Data Quality in Malawi’s Health Information System: A Qualitative Perspective
Prepared in March-June 2017

Prepared by:
• Tiope Mleme, National Statistical Office, tmleme@yahoo.co.uk
• Ernest Kaludzu, Ministry of Health, ernestkaludzu@yahoo.com
• Emmanuel Chimbalanga, Save the Children/IMCI Malawi, emmanuel.chimbalanga@savethechildren.org
• Collins Mitambo, Ministry of Health, cmitambo@gmail.com
• Melody Sakala, African Institute for Development Policy, msakala@mlw.mw

Acknowledgments
The Ministry of Health (Dr. Charles Mwansambo, Mrs. Emma Mabvumbe, Mr. Isaac Dambula, Mr. Patrick Naphini, Mr. Simeon Yosefe), the National Statistical Office (Mrs. Mercy Kanyuka, Mr. Jameson Ndawala), and the Institute for International Programs at Johns Hopkins School of Public Health (Melissa Marx, Amos Misomali, Karen Finnegan, Tricia Aung, Rebecca Heidkamp) provided technical and logistical support to the study.

Financial support for this study was provided by Global Affairs Canada through the National Evaluation Platform (grant number 7059904).

Recommended Citation
National Statistical Office [Malawi], Ministry of Health [Malawi], Save the Children/IMCI Malawi, African Institute for Development Policy, and Johns Hopkins Institute for International Programs. 2017. Qualitative Data Quality Assessment Study Report. Lilongwe, Malawi, and Baltimore, Maryland, USA.
# Table of Contents

List of Acronyms, Tables, and Figures ................................................................. 4  
Executive Summary ............................................................................................ 5  
Introduction ......................................................................................................... 7  
Background ......................................................................................................... 7  
Study aims and research questions ...................................................................... 7  
Methodology ......................................................................................................... 8  
  Data Collection Methods .............................................................................. 8  
  Data Analysis ................................................................................................. 9  
  Ethical Considerations .................................................................................. 10  
Results ................................................................................................................ 10  
  Participant demographics ........................................................................... 11  
  Reporting process ......................................................................................... 11  
    Data collection ......................................................................................... 11  
    Aggregation ............................................................................................... 11  
    Submission ................................................................................................... 12  
    Data use ..................................................................................................... 13  
    Feedback ..................................................................................................... 14  
  Understanding data quality .......................................................................... 14  
    Definition .................................................................................................... 14  
  Data quality barriers ...................................................................................... 14  
    System issues ............................................................................................. 14  
    Training ....................................................................................................... 16  
  Data quality facilitators .................................................................................. 17  
    Training ....................................................................................................... 18  
    Innovation ................................................................................................. 18  
    Systems ....................................................................................................... 19  
    Partner Role ............................................................................................... 19  
Conclusions ......................................................................................................... 20  
Key Findings ...................................................................................................... 20  
Recommendations ............................................................................................... 20  
References .......................................................................................................... 21  
Appendix 1 - Consent forms .............................................................................. 22  
Appendix 2 - Data collection instruments .......................................................... 26  
Appendix 3 - Malawi IRB Waiver ..................................................................... 30
List of Acronyms, Tables, and Figures

Acronyms

<table>
<thead>
<tr>
<th>Acronym</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>ARI</td>
<td>Acute respiratory infection</td>
</tr>
<tr>
<td>CMED</td>
<td>Central Monitoring and Evaluation Division of the Ministry of Health</td>
</tr>
<tr>
<td>DHS</td>
<td>Demographic and Health Survey</td>
</tr>
<tr>
<td>DHIS-2</td>
<td>District Health Information System-2</td>
</tr>
<tr>
<td>DHMT</td>
<td>District health medical team</td>
</tr>
<tr>
<td>DHO</td>
<td>District Health Office</td>
</tr>
<tr>
<td>DQA</td>
<td>Data quality assessment</td>
</tr>
<tr>
<td>FGD</td>
<td>Focus group discussion</td>
</tr>
<tr>
<td>HIS</td>
<td>Health information System</td>
</tr>
<tr>
<td>HMIS</td>
<td>Health management information system</td>
</tr>
<tr>
<td>HMIS-15</td>
<td>Health management information system-15 (reporting form)</td>
</tr>
<tr>
<td>HSA</td>
<td>Health surveillance assistant</td>
</tr>
<tr>
<td>HSSPII</td>
<td>Health sector strategic plan, version 2</td>
</tr>
<tr>
<td>IDI</td>
<td>In-depth interview</td>
</tr>
<tr>
<td>IIP</td>
<td>Institute for International Programs</td>
</tr>
<tr>
<td>MICS</td>
<td>Multiple Indicator Cluster Survey</td>
</tr>
<tr>
<td>MOH</td>
<td>Ministry of Health</td>
</tr>
<tr>
<td>M&amp;E</td>
<td>Monitoring and evaluation</td>
</tr>
<tr>
<td>NEP</td>
<td>National Evaluation Platform</td>
</tr>
<tr>
<td>WHO</td>
<td>World Health Organization</td>
</tr>
</tbody>
</table>

Tables and Figures

Table 1  Study Participants by Cadre
Table 2  Themes and sub-themes applied to DQA transcripts
Executive Summary

Background

The National Evaluation Platform (NEP) aims at developing capacity among Government of Malawi staff to evaluate health and nutrition programs by identifying, systematically compiling, and rigorously analyzing data from diverse sources. Empowered with evidence, national and district leaders can make strategic decisions that will facilitate achieving maximum health and nutrition impact for the women and children of Malawi. From 2014 through 2017, Malawi has been building the NEP, with technical guidance from the Johns Hopkins University Institute for International Programs (IIP) and funding support from the Government of Canada.

