Quality Assurance Process

Data Collection/Database

Data entry databases were created using Access, and in a few cases, Excel

Where possible, drop down boxes were provided in the databases to standardize answers and facilitate data entry (e.g. fields for region, type of facility, and satisfaction scale questions)

Validation formulae were used where appropriate to increase the quality of the data entry (e.g. designating fields as numeric where appropriate and restricting the range of permissible answers)

Data entry supervisors were provided with detailed written instructions and screenshots on how to navigate and enter data into the databases

If necessary, a conference call was made with the data collection/entry supervisors to go over the data collection and entry process and answer any outstanding questions. In Ghana and Tanzania, this was facilitated by in-country MSH staff.

Data Collector Debriefing (Post-data collection)

In-country data collection coordinators were requested to organize debriefings with data collectors upon returning from the field. Debriefings served to identify any problems with the data collection process, the need to return to certain facilities for follow-up, and to identify any information about facilities and the process that would be relevant to data collection. The debriefings were documented and included information on the following:

- Reconciliation of facilities visited and forms collected - be sure to check this according to the analysis that is to be done to make sure that control and intervention groups have a sufficient number of each form collected.
- Facilities where data collection was interrupted or cancelled – identify if a replacement facility needs to be selected and visited in replacement
- Issues with forms – for documentation and lessons learned

Data Quality Checking

Data collection coordinators were requested to send copies and/or originals of the data collection forms to the MSH/Arlington office. They were asked to keep a set for themselves. This facilitates checking of values on both ends, as well as data entry QA. While this was an expense for the field office, it helped to provide a way to check any specific entries that looked incorrect.

Data in the databases were double checked against the copies of the data collection forms

Facility information (e.g. facility name, facility type, region) for each form was verified, and corrected where necessary. This was done because in some cases the field staff did not keep track of how many facilities were visited, the locations, or the facility type and as a result the information was often entered inconsistently. For example, a facility was entered as an ADDO when it was really a Duka la Dawa Baridi. By checking the facility information across forms as a first step, the facility identifiers could be checked for consistency, and counts could be made on the total number of forms collected and facilities surveyed.

Data collection/entry supervisors were contacted to address any queries that arose from the process of verifying the data in the databases

Detailed notes were created to keep track of the data validation process – there is one set of notes per form used in an assessment, and each keeps track of any queries, clarifications, corrections and decisions made during the QA process. The QA documentation process included the following elements:

- Each set of notes included general comments about the set of forms
- Corrections made to the database when data were not consistent with the data collection forms, queries about specific forms (inconsistent data, data collection issues, missing information and data) were noted.
- Queries and questions were sent to the field to be addressed.
- Clarifications, decisions and actions made to the database were added to the notes and highlighted to keep track of what was resolved and what was still pending.
- The QA process ended when all pending questions were reviewed and answered, and updates were made to each database. If clarification of any question was not possible, in most cases the information was left in the database as it appears on the form. In some cases, information that was inconsistent or redundant was deleted and a note of this was made.
- Notes were also made on the actual data collection forms to keep track of any data collection errors, as well as changes made to the original data as a result of the QA process (e.g. standardizing facility identifiers).

The most common data errors were:

Data collection errors

- *Incomplete or missing data* - data collectors often did not complete all fields on a data collection form, or did not provide pertinent information in the Comments section in instances where comments were warranted
- *Comprehension problems* – data collectors often had a problem filling out the price form, and often wrote data in the wrong columns, or

recorded the units per pack data using the wrong units for a given tracer item

Data entry errors

- *Incorrect data entry* – when data on the forms did not match what was typed into the database, these corrections were made and noted in the QA tracking document.
- *Incomplete data entry* – e.g. for an entire set of satisfaction forms, half of the data for the qualitative questions had not been entered
- *Typos*

Report Tables and Graphs (Combine this with the section below – they seem to overlap)

Most data were entered into the Access databases. In some cases, as noted above, Excel spreadsheets were used for primary data entry.

Upon completion of the QA process, the data were generally exported to Excel to facilitate data analysis and to add any additional coding that was required.

