1.2 Table of Key Indicators

Key Indicator	Source	Data Collection Tool	Frequency of Collection
Number of individuals who received Testing and Counseling services for HIV and received their test results	HTC Client Record Forms	HTC Client Record Forms	Daily
Number of HIV-positive adults and children receiving a minimum of one CLINICAL service	HTC Client Record Forms	Adherence Support Register	Weekly
TB/HIV: Percent of HIV-positive patients who were screened for TB in HIV care or treatment settings	TB Screening Tool	TB Screening Tool	Daily
Number of adults and children who were provided with a minimum of one care service (de-duplicated total of the CLINICAL CARE and SUPPORT CARE targets).	HTC Client Record Forms	HTC Client Record Forms	Daily
Number of HIV positive individuals receiving ART adherence support services	Adherence Support Register	Adherence Support Register	Weekly
Number of TB suspects identified and referred for treatment	TB Referral Forms	TB Screening Tool	Weekly
Number of TB patients receiving TB treatment adherence support	Adherence Support Register	Adherence Support Register	Weekly
Number of condoms distributed	Condom Distribution Forms	Condom Distribution Forms	Daily

1.3   Key Stakeholders

Stakeholder	Data Management Stage	Responsibility
SNAP Lubombo	National	Inform national standards and strategies in data management
Program Officers (Nurse & HTC Counselors)	Data Collection	Quality data collection by using adequate and right tools capturing correct and reliable information
M&E Officer	Data Collection	Quality data collection, collation, analysis, and management
Programme Director	Data Quality Management	Aligning data management with international and national ISOPs
CANGO/PACT	Data Quality Management	Capacitating partner in best practices of quality data management

Section 2 – Data Management System

2.1 Data Flow Map and Data Management Process

Data Flow Map
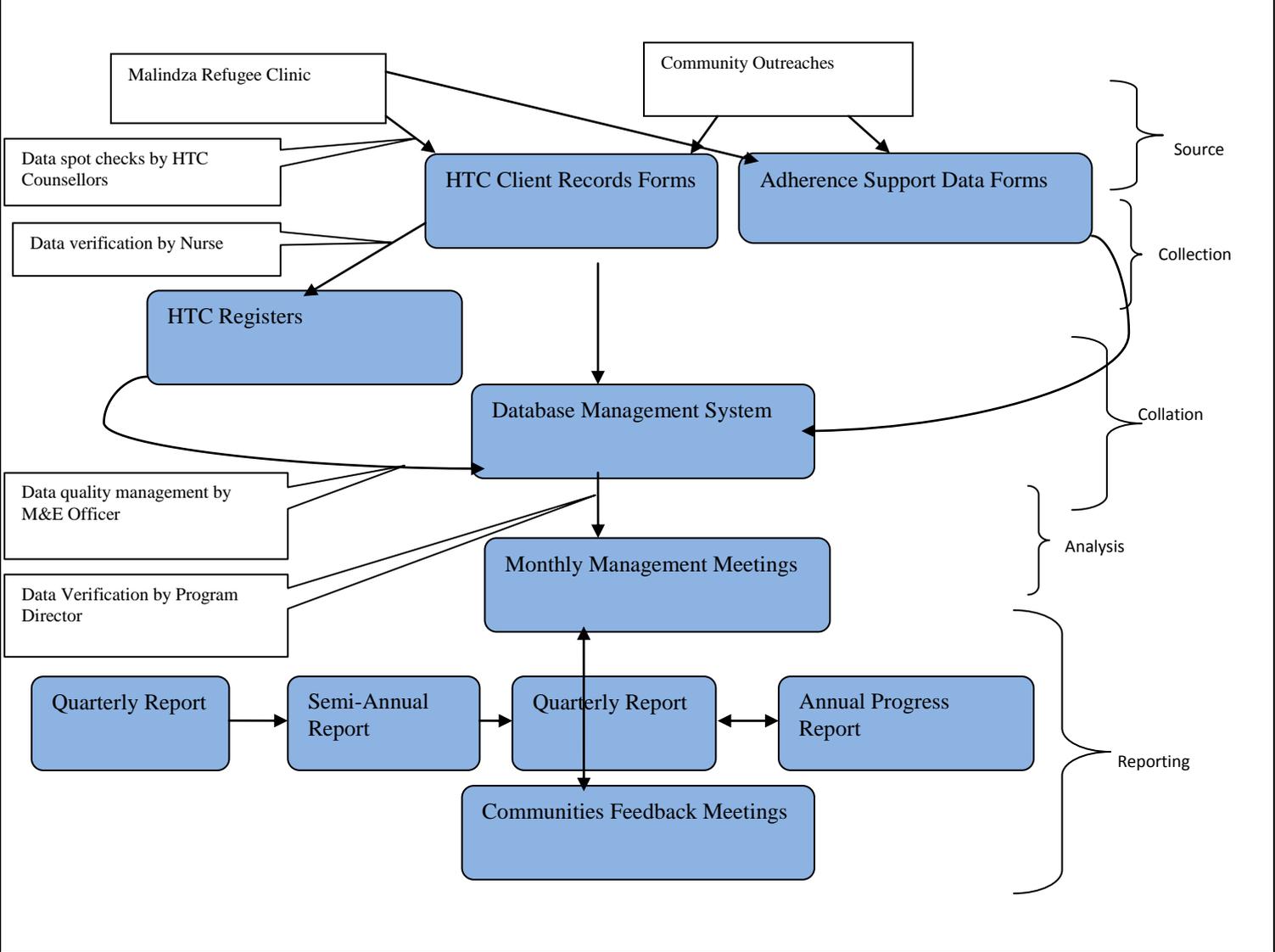
2.2 Standard Operating Procedures: Routine Data Quality / Mgmt Practices to be implemented by Pact