In August 2016, the Malawi NEP team conducted a data quality assessment (DQA) which sought to characterize the quality of routine health data in Malawi’s health system and to identify systems-level factors associated with quality. This report focuses on the results of qualitative research, which explored the barriers and facilitators of data quality, as well as the understanding of data quality by different cadres. The results of a quantitative data quality assessment are presented in a separate report.

Methodology

This study collected qualitative data from participants in the data collection and reporting process: statistical clerks based at health centers, district health office-based HMIS Officers, monitoring and evaluation (M&E) staff, and CMED staff. We used focus group discussions (FGDs) and in-depth interviews (IDIs) to obtain information from users and stakeholders on the process of collecting data, how it shapes data quality, and how perceptions of quality impact data use. We visited four districts: Nsanje, Blantyre, Dowa, and Lilongwe. In each district, we conducted an interview with the HMIS Officer and a focus group discussion with the statistical clerks or other facility staff responsible for data collection and reporting. In-depth interviews were conducted with CMED staff at the central level and with clinical program staff. We analyzed transcripts using the framework analysis method. We analyzed data under four themes: the reporting process, data quality barriers, data quality facilitators, and understanding of the term “data quality.”

Results

Reporting process

- Steps in the reporting process include data collection, aggregation, submission of the report, data use, and feedback.
- Each step of the reporting process plays a key role in shaping data quality.
- Data use is limited at all levels of the health system, with different constraints influencing use across the cadres.

Data quality barriers and facilitators

- Systems issues both support and inhibit the collection of high quality data.
- Resource constraints, namely material shortages, transportation challenges, and limited training opportunities, play a pivotal role as a data quality barrier.
- Training is important in ensuring the collection of high quality data.
Defining “data quality”

- All cadres have a good understanding of the technical term “data quality” and can explain its definition.
- While data quality is generally understood in theory, its understanding does not translate to higher quality data

Conclusions

Data quality is shaped during the stages of data collection and reporting. At each step, there are barriers and facilitators, which contribute to the quality of data available in the HIS. Innovation at an individual and systems level can contribute to improving the quality of reported data. Additionally, training and support by partner organizations are key components in improving data quality. Although there are challenges with the collection and use of high quality data, Malawi is making progress in ensuring that high quality data are consistently available to end-users at all levels of the health system.
Introduction

Globally, increased attention has been given to improving the quality and accelerating the use of Health Information Systems (HIS) data. Effective management of health programs requires the consistent availability of high quality routine health data. These data enable program managers and policy makers to understand, monthly, how their programs are performing relative to targets and where intervention may be needed. In addition, high quality HIS data enable managers to identify high and under-performing districts, allowing a reaction to better allocate resources or technical support. The ability to monitor and manage performance on an ongoing basis is critical to achieving desired program outcomes and national and global commitments; however, it is highly contingent on the availability of high quality data.

While survey data from sources such as the Demographic Health Surveys (DHS) and Multiple Indicator Cluster Survey (MICS) permit the periodic monitoring of the performance of key indicators, surveys are conducted infrequently and are costly. HIS data are readily available and present nearly real-time estimates of intervention coverage and utilization. However, stakeholders have expressed mixed feelings regarding the quality of routine health data, and its suitability and credibility to support effective program monitoring and performance management actions.

This data quality assessment work was undertaken in August 2016, to support the Ministry of Health’s Central Monitoring and Evaluation Division in conducting a comprehensive and in-depth diagnosis of Malawi's health information system. The objective was to develop a better understanding of the quality of routine health data in Malawi, and the factors which are associated with data quality. This was a first step in identifying areas for intervention in order to improve data quality and increase data use. Thus, the Malawi NEP team, in coordination with the National Statistical Office and the Ministry of Health, conducted a qualitative assessment of the barriers and facilitators influencing the quality of routine health data.

Background

The quality of HIS data depends on underlying data management and reporting systems. In other words, for good quality data to be produced by and flow through a data-management system, key functional components need to be in place at all levels of the system - the points of service delivery, the intermediate levels where data are aggregated (e.g. health facility, districts, partner offices), and the M&E offices to which data are reported. There has been increased interest, globally and in Malawi, in understanding and improving the quality of routine health data. Data from a health information system can be a valuable source of information for health systems. However, the perceived and actual quality of routine health data in low- and middle-income countries hinders its use for policy and programming. Qualitative data collection was coupled with a quantitative data quality assessment and facility survey to understand the data collection process; qualitative data explore the process of creating HIS data from the perspectives of stakeholders and provide context for the DQA quantitative results. Understanding data collection, reporting, and use from the perspectives of stakeholders is important to understanding where CMED and partners can intervene to strengthen the process and better meet user needs.

Study aims and research questions

This qualitative study aims to explore how data are collected and where intervention can help improve data quality. We focus on understanding data quality barriers and facilitators from the perspectives of systems actors at many levels. Additionally, we explore the data collection and reporting processes, and data use.
Data collection and analysis were guided by the following research questions:

1. How does the data collection process contribute to data quality in Malawi’s health information system (HIS)?
2. What are the barriers and facilitators driving the quality of data in Malawi’s HIS?

The results from this research are intended to help the Ministry of Health (MOH) and partners appreciate the quality of routine health data and opportunities for quality improvement, as they finalize the Health Sector Strategic Plan (HSSPII).