Data analysis files for each assessment form contain raw data, all analyses, summary tables, and (frequently) any graphs that were made for reports or presentations

In Excel, most of the data have been summarized using Pivot Tables, which can easily be updated (they do not automatically update – each worksheet containing pivot tables must be updated individually, though updating the worksheet will update all the pivot tables contained in it)

Tabs in the Excel Files labelled "Summary Tables" contain all data needed to calculate the indicators of interest. The data in these tabs were copied over from the pivot tables and formatted into a more useful presentation.

Graphs were created in individual sheets within each Excel spreadsheet. Especially in later assessments, graphs were linked back to the summary tables so that they would automatically update if any changes were made to the results

Methods and Programs Used for Data Analysis

General (Combine with the section above – see note above)

Most data were analyzed in Excel (although the earlier versions of some files, e.g. for ADDO, are in SPSS) – data were exported from Access to Excel where necessary

A "Notes" or "Assumptions" worksheet was included in most Excel data analysis files. These pages included comments on who had worked with the data, the dates, and what changes/additions/corrections/analyses had been performed. This page also included any assumptions that were used in coding the data, such as decision matrices for coding STG decisions about simulated clients.

In some cases, a deleted data tab was added to the Excel sheet. For example, in one assessment a facility was visited that was not the correct facility to include. The data were not to be included in the analysis. In the interest of conserving the information, the data were deleted from the main data spreadsheet and saved in a "deleted items" tab, along with an explaination.

The original raw data are typically included in Excel data analysis files for easy reference (in the case of the CPS medical record review raw data, there are so many individual raw data sheets that the worksheets have been hidden, but instructions in the file explain how to display them)

Common Problems*

***That caused delays, affected the quality of the analyses, or resulted in a loss of data**

Incorrect or conflicting facility type and region information

Spelling (e.g. drug and facility names) – pivot tables cannot effectively summarize data that are not standardized – provide an example

Incorrectly recorded price information - misunderstandings were often not apparent until data collection was over, which increased the time required for the quality assurance process as well as the potential for the loss of data – provide an example

Data collected for forms of a drug that weren't specified on the tracer list - e.g. collecting data for the tablet form of a tracer item when the tracer list specifies that the tracer item is in suspension form

Incorrect or inconsistent recording of pack size units for tracer items - e.g. recording the pack size of a 100ml bottle of a tracer item as 1 unit when the tracer list specifies that 1 unit is equivalent 500ml bottle (this was a frequent problem for suspensions and oral contraceptive pills)

Determining the generic names for obscure local brand name drugs – for simulated client data, all the antibiotics and antimalarials had to be identified. In some cases there was space for this on the data collection forms, but the data collectors did not

know the generic names and were unable to fill the information in. Ascertaining the generic names of medicines required extensive support from country staff.

Fields that are left blank, or marked as unavailable or not applicable – the formulas used to calculate indicators had to be adjusted to account for missing data

Poor photocopies of the original forms - records in pencil were often illegible and it was often difficult to get revised copies mailed or faxed

Lengthy turnaround time on queries to local resources

Common Problems Specific to the Baseline Assessments

Original baseline data collection forms (filled out) were not available – Any attempts made to locate the forms were unsuccessful

Poor recording keeping/documentation on the QA and data analysis processes for most of the baseline assessments – It was difficult to locate the final raw data and data analysis files. When the files were finally received, it was unclear how much QA had been done on the data, and whether the data and analyses contained in those files were the most current (none of the files were accompanied by any documentation on the QA or analyses). In some cases where coding was required, there were no files available with the original data coding, so the baseline had to be re-coded to ensure consistency.

Length of time it took to obtain baseline files – in some instances, the baseline data were received after the endline data collection and analyses had been completed. This increased the length of time it took to complete the baseline-endline analyses, and reduced the amount of time available for queries to local resources regarding the baseline files.