2.2.1 Routine Data Quality Procedures – Against Data Quality Criteria

Data quality refers to the worth/accuracy of the information collected & focuses on ensuring that the process of data capturing, verifying and analysis is of a high standard. There are 5 key components of quality data: 1) Validity 2) Reliability 3) Timeliness 4) Precision and 5) Integrity. To ensure quality standards are developed and maintained issues and risks relating to data quality need to be thought through and documented. Thus we develop and document our means to implement:

1   **Standard procedures to ensure the data adequately represent performance and are valid.**

As standard practice Pact works to ensure:

Face Validity:
- There is a solid, logical relation between the activity or program and what is being measured…

Transcription Validity:
- Transcription / data entry and collation procedures sound and is the data entered/ tallied correctly.

Measurement Validity:
Measurement tools / procedures well designed, and defensible and to limit potential for:
- Non sampling errors
- Sampling / representation errors
- Memory errors
- Self-presentation bias

Our critical validity issues are:

To ensure <u>the data adequately represent performance</u> are thus valid we either undertake the following SOPs ourselves or mentor our partners and grantees to routinely:

- ☑ Clarify Results Statements and select indicators that meet the criteria of being "Direct".
- ☑ Carefully select data collectors. Consider age, marital status, gender to better match cultural norms on target subjects (sexual communication, age practices…).
- ☑ <u>Carefully train data collectors / grantees on: methods, procedures, ethics, and interview techniques.</u>
- ☑ Inform the community, individuals, and households in advance with letters, radio, or community leader visits explaining the purpose of the data collection effort.
- ☑ Conduct pre-tests /pilots of any variables to be incorporated into questionnaires or interviews – check the meaning of the language used and the accuracy of respondents recall.
- ☑ Double check all sampling methodologies prior to data collection.
- ☑ Structure questionnaires so that there are "warm up" questions and that sensitive questions are asked later in the process and only to appropriate respondents.
- ☑ Select questions that have been internationally "standardized".
- ☑ Understand and clarify the effective range of scale – questions for children will differ than questions for adults.
- ☑ When answers to questions require recall provide specific reference points for respondents such as a month ago or a year ago (rather than an artificial period such as 3-6 months ago).
- ☑ Carefully word questions to address self presentation bias
- ☑ State (and follow) your confidentiality / privacy practices.
- ☑ <u>Routinely randomly check data for transcription errors.</u> If absolutely necessary double key information.
- ☑ <u>Track errors to their original source and correct mistakes.</u>
- ☑ Always double check that final numbers are accurate!

2   **Standard Procedures to ensure the data the collection processes are stable and consistent over time and are thus reliable.**

As standard practice JOYFUL HEARTS ORGANIZATION works to ensure:

Consistency:
- Consistent data collection processes are used from year to year, location to location, data source to data source.
- The same instrument is used to collect data from year to year, location to location.
- If data comes from different sources the instruments are similar enough that the reliability of the data are not compromised.
- The same sampling method is used from year to year, location to location, data source to data source.