Methodology

Data Collection Methods

This qualitative research study used focus group discussions and in-depth interviews to understand the perspectives of various actors in the data collection, reporting, and use process. Participants included representatives of different cadres.

Table 1. Study Participants by Cadre

<table>
<thead>
<tr>
<th>Cadre</th>
<th>Health system level</th>
<th>Role in data collection and reporting</th>
</tr>
</thead>
<tbody>
<tr>
<td>Statistical clerks, senior health surveillance assistants, and data focal persons</td>
<td>Facility—health center and district hospital</td>
<td>Data collection, aggregation and completion of monthly forms, form submission</td>
</tr>
<tr>
<td>HMIS Officer</td>
<td>District health officer</td>
<td>Review of facility forms, entry into DHIS-2, presentation of data for district health management team (DHMT)</td>
</tr>
<tr>
<td>Clinical programs M&amp;E staff</td>
<td>Central-level</td>
<td>Review of electronic data from DHIS-2, use of data for program planning and management</td>
</tr>
<tr>
<td>Directors and technical staff from the Central Monitoring and Evaluation Division</td>
<td>Central-level</td>
<td>Oversight and design of health information system</td>
</tr>
</tbody>
</table>

For the qualitative study, we purposively selected four districts (Nsanje, Blantyre, Dowa, Lilongwe) that were included in the quantitative DQA study (provide reference) to represent a range of data quality for the indicator: number of children under 5 with ARI (ARI indicator), specifically one district with high, one with low, and two with average data quality. Data quality was defined as a combination of completeness, consistency, and percentage of non-zero values for the ARI indicator in HMIS-15. We recruited statistical clerks and HMIS Officers from these four districts.

In each district, we held an in-depth interview (IDI) with the HMIS Officer and a focus group discussion (FGD) with data clerks, senior health surveillance assistants (HSAs), or others responsible for the collection and reporting of HIS data.

The Central Monitoring and Evaluation Division made a formal request for staff participation in data collection activities. A letter was sent to the district health office requesting participation...
and assistance with recruitment for this study. The HMIS Officer contacted facilities and requested participation of staff. The HMIS Officer or his/her designee was provided with airtime to recruit participants, remind them of the upcoming data collection, and communicate about logistics. The quantitative data quality assessment (DQA) study team reminded district health office leadership of the upcoming focus group discussion during their visit to the DHO at the start of district data collection. The DQA teams also reminded the health center in-charge and the data clerk about their requested participation during the facility visit. The day before the focus group, the qualitative research team called participants to remind them of their requested participation and confirmed the time and location.

Each focus group was made up of 10-13 participants. In Nsanje and Dowa, statistical clerks/data focal persons from all public sector health facilities in the district were invited to participate in the focus group. In Blantyre and Lilongwe, participants were recruited based on facility location, with priority given to those who worked at sites closest to the FGD location. We completed four focus group discussions, one FGD per district selected.

A team of three researchers were present for all FGDs. The team rotated roles throughout data collection with one member acting as group facilitator, one taking notes, and a third member managing logistics (welcoming late arrivals, handing out consent forms). All participants were read a consent document, received a copy for their records, and provided verbal consent before the study began. FGDs were audio-recoded with two recorders. Participants all provided verbal consent to be recorded before taping began.

In-depth interviews were conducted with central-level CMED staff (n=5), district-based HMIS Officers (n=4), and clinical programs M&E staff (n=4). Each of the HMIS Officers at the districts were willing to be interviewed. In one district, however, the HMIS Officer was away and so the interview was conducted with the HMIS officer’s designee.

All recorded data collection activities were transcribed from the recordings. Transcription was completed by a local research firm. As interviews and focus groups were primarily conducted in English, there was no need for translation. A sample of transcripts were checked for correctness before analysis, and edits were made for clarity using notes collected during data collection.

Data Analysis

Data were analyzed using the framework method. When using framework analysis, researchers summarize the content of each data collection activity according to the researchers’ major themes. Our themes were 1) steps in the data collection process, 2) understanding of “data quality”, 3) data quality facilitators and 4) data quality barriers.

Transcripts were independently coded by a team of researchers during a workshop held in Lilongwe, Malawi, 3-11 November 2016. Initially, all members of the research team opened one transcript. The collaborative coding process identified four main themes with detailed sub-themes; codes generated during open coding were used to create standardized definition of the themes for our analysis. For example, mention of lack of transportation was used to help define data quality barrier—systems issues. Very specific varied codes helped provide the definitions of the themes that we used during analysis. Major themes and sub-themes identified during the analysis workshop are defined in Table 1.
Table 2. Themes and sub-themes applied to DQA transcripts

<table>
<thead>
<tr>
<th>Theme</th>
<th>Sub-theme</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data collection and reporting process</td>
<td>Collection</td>
</tr>
<tr>
<td></td>
<td>Aggregation</td>
</tr>
<tr>
<td></td>
<td>Reporting and submission</td>
</tr>
<tr>
<td></td>
<td>Use</td>
</tr>
<tr>
<td></td>
<td>Feedback</td>
</tr>
<tr>
<td></td>
<td>Overview</td>
</tr>
<tr>
<td>Data quality definition</td>
<td>How do cadres define it?</td>
</tr>
<tr>
<td></td>
<td>Other thoughts</td>
</tr>
<tr>
<td>Data quality barriers</td>
<td>Systems issues</td>
</tr>
<tr>
<td></td>
<td>Lack of training/capacity</td>
</tr>
<tr>
<td></td>
<td>Others</td>
</tr>
<tr>
<td>Data quality facilitators</td>
<td>Training</td>
</tr>
<tr>
<td></td>
<td>Innovation</td>
</tr>
<tr>
<td></td>
<td>Systems or infrastructure</td>
</tr>
</tbody>
</table>

Members of the research team coded transcripts independently. Codes were applied to a range of text, from a phrase to multiple paragraphs of text. During the coding process, researchers also identified possible quotations of interest. After coding the data, the research team used the analytical framework (major themes and sub-themes) to summarize findings in a matrix. A matrix was generated for each theme with columns for sub-themes; each data collection activity, a focus group discussion or interview, had a row. After coding a transcript with the themes and sub-themes, researchers summarized the findings in the matrix. Key findings and conclusions were generated from the data by reviewing the matrix and making connections within and between participant and categories.