Quality Assurance Problems Specific to the Baseline Data

Quality of baseline data – the baseline data could not be validated (see exceptions note below) without the original data collection forms. One way to estimate the quality of the baseline data is to assume that the prevalence of errors in the baseline data were similar to what was found in the endline data through the QA process, though this method is probably an underestimation. Many of the data entry quality control steps put in place for the endline assessments (e.g. designing databases to minimize data entry error, and providing detailed written instructions for data entry) were not present for the baseline assessments, so it is likely that the frequency of data entry errors would be greater than estimated. Regardless, the baseline-endline analyses are comparing unvalidated baseline data to clean endline data, which needs to

be taken into consideration when interpreting the results of the baseline-endline analyses.

Correcting errors in the baseline analyses – In rare cases, it was possible to check the quality of the baseline data and make the necessary corrections without the original forms. This was the case for the price data – provided that the data collectors recorded data in the appropriate columns, and provided details where necessary, the units for each tracer item could be standardized (In many cases they had not been. For example prices for a quantity of 10 tablets had been combined in some cases with prices for 1 tablet. This had to be re-coded and standardized at the time of the endline analysis.).

Standardizing the baseline data for the purpose of the baseline-endline analyses - All baseline data were reanalyzed to ensure that they went through the same QA (to the extent possible), formatting, analysis, and documentation process as the endline data. The following are examples of how the baseline data were standardized:

- Raw data and data analysis worksheets were arranged and formatted in the same manner as the endline files
- "Notes" or "Assumptions" worksheets were included in the revised baseline data files to document any queries, corrections, or alterations
- Outliers were corrected or removed based on the same criteria used in the endline analyses
- Dose instructions for antimalarials and antibiotics dispensed during simulated client visits were evaluated based on the same STG criteria used in the endline analyses
- Units for tracer items in the price per unit analysis were standardized to match the units used in the endline analysis

Price

1. Price data were standardized so that the comparison units were identical for a given tracer item (e.g. 1 tablet/capsule, 1 pack of 28 oral contraceptive pills, 1 bottle of 100ml)

2. Any data collected for incorrect dosage forms or strenghts of a tracer item (for example suspension versus tablet or 200mg versus 500mg) were excluded from the analyses

3. Price per unit data were calculated and outliers were identified

4. Outliers were checked against the data collection forms to see if they were the result of errors in the data entry (in which case the outliers could be easily corrected by referencing the original data collection forms)

5. Data collection/entry supervisors were contacted to help resolve or evaluate any remaining outliers

6. Rarely, outliers were excluded from analysis, however, the raw data for these outliers can be found in separate worksheets in those data analysis files

Availability

If a tracer item was found to be expired it was considered "Not Availble". In cases where expiry data were collected, data were checked to make sure that tracer items were correctly marked as unavailable if they were expired (though price analyses might include data for tracer items that were expired or otherwise unavailable)

Simulated Client

Malaria Simulated Client
Upper Respiratory Tract Infection (URTI) Simulated Client

1. Columns were added to the raw data to categorize the drugs that were dispensed or recommended but not purchased. The categories of interest were:

   Malaria
   - 1st line antimalarial sold
   - 2nd line antimalarial sold
   - 3rd line antimalarial sold (if applicable)
   - Antibiotic sold
   - other drug sold
   - # of drugs sold during the encounter
   - Antimalarial recommended but not sold
   - Antibiotic recommended but not sold
   - # of drugs recommended but not sold during the encounter

   URTI
   - Antibiotic sold
   - Antimalarial sold
   - other drug sold
   - # of drugs sold during the encounter
   - Antimalarial recommended but not sold
   - Antibiotic recommended but not sold
   - # of drugs recommended but not sold during the encounter

2. Columns were added to categorize the dispensing of medicines according to the following criteria:

   Malaria:

- exactly according to STG
- consistent with STG
- sold in sufficient quantities to be able to comply with STG dose (regardless of whether or not the dose instructions were correct according to STG)
- ***Note:*** *The STG criteria used to code these columns are documented in the data analysis files. They are specific to each country's STGs.*

URTI

- exactly according to STG for pneumonia
- consistent with STG for pneumonia
- in sufficient quantities according to the STG for pneumonia (regardless of whether or not the dose instructions were correct)
- ***Note:*** *The STG criteria used to code these columns are documented in the data analysis files and were only used in the ADDO analysis. The country's specific STGs were applied.*