Reliable Internal Quality Controls:
- Procedures exist to ensure that data are free of significant error & that bias has not been introduced.
- Procedures are in place for periodic review of data collection, maintenance, and processing.
- Procedures provide for periodic sampling and quality assessment of data.

Transparency:
- Data collection, cleaning, analysis, reporting, & quality assessment procedures are documented in writing.
- Data problems at each level are reported to the next level.
- Data quality problems are clearly described in reports.

To ensure the data the collection processes are stable and consistent over time and are thus <u>reliable</u> we either undertake the following SOP ourselves or mentor our partners and grantees to routinely:
- ➢ SWAZILAND NATIONAL HIV TESTING AND COUNSELLING GUIDELINES August 2010
- ➢ NATIONAL STRATEGY FOR COMBATING HIV AND AIDS RELATED STIGMA AND DISCRIMINATION

☑ <u>Develop indicator protocols--clarify instructions/definitions for all data collection efforts at all levels of the process.</u> Document data collection, cleaning, analysis, reporting, and quality assessment procedures.

☑ <u>Make sure everyone is aware of the procedures and definitions</u> and follow the procedures from reporting period to reporting period!

☑ Maintain official MER records with copies of all protocols and instruments.

☑ When possible using multiple (3) independent raters when data is based on human observation

☑ Provide guidance to data collectors and transcribers on what to do if the see a data error and who to report it to. Develop an error log to track data errors and document data quality programs.

3   Standard Procedures to ensure the data are collected frequently enough, are current, and are thus <u>timely</u>.

As standard practice J OYFUL H EARTS O RGANIZATION works to ensure:

Frequency
- Data are available on a frequent enough basis to inform program management decisions.
- A schedule of data collection, collation, analysis and reporting is in place that meets program management needs.

Currency
- The data are reported in the most current timeframe practically available
- The data are from within the policy period of interest
- The data are reported as soon as possible after collection
- The date of collection is clearly identified in reports

To ensure the data are collected frequently enough, are current, and are thus <u>timely</u>. We either undertake the following SOP:

☑ Make a schedule and keep to it!

☑ <u>Expect/include time in data collection efforts for a large number of call-backs</u>.

☑ Have a calendar for when data is due at each level in the data mgmt process

☑ Clarify dates of data collection periods when presenting results.

☑ Clarify needs of MER audiences and meet those timelines

4   Standard Procedures to ensure the data have an acceptable margin of error and are thus <u>precise</u>.

As standard practice J OYFUL H EARTS O RGANIZATION works to ensure an acceptable:

Margin of Error
- The margin of error is less than the expected change being measured.
- The margin of error is acceptable given the likely management decisions to be affected. (consider the consequences of the program or policy decisions based on the data)
- Targets have been set for the acceptable margin of error.
- The margin of error has been reported along with the data.
- Clarity on if an increase in the degree of accuracy be more costly than the increased value of the information

To ensure the data have an acceptable margin of error and are thus <u>precise</u> we either undertake the following SOP:

- ☑ Calculate and know our margins of error
- ☑ Set targets for our margin of error
- ☑ Report any issues around precision/ develop an error log
- ☑ <u>Consider the cost benefits of the various levels of precision / confidence intervals.</u>

5 Standard Procedures to ensure data are free of 'untruths' (introduced by either human or technical means, willfully or unconsciously) and therefore have <u>integrity</u>

As standard practice JOYFUL HEARTS ORGANIZATION works to ensure there is no:

Manipulation / Bias
- Mechanisms are in place to reduce the possibility that data are manipulated for political or personal reasons.
- There is objectivity and independence in key data collection, management, and assessment procedures.
- There is independent review.
- If data is from a secondary source, we are confident in the credibility of the data

To ensure the data are collected frequently enough, are current, and are thus <u>timely</u>. We either undertake the following SOP:

2.2.2 Routine Data Management Procedures

Data Management refers to the standard operating procedures/ routines and actions put in place to track data flow and reduce risk to ensure the data does not become corrupted at any stage of the data flow process there are 6 key processes in the data mgmt chain:1) Source 2) Collection 3) Collation 4) Analysis 5) Reporting 6) Usage. To ensure quality standards are developed and maintained issues and risks relating to data quality need to be thought through and documented. Thus we have developed and documented our means to implement:

1 Standard procedures to ensure data quality at the source

The Source is where the data originates, e.g., HTC Client Records Forms(primary) or Ad(primary); or patient files in a hospital (primary), or hospital excel spreadsheets with patients data (secondary), or National HIV/AIDS Surveillance Report (tertiary), etc. The potential for risk for poor data quality increases with secondary and tertiary data sources. Examples of data quality issues at source:
Could be Validity type issues: data could be incomplete (incomplete Drs Notes, ineligible notes in patient files); or for instance Reliability type issues: inconsistent recording of information by different staff depending on skills levels, etc. To help ensure data quality at the source we either undertake the following SOP:

- ☑ Design instruments correctly
- ☑ Include data providers (community stakeholders) and data processors in decisions to establish what feasible, review process is, and draft instruments.
- ☑ Ensure all personnel are trained in their assigned task use -1 trainer if possible
- ☑ Develop & document instructions for the data collectors, on the collection forms, and for computer procedures
- ☑ Develop an appropriate sample
- ☑ When possible comply with professional standards for data set up.
- ☑ Ensure all data collection, entry and analysis needs are available (pens, paper, forms, computers…).

2 Standard procedures to ensure data quality during collection

Collection is the process of gathering data generated from the various activities implemented by an organization and relevant to an organization's M&E framework. Collection Involves obtaining data from original source and transferring it into tools (paper or electronic) from which it can be analyzed and reported

To help ensure data quality during the collection stage we either undertake the following SOP:

- ☑ Train data collectors in how to collect information
- ☑ Routinely check to see if instructions are being followed
- ☑ Identify what to do if you (or someone) wants to make a change to the data collection process or if you have problems during data collection (change mgmt process)
- ☑ Check to see if people follow the change mgmt process
- ☑ Develop SOPs for managing the collected data (moving data from 1 point to the next).
- ☑ Develop SOPs for revising the collection tool
- ☑ Communicate the process
- ☑ Conduct on site reviews during the process

3   Standard procedures to ensure data quality during collation

Collation refers to aggregation of data into summarised formats. Collation can be done electronically or manually. Typically data quality concerns during collation include data being inconsistently or wrongly entered and/or margins of error being compounded particularly in cases where estimates and sampling have been used. To help ensure data quality during the collation stage we either undertake the following SOP:

- ☑ Develop check lists and sign off for key steps
- ☑ Conduct reviews during entry process
- ☑ Create an electronic or manual format that includes a data verification process by a second individual who is not entering the data/ Randomly sample data and verify
- ☑ Ensure problems are reported and documented, corrected and communicated and tracked back to the source of the problem

4   Standard procedures to ensure data quality during analysis

Analysis refers to processes that improve understanding of how the data relates to program objectives and targets, it iinvolves the review & manipulation of data and enables data users to compare data results against targets, associate variables, predict relationships, indicate confidence in results, aggregation of data into summarised formats.etc. Collation can be done electronically or manually. Typically data quality concerns during analysis include either incorrect analysis, or misrepresentation of results To help ensure data quality during the collation stage we either undertake the following SOP:

- ☑ Run audit report for review by experts with knowledge for reasonableness
- ☑ Ensure analysis techniques meet the requirements for proper use
- ☑ Disclose all conditions /assumptions affecting interpretations for data
- ☑

5   Standard procedures to ensure data quality during reporting

Reporting refers to compilation of descriptive information, presenting raw data and information as useful knowledge and provides an opportunity for project implementers and others to inform themselves of progress, problems, difficulties encountered, successes and lessons learnt during implementation
Typical data quality concerns include: too little reporting or feedback, not presenting information as useful knowledge, misrepresentation of results.To help ensure data quality during the reporting stage we either undertake the following SOP:

- ☑ Synthesize results for the appropriate audience
- ☑ Maintain integrity in reporting – don't leave out key info
- ☑ Have multiple reviewers within the organization - prior to dissemination!
- ☑ Protect confidentiality in reports/ communication tools
- ☑ Review data / provide feedback with those who have a stake in the results

6   Standard procedures to ensure data quality during usage

Usage is about making timely data informed or data driven decisions within our programs..To help ensure data quality during the reporting stage we either undertake the following SOP:

- ☑ Consistently promote use of the data
- ☑ When possible provide data in terms of spatial or graphic representation
- ☑ Routinely make available data at key decision points (staff mtgs, review etc)
- ☑ Understand your data's strengths and weaknesses and incorporate results appropriately