**Ethical Considerations**

We submitted the qualitative data quality assessment application for ethical review, as part of an overarching data quality assessment (DQA) study, which included both the quantitative and qualitative studies. The Malawi National Health Sciences Research Committee (NHSRC) granted a waiver from a full scientific and ethical review, and approved the study for implementation. All data collectors were trained in ethical research as required by international protocol on research affecting human subjects, to ensure they were well oriented to the rights that accrue to human subjects in standard research work and proper approach to undertaking research involving human subjects.

During data collection, the study team administered verbal consent to all participants. A consent form that was pre-approved by the NHSRC was read aloud to all participants before IDIs and/or FGDs commenced. Thereafter, verbal consent was obtained. Participants were provided with a copy of the consent document for their records.

**Results**

We present results by theme and sub-theme. Results include findings from all cadres of health workers, and from focus group discussions and in-depth interviews.
Participant demographics
In total, we had 52 participants. 85% were men with 15% women. On average, FGD participants had 4.4 years of experience in their positions.

Reporting process
We explore the steps in the reporting process to understand how the sequence of collection and reporting activities contribute to data quality. Through open coding, we identified major steps in the process: collection, aggregation, report submission, data use, and feedback. General comments on the collection and reporting process were coded as “overview.” We primarily focused on these steps for pediatric care with an emphasis on the outpatient department. However, respondents discussed issues outside of this clinical area and these comments are reflected in the results.

Data collection
In the outpatient department, diagnosis and treatment by clinical staff are documented in health passports, which patients carry to the data clerk or other designated focal person for entry into a register. In some cases, information may be missing from the passport; in these instances, clerks use available information to impute missing data (e.g. use medication dose to determine age), ask the caregiver (e.g. what is the child’s age?), or follow up with the clinician for clarification. Clerks report that quality of documentation varies by provider. Data on patient demographics, diagnosis, and treatment are entered into the outpatient department (OPD) register. In some facilities, this register is paper-based, in others the OPD register is electronic. The use of an electronic system at the facility depends on number of clients (should be over 3,000) and the services provided.

Data are collected throughout the month, and the goal is to have a record of all visits to the outpatient department. Challenges to this are discussed in the section entitled “Data quality barriers.”

The majority of FGD participants reported that they were responsible for completion of the outpatient department register and the HMIS-15 report, so they feel more confident of the quality of that information. Data collected in other departments by other staff may be of more variable quality. Data clerks generally thought that Health Surveillance Assistants (HSAs) submitted data that is more error prone, resulting in concerns that data were unreliable.

Aggregation
Health care utilization reports are required by the Ministry of Health. The HMIS-15 is a summary report that has measures of many clinical programs. HMIS-15 data come from different data sources; the report integrates most programs implemented by the health facility. Statistical clerks are responsible for aggregating HMIS-15 data. The statistical clerk’s role in aggregation varies by report and in accordance with who works at the facility. Some reports fall under the purview of clinical staff or other clerks. While some reports require that the data clerk tally and consolidate data manually, the electronic system produces a monthly summary for reporting.

During the aggregation step, clerks review the register and sum data. Aggregation from register to report is an area where many feel that errors are introduced.
“So we have different levels where errors are being introduced; it’s either during the aggregation from the register and form, for example, if I am tallying from the register and putting some tallies there, then I put numbers here. During the time I am transferring this data to the form, I can also make mistakes because of the errors of transcription. From this paper and the other one, now from the form. It is said the district and the data entry people will also have to enter data into the computer, so they will also introduce their own type of errors as they are typing into the DHIS-2. So the errors are introduced at different levels…” (District staff)

Aggregation may require clerks to parse out which patients fulfill criteria for inclusion in the report. For example, the number of children diagnosed with ARI, requires that clerks identify new cases of ARI in children within a certain age range. The complexity of aggregation can influence data quality. There is some concern about the correct aggregation of Helping Babies Breathe report figures:

“I will give an example of Helping Babies Breathe… some providers, I think they still don’t understand the indicators that they are reporting on….We are interested to know, out of the resuscitated babies that are captured at facility level, how many were resuscitated successfully…but you find that numbers [of babies resuscitated by specific methods] are always bigger than the total number of resuscitated babies of that particular facility.” (Clinical program staff)

**Submission**

Reports generated at health facilities are sent to the district by the 5th of the following month, and data are expected to be captured in the DHIS-2 by the 15th of the month of submission. Reports are entered into DHIS-2 by HMIS Officers at District Health Office. District staff track submission to allow for follow up with facilities that have not reported by the deadline.