3. The Access database used in these evaluations resulted in several lines of drugs for each facility. So, if a attendant at one facility sold 5 items, there was a line for each in the database. Because we wanted the denominator in the simulated client analysis to be the number of encounter, we had to cut down the information to one line per facility and code for the item of interest, in this case an antimalarial and/or antibiotic. To achieve this, in the excel file the original data was maintained and a new worksheet was added that summarized the raw data to show only one record per facility:

    - If the facility dispensed a drug of interest (an antimalarial or an antibiotic), the record for that drug was the one kept for that facility
    - If a facility dispensed both an antimalarial *and* an antibiotic, the record for the antimalarial was the one kept for that facility, and the appropriate column was marked indicating that an antibiotic was also dispensed
    - ***Note:*** *There was only one occasion where two antimalarials were dispensed by the same facility (CAREshop assessment), and notes on how the data were handled can be found in the data analysis file for this assessment*

4. Columns were added to the worksheet containing only one record per facility to categorize dispensing and referral practices according to the following criteria:

    - an antimalarial was dispensed and the client was referred
    - an antimalarial was dispensed, but the client was not referred
    - no antimalarial was dispensed, but the patient was referred
    - no antimalarial was dispensed, and the patient was not referred
    - at least one drug was dispensed and the client was referred

- at least one drug was dispensed, but the client was not referred
- no drugs were dispensed, but the patient was referred
- no drugs were dispensed, and the patient was not referred

5. A column was added to the worksheet containing only one record per facility to indicate facilities where dispensers asked the simulated client about *both* of following:

    - the symptoms of the child
    - other medications the child may have taken

Registration

1. Data were checked to make sure that all the items surveyed were appropriate for analysis (e.g. medical supplies, herbal products, and other non-drug items were removed – a list of these items, if applicable, can be found in the data analysis files as "deleted items").

2. Data were checked to make sure that the registration and approval status of the drugs surveyed were entered consistently and correctly. For example, data were sorted to show the drugs listed by generic name. In some cases a drug from a certain manufacturer was listed as registered, but in another entry it had been left blank. This was re-checked with the field staff and corrected as appropriate.

3. Data were summarized according to the following criterion where applicable (different categories were used in different assessments):

    - Registered
    - Unregistered
    - Approved
    - Unapproved
    - Locally manufactured
    - Country of origin

Stockout

The total number of days possible for each tracer item to be out of stock (i.e. the total number of days in the time period surveyed) was adjusted to account for months where data were missing or unavailable (data were mistakenly not recorded, or the data collectors did not have access to the stockout records for that month)

Malaria Medical Record Review (Ghana CPS/DTC Evaluation Only)

1. The following columns were added to the data to characterize the drugs dispensed:

    - Was an antimalarial prescribed?
    - What type of antimalarial was prescribed (i.e. what was the generic name)
    - Was the antimalarial an injectable?

2. The drugs prescribed were categorized as follows:

    - Chloroquine
    - Amodiaquine
    - S&P
    - Artemether derivative
    - Other antimalarial
    - Other drug

3. A new worksheet was created that summarized the raw data by patient encounter

4. A new worksheet was created that summarized the data by patient encounter so that for each column containing data, the following question was asked: Did at least one of the drugs prescribed meet this criterion?

Hypertension Medical Record Review (Ghana CPS/DTC Evaluation Only)

1. The following columns were added to the data to characterize the drugs dispensed:

    Was an antihypertensive prescribed?
    What type of antihypertensive was prescribed (i.e. what was the generic name)

2. The drugs prescribed were categorized as follows:

    - Atenolol
    - Bendrofluazide
    - Lisinopril
    - Methyldopate
    - Nifedipine
    - Propanolol
    - Other antihypertensives
    - Other drug

3. A new worksheet was created that summarized the raw data by patient encounter

4. A new worksheet was created that summarized the data by patient encounter so that for each column containing data, the following question was asked: Did at least one of the drugs prescribed meet this criterion?

Satisfaction

????