2.2.3   Routine data management procedures to ensure data processes are ethical

Data ethics are the rules or standards governing the conduct of a person collecting, collating, reporting, or utilizing data and represent our standard of "right" conduct. The data management process involves a series of decisions that are made by the organization to ensure ethical data collection. To help ensure ethical data collection we either undertake the following SOP:
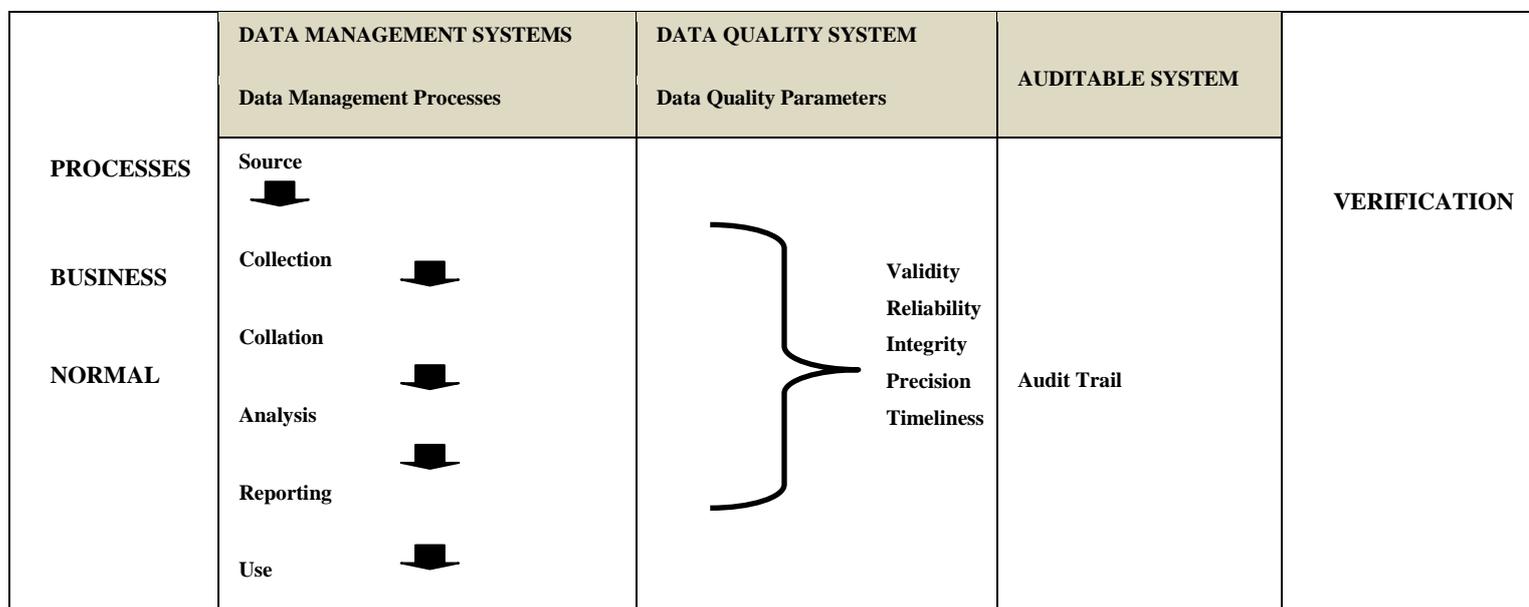
- ☑ Ensure Informed Consent--Participants (individuals & communities) are always provided enough information about the data collection effort to enable them to make informed decisions about their participation. Get a written or spoken consent statement as applicable. Consider the rights of children and their capacity to provide informed consent – get permission from parents / guardians
- ☑ Protect Privacy--Participants are always made aware of how much privacy they can expect about their responses (be it confidentiality or anonymity). Do not discuss an individuals answers and be sure the staff / community volunteers do not as well, Use "blind data" do not put names to the data (files, reports..) Use codes if possible
- ☑ Misrepresenting or Falsifying Data is not allowed and Pact staff can be dismissed for these reasons (or grantees can lose funding)
  - In-appropriate scale up- Have incomplete, anecdotal or case study data but presenting it as a completed data set or representative of a larger population
  - Failure to note the results are due in whole or part to a partner
  - Adding missing data (during collation)- making up data or changing responses
  - Limitations not made clear
- ☑ Ensure staff and grantees are aware that all participation in data collection efforts is voluntary, if someone one does not want to provide information they may not be refused servces for this reason. The subject has the right to refuse to participate in data collection efforts/ Do not coerce people into participating
- ☑ Have a signed code of conduct for all individuals assisting in the data collection, collation, analysis, and reporting effort.
- ☑ Empower and include communities: Discuss with communities/ participants how you plan to carry out the collection effort & why you are doing the research, Discuss with communities/ participants how you will use the data. Know and follow cultural norms / practices as appropriate
- ☑ When presenting results explain the limitations of the findings. Provide results / feedback to the findings to the participants and the communities