“Once I get the reports from the facilities, I am supposed to have a chart to record all the reports that I have received from the facilities per program.” (District staff)

It is expected that private clinics report data through a nearby MOH facility, however reporting rates vary by private facility. We heard concerns from statistical clerks that they do not oversee the collection and aggregation of data at private facilities, so they do not know how trustworthy the data are. Additionally, the Ministry of Health and public facilities cannot force private facilities to report or adhere to reporting deadlines, so they may only receive data sporadically or after a month’s report has already been submitted. The number of private clinics falling within the catchment area of an MOH facility varies quite a bit within and across districts; there was variability in how concerned district staff were with private facility reporting.

The process of submitting the reports varies by district and within districts and even within
facilities depending on the report. In some areas, a partner project, SSDI, helps with collection of reports from the field on a monthly basis. For some reports, submission is done electronically through phones. In most cases, reports are delivered to the district through ambulances that commute between the facility and the district or carried by facility staff traveling through the district center.

Data use

"Data use is the catalyst for data quality." (CMED staff)

HIS data are used by program coordinators to understand how many people the program has reached, monitor program performance, identify what kind of supplies are needed, and track supplies used. Program coordinators should review monthly paper-based reports for correctness before entry into the DHIS-2. Data use and interest in data quality vary by coordinator. Data use improves data quality.

At the facility, specific examples of data use include review meetings and follow-up on concerning patients; facility staff may review aggregate or patient-specific data during meetings. Specifically, one clerk reported that his facility calculates the expected number of deliveries in a month based on antenatal care data, tracks home deliveries, and identifies if a lower than expected number of women are presenting at the health facility for labor and delivery, indicating possible issues with access and education. Another facility gave an example of using patient-level data to identify ART defaulters. Data use at the facility-level varies widely across facilities, with some staff reporting low levels of use and others providing examples of robust data use.

"We don’t use the data... Health facilities don’t have review meetings, just because of maybe resources, but we don’t use [the data].” (District staff)

"...not all districts are making best use of the data they collect. We have noted... that some districts only put data into the system and they think it belongs to the national level, not for them to make decisions." (District staff)

CMED and central-level clinical program staff use data for planning.

"The data is used in a number of ways, at times when we have the national level reports, we use this data for the production of national reports. In review meetings, we use the same data. And for planning at national level...even for decision-making.” (CMED staff)

Clinical programs staff discussed using data for forecasting and ordering medications. The malaria program used diagnosis data to match antimalarials dispensed to ensure that drugs were distributed appropriately and to check data quality. Additionally, data were used for monitoring trends over time, looking back over the last year, and planning programs for the future.
Feedback
The District Health Office, program coordinators, and HMIS Officers may all give feedback on the reports submitted. There may be specific review meetings or feedback can be communicated by phone. For example, HMIS staff at the DHO may ask for clarification if data trigger a verification error on entry into DHIS-2.

DHO staff provided specific examples of feedback from CMED central-level staff and CMED also discussed the process of providing this feedback. CMED alerted district staff to possible errors in the DHIS-2 data and requested verification and correction. Data correction is the responsibility of district-based staff, so CMED staff may communicate about questionable values and ask for verification, but will not edit implausible values.

Understanding data quality
Definition
The study sought to understand participants’ understanding of the term “data quality.” The respondents’ understanding of “data quality” aligns with a technical definition of data quality. Some participants used technical terms to describe data quality: completeness, accuracy, timeliness and outliers. We frequently heard that for data to be of high quality it must reflect what is “happening on the ground.” We also heard that data should be meaningful. Some respondents used “truthfulness” and “verifiable” to describe high quality data.

Outliers were understood as extreme values, data that cannot be understood, and questionable values. Understanding of data quality and its definition was not necessarily associated with good data quality, as described in the DQA quantitative report, when assessing data for completeness and consistency.

“Data quality covers a number of things. The data has to be accurate, consistent, timely, and the data should be reliable. Once the data is not reliable, then quality is not there.” (District staff)

Data quality barriers
Systems actors at all steps of the data collection and reporting pathway provide feedback on barriers to high quality data. These challenges have been categorized as systems issues, widespread challenges which relate to how the overall HIS functions or resource challenges which span across districts; training, where a gap in training contributes to lower quality data; and other barriers, a general category of important points which were not clearly categorized.

System issues
Systems challenges that affect data quality exist at all levels of the health system, and include workload, resource constraints, communication, and governance issues.

Within health facilities, many clerks cited high workloads as a hindrance in the collection of quality data. At the start of the month, data clerks and providers are expected to complete registers or forms for different disease programs, while also aggregating data from the previous month and preparing monthly reporting forms. As mentioned previously, there is a report submission deadline shortly after the start of the month. There were specific challenges mentioned in facilities with electronic OPD register: most providers were not well oriented to Baobab Electronic Data System (EDS) and may struggle with providing correct data.
At an institutional level, the system barriers included: gaps in communicating updates to tools, lack of proper terms of reference (ToRs) or indicator description, and complex registers or forms. An example given of the latter was the family planning reporting form which requires reporting on number of users by age group; to correctly transfer this data from the register to the report, clerks create ad hoc tally sheets to sum users in each age group. The process is laborious and not easily replicated to check the addition. Another challenge was around communicating change in data collection sheets and reports to users at all levels of the health system. This gap in communication had effects at facility and district. One district health office indicated that a new form was supposed to be implemented during this study’s data collection but they had not yet had time to communicate these changes to the facility staff. This gap in communication also contributed to another identified challenge: some reports and forms are not consistent. As with the previous family planning example, the report may require information that is not easily extracted from the register or other primary source document.

A concern voiced by many was the lack of verification/validation of the data by the supervisors before submission to higher level and a lack of feedback by supervisors.

“Ideally they [facility heads] are all supposed to do the validation and the validation is supposed to start from the facility. When you see the data where is not clear, you are supposed to go back, you are supposed to go back to validate. Sometimes because of the challenges in time, or sometimes people are busy and do not do that.” (Clinical program staff)

There was concern that data checks were not a routine part of the submission process because of limited time and training.

Inadequate time was one resource constraint, but participants also mentioned the lack of tangible resources as a barrier to high quality data. The study also found the following resource barriers:

• Most facilities have stock outs of materials including airtime for electronic data submission, registers and report templates.

“The moment you review [the] register, you also review the manuals and manuals contained information on how to record in each register, how build indicators, where to get the numerator for this indicator, where to get the denominator for this indicator, all those things. Because of this absence of this manual at the facility level, if am a new recruit…I would start doing my own things because there is nothing to guide me on how I should record in the register. Yes, some registers come with instructions on how to record the register, but then these do not last long. Maybe six months down the line, the first paper or cover is ripped off. Now you remain with the register without instructions. When you are stuck, where do you refer to?” (CMED staff)

• There is no dedicated, secured space in which clerks can sit at facilities and where historical data can be filed and stored.
• Lack of reliable transportation, which affects the submission of reports to the district office. Transportation of the report to the district office presents a challenge because, for example, facilities may rely on passing ambulance, clerks might pay out of pocket to travel to the facility to submit the report, or employ other piecemeal solutions to inadequate transportation.

Additionally, many respondents mentioned inadequate funds to support the district HMIS team to conduct data quality supervision.

“Due to transport problems, we often go out [to facilities for data quality checks and support] two times every month, but it was supposed to be 50/50 with 50% at the office and 50% at the facility. Or even 75% at the facility and 25% at the office. But due to [the cost of] transportation and fuel, it is often 30% at the facility and 70% at the office.” (District staff)

As mentioned previously, in private facilities, the barriers that affect data quality include lack of supervision by data clerks on collection and aggregation of data as MOH facilities cannot force private facilities to report or adhere to reporting deadlines so they may only receive data sporadically or after a month’s report has already been submitted.

“...we are supposed to include those private facilities, but there is a problem... When it comes to meetings, we are not at the gathering, so there is that break. On some indicators, yes, [they report], like family planning, but when it comes to HMIS-15 that is not happening. And it also depends on the personnel in that facility. If some are conversant in the reporting to government facilities, they come and help us, but there are some that don’t even know how to report, they just treat patients” (District staff)

Although many of the systems barriers are focused on the first steps in data collection and reporting, there are also challenges in storing and accessing the data electronically. Logistics challenges include an unreliable internet network that hinders data entry into DHIS-2. When data are entered, not all reports have had error checks completed to identify implausible values and outliers, which means that suspect data are not flagged for follow-up until review by anyone.

Users of the data cited challenges in accessing the data and concerns about quality. Not all central-level program managers have functional DHIS-2 login and this limits data use. They may have an account but are unable to log in regularly. Lack of confidence in the quality of the data spurred some clinical programs to create parallel databases, which also provided easier access to the data but required relying on district-based clinical program coordinators to collect and enter data into a program-designed database. These data are easily accessible by program managers and they feel that they have control over content, including information which may not be in DHIS-2, and content. The access issues and concerns about quality create a disconnect between end users and the data.

**Training**

Gaps in both pre-service and in-service training inhibit data quality.
There is variability in how much pre-service training facility-based statistical clerks and data focal persons received, depending on when they were hired. Newer staff have not undergone the same extensive training which was implemented for previously hired cadres.

“Statistical clerks were just employed, they didn’t undergo any training. Like the pre-service training. So they were just recruited and what is actually happening is that they are participating in on the job training, they have never undergone the pre service training...” (District staff)

On the job and in-service training play an important role for new hires.

“For instance, if a partner is conducting a training in data management and use, [newly hired clerks] are also included in the list. That’s how they learn things. So basically, I would say they were not initially trained or given specific training in terms of what they are supposed to do for the Ministry of Health” (District staff)

In-service training opportunities may vary across districts and range in length from a few hours to a week. The in-service trainings are often focused on a specific topic or tool. Although important in providing ongoing, updated information, in-service trainings were criticized as being short and too narrowly focused. Additionally, trainings may not always include multiple staff from facilities. When new programs or tools roll out, clinical staff may be trained, but statistical clerks are not always invited or sent to such trainings. The burden of training clerks on data collection and aggregation then falls to clinicians when they return to facilities.

“There were some changes to the antenatal reporting forms. Previously, we were using version 2 and that is what most of us were trained on. But now, there is a new version which has certain indicators that we don’t know about.” (District staff)

In-service trainings are often funded through partner or donor relationships.

Participants identified possible areas for future training as mathematics, statistics, data visualization, and the use of computers. Content-specific trainings included DHIS-2 and sensitization to new tools and reporting forms as they are introduced. There was concern about the introduction of new data collection and reporting tools without adequate training on how to complete them.

Data quality facilitators

Facilitators of data quality are actions, systems, or tangible objects which help improve the quality of collected and reported data. We explored data quality facilitators under four sub-themes: training, innovation, systems and infrastructure, and the role of partners. Training was defined as any formal or informal capacity-building activity which enhanced understanding or ability of staff involved in data collection. Training included both pre-service and in-service opportunities. Innovation was defined as unique tools or procedures which we heard about from only one individual, something that he/she had adopted or adapted in response to a
perceived quality or systems gap. The sub-theme of systems and infrastructure captured institutionalized processes which supported high quality collection and reporting.