Section 3 -   Evaluation of the Current Data Quality Status / Risk Analysis of the Indicator Set:

In addition to following our standard operating procedures CANGO conducts a Data Quality Assessment (DQA) of each indicator every three months (Quarter) data collection efforts and completes a risk analysis. The DQA process is focused on asking a series of pointed questions on data quality elements and subcomponents and researching the answers. Based on our research, risk is assessed at the indicator level utilizing the data quality criteria to determine the potential level of error that may exist in the data and the overall effect a given error would have on our ability to use the data. Determining the level of risk is ultimately a judgment call which helps us prioritize issues and determine where to spend our time and resources. Where we identify moderate or high risks to data quality we develop mitigation actions to bring the risk within an acceptable level. Our Procedures for DQA are as follows:

3.1 Procedures for Conducting a Data Quality Assessment of Partner Data

Each of the data quality parameters must be monitored at each point in the data management cycle. In addition there must be an auditable trail to indicate that the data maintains its quality at each point in the data management cycle. This relationship is illustrated in the diagram below.

	DATA MANAGEMENT SYSTEMS Data Management Processes	DATA QUALITY SYSTEM Data Quality Parameters	AUDITABLE SYSTEM	
PROCESSES **BUSINESS** **NORMAL**	Source Collection Collation Analysis Reporting Use	Validity Reliability Integrity Precision Timeliness	Audit Trail	**VERIFICATION**

- Bi-annual data quality audits at 3 levels
    - Program-Database and data in tools sent
    - Facility-Reported and source
    - Community-Reported and source
- Routine data quality checks on received data.

Procedural Guidance

1) Undertake a visit (or a series of visits) to begin the verification and validation process of the pre-selected indicator. The assessor usually takes a piece of data (from the source) and tracks it through the organizations data management system -- i.e. identifying issues at 1) the Source; 2) the Collection process; 3) the Collation process; 4) the Analysis process; 5) the Reporting process; to assess if there are any issues in the quality of the data in terms of its' validity, reliability, integrity, timeliness, and precision. The assessor often visits communities to help determine if they can vouch that services were actually delivered

2) At each stage of the process data quality assessment questions are asked and findings are documented capturing any issues that might be helpful to develop a risk assessment score

3) Risk scores are determined and recommendations are made which are tied to the risk score (high risk data will either be disallowed or the organization will be provided a grace period to fix it).

4) A feedback meeting with the organization is held on the findings and recommendations.

5) A formal report is generated.

6) Follow up visits are usually conducted within 6-12 months to ensure recommendations are being implemented.

3.2 Data Quality Assessment Tool

Please insert or attach the Data quality assessment tool which you are going to use. You can adapt the following tools:

- JOYFUL HEARTS ORGANIZATION Data Quality Assessment Tool
- Routine Data Quality Assessment Tool
- Data Quality Audit Tool (Phase 2)

Section 4- Summary of Corrective Actions to be Undertaken

Where data quality is found to be of concern Pact identifies corrective actions it will take (or asks our partners to undertake) to improve quality and accuracy of the data set (our mitigation plan).

Normally level of response action is guided by the risk score where:

1 – 3 = Low Risk: No immediate action required; risk can be managed through normal internal data quality controls.
4 – 8 = Medium Risk: Contingency plan established to reduce risk. Data must then be verified and validated prior to our annual report, and a strict audit trail maintained.
9 – 16 = High Risk. Contingency plan established to reduce risk. Data must then be verified and validated prior to each reporting episode, and a strict audit trail maintained.

4.1 Summary of Key Data Quality Issues and Mitigation measures

Update this section as appropriate.

Name of Indicator	Risk Score	Data Quality Challenge	Mitigation Action Planned	Comments/ Updates
M&E Systems		Data Management Backup Documentation	Work on developing electronic management system. Documentation of backups	Database now available and being developed. A tool for documentation of backups being developed
Program Implementation		Data collection gaps		