Training
Although lack of training was identified as a barrier to data quality by all cadres who participated in this study, there were examples of trainings which helped improve data quality. Clerks who had been employed for a longer time received a comprehensive introduction to HMIS. More recently hired statistical clerks received in-service training. The most intensive training occurred during employment orientation, prior to beginning in the position of statistical clerk. As new tools are introduced for data collection or there were changes to existing tools, there has not been routine follow-up training.

There have been discrete opportunities for training. We heard a variety of examples from across the districts. HSA supervisors, working as data focal persons, in one district, reported that they were trained in statistics and mathematics as part of on-the-job training. Most mathematics training was done before commencement of service. Training in mathematics was reported as important for data management. Some districts, but not all, conduct annual review meetings and health workers received training after data quality self-assessments. Some programs provided on-the-job data management training. Additionally, implementation of the electronic OPD register called for training on the Baobab system.

Program coordinators have also been trained. Coordinators were trained on DHIS-2. There have been additional trainings on program-specific documentation and tools.

Innovation
Innovative methods of improving data quality have been implemented by districts and individuals. These innovations respond to specific needs as identified by those who work with data on a daily basis. Districts and individuals have implemented changes in an effort to improve data quality. These examples come from across the districts and cadres of staff interviewed:

- To address the limited use of data, a facility has set a schedule to review data, one indicator at a time. The set schedule ensures that all indicators are reviewed in a systematic fashion.
- A district developed a sheet to help with completing the HMIS-15 report. The sheet has a section for each of the indicators on the monthly form and is organized by data source to help with aggregation and minimize repeat visits to clinical departments.
- At one facility, the statistical clerk developed a tally sheet to aggregate family planning data by age groups to help complete the revised family planning report while collecting data in the old registers.
- There has been some success in using cellphone reporting to overcome reporting challenges. Respondents stated on the improvements that have resulted from use of cell phones in data collection:

  “…for people to come at a period and collect data, it can take a long time, so the introduction of phones makes it easy for us to ... send data directly without them coming to our facility.” (District staff)

- At the national level, the malaria program has placed an emphasis on high quality data. This message has been communication at all levels of data collection and reporting.
“...Now things are improving because of [their] approach that if you don’t record [data] well, if you don’t provide data of quality—good quality data—then you will not be given drugs for malaria.” (CMED staff)

Systems

There are district-based staff who participate in strengthening data quality. Program focal persons or coordinators play a key role in some districts in checking data quality and using the data for planning and decision-making. In the Expanded Program on Immunization (EPI), coordinators help prepare the report by the second of each month. Additionally, coordinators may work closely with district HMIS Officers to check data before it is entered into DHIS-2.

“If I see that in the report there is something that we need to improve, I have the knowledge [on the topic], and I call the person responsible where I see that [data] is missing.” (Clinical program staff)

Facility in-charges are supposed to check reports before submission to the district for entry into DHIS-2. Additionally, HMIS officers are supposed to check data quality before submitting reports. There is variability in how frequently in-charges and HMIS Officers complete these tasks, but we did hear from some district-based staff who reported that their work was checked for accuracy at the time of submission. In some districts, HMIS office staff travel to facilities to collect reports that have not been submitted, helping to reduce missing data.

CMED’s role is to facilitate communication between central-level programs and district and facility staff; this includes identifying and communicating about data quality issues, including following up with district staff about edits to DHIS-2 data. CMED staff also alert district-based staff to changes in data collection instruments and reporting practices. Additionally, CMED helps support clinical programs in developing usable tools for data collection to see that program needs are met and facilities are able to collect and provide data.

Partner Role

In focus group discussions, SSDI was the most commonly cited partner supporting data quality improvement. SSDI contributed to purchasing materials for data collection, providing transport for district teams to go to health facilities for supervision or data verification, funding the transport of reports from facilities to the district health office, and funding or conducting trainings. It was also be involved in data verification and supervision of staff at health centers. SSDI support varied across districts; across districts, we heard different specific details about how SSDI supported their work.

Additional partners played key supporting roles in one or two districts. One district mentioned that Riders for Health was used to help transport completed reports to the district health office from facilities. Partners have at times provided funds to hold orientation or training when new tools are introduced.

Partners also support central-level staff. They provide salary support for many staff seconded to CMED. Save the Children was mentioned as a partner which provided financial and technical
support for a data quality assessment. Partners have supported quarterly review meetings of HMIS, by providing funds, and through collaboration with CMED and clinical programs.

Conclusions

Key Findings

Data quality is shaped during the stages of data collection and reporting. At each step, there are barriers and facilitators which contribute to data quality. Innovation at an individual and systems level can contribute to improving the quality of reported data. The initiative that people take to improve data can help improve the quality of select indicators or the data collection process overall.

Each step in the data collection and reporting process contributes to the quality of HIS data. Data collection at facility-level is challenged by a heavy workload for frontline staff, limited training, and missing materials. Data collection is supported by staff initiative in following up on missing data and record reviews. Aggregation of collected data occurs monthly and is an area in which errors can be introduced. Data focal persons have developed tools to support aggregation. Submission of monthly reports are resource-intensive, requiring the transport of paper forms to the district for data entry, and may be supported by partners. We found low levels of data use across the health system. Use was shaped by perception of data quality, ease of access, and relevance of the collected data. Facilities often had limited time and technical ability to use their reported data. Although districts use data regularly during management meetings and for annual planning, they are limited in scope to the indicators which they use.

Resource constraints play a key role in influencing data quality. Limited funding may impact material availability, staff availability for data collection and reporting, the on-time transport of completed reports to districts, and the ability of district HMIS staff to travel to health centers for data-related supervision. Related to resource availability, insufficient training was a theme that emerged across the study. There is limited training for all stages of data reporting; pre-service training for staff involved in data collection and reporting (statistical clerks at facilities and HMIS staff at districts) and refresher course after introduction of new instruments and methodologies. Districts and clinical programs rely heavily on partner support for funding on-the-job trainings.

Recommendations

Data quality could be improved through systems interventions and targeted improvement activities. Key areas for action, include:

- Data use is at the discretion of individuals. Building data use supports into the DHIS-2 and requiring it as part of the job descriptions of key staff may improve use. Program indicators were under review at the time of this study. This review of indicators, a process led by CMED, allowed clinical program managers to review data in DHIS-2 and propose meaningful additions. This process may help program officers feel ownership of their data and could be further nurtured with data use supports built into DHIS-2.

- Standardized training, aligned with job description, is needed during the orientation of statistical clerks and health surveillance assistant supervisors who will act as data focal persons. A combination of pre-service and on-the-job training would orient new staff to data collection and reporting and update them as tools change. A supplementary training curriculum is needed as registers, forms, and other data collection or reporting tools are revised. Additionally, facility managers require training in reviewing data, identifying data quality issues, and using data effectively.

- Reporting and submission of data is labor- and resource-intensive. There may be ways that
cost-effective technology can be leveraged to reduce the burden of data submission. A robust DHIS-2 will help reduce parallel reporting systems, ensuring that the information needs of clinical programs are satisfied with high quality, available data.

- Reporting from private facilities should be systematized as much as possible given current policy. Private facility reporting is sporadic, may not be completed before their affiliated public sector facilities report, and is of unknown quality. Implement agreed upon reporting standards and require reporting by these facilities.

Malawi’s HIS struggles with challenges that have been found in many low and middle income countries. Although there are systemic challenges to the collection and reporting of high quality data, there are also systems components and micro-level innovations which support collection of trustworthy, accurate, timely, and complete data.

References

Appendix 1 - Consent forms

Consent Documentation—Focus Group Discussion
Data quality assessment: qualitative data collection

Introduction
I am [your name], here on behalf of the National Evaluation Platform, a collaboration between the National Statistical Office, Ministry of Health, and Johns Hopkins University. I am doing research on data collection and quality as part of the health information system. I am going to give you information and invite you to be part of this research. Before you decide, you can talk to anyone you feel comfortable with about the research.

This consent form may contain words that you do not understand. Please ask me to stop as we go through the information and I will take time to explain. If you have questions later, you can ask them of me or of another researcher.

Purpose of the Research
We are interested in collecting information on your experiences with submitting and using data as part of your work. The information that you share may be used to help provide recommendations to the Ministry of Health for improving the work that they do. We believe that your experiences are important in understanding how the process currently works and how it could be better.

Type of Research Intervention
You will be asked to participate in a focus group discussion which will involve answering questions about your experiences. This will take about one hour.

Participant Selection
You are being invited to participate in this study because of your work with the government. You have been selected based on your position title and the district in which you work.

Voluntary Participation
Your participation in this research is completely voluntary. You do not have to participate if you do not want to. The choice that you make will have not influence your job or on any work-related evaluations or reports. You may change your mind later and stop participating even if you agreed earlier to participate. You can leave the discussion or decline to answer any question.

Procedures
We are asking you to participate in a focus group discussion. If you agree to participate, we will ask you questions about your experiences collecting and submitting data. For instance, we may ask you to tell us about the process of submitting data to the district each month. I can answer questions about the research that you might have. Then I will ask you questions about your experiences with data and give you time to share your knowledge. You do not have to answer any question that you are uncomfortable answering. The discussion will be tape-recorded, but you will not be identified by name on the tape. The tape will be kept on a password protected computer. The information recorded is confidential, and no one else except approved study staff will have access to the tapes.
Duration
It is expected that this will last for no more than 90 minutes. We may have a few follow-up questions in the weeks after the discussion, but you are under no obligation to answer them. All data collection should be concluded by October 2016.

Risks
There is a risk that you may share some personal information by chance, or that you may feel uncomfortable talking about some of the topics during our discussions. If you feel uncomfortable during our discussions, you can skip the question, stop answering questions, or leave. All information that you provide will be confidential.

Benefits
We believe that the information will help improve the health system’s data. Better data can be used for resource allocation, district planning, and policy development.

Confidentiality
We ask you and others in the group not to talk to people outside the group about what was said in the group. We will ask each of you to keep what was said in the group confidential. We cannot promise that other group members will not repeat what you say. Your name will not appear in our data. You will be assigned a code and we will use that to identify your contribution. When we present results, we will not identify you by name if we use a quote from you.

Sharing the Results
The knowledge that we get from this study will be shared with your district in a written report. We will publish the results so that other interested people may learn from the research.

Right to Refuse or Withdraw
It is your choice to participate in this study. You do not have to take part in this study if you do not wish to do so, and choosing to participate will not affect your job or job-related evaluations in any way. You may stop participating in the group discussion at any time that you wish without your job being affected.

Who to Contact
If you have any questions, you can ask them now or later. If you wish to ask questions later, you may contact Amos Misomali, Resident Advisor for Johns Hopkins University in Malawi at +265 888 209 405 or amos.misomali@yahoo.com. This proposal has been reviewed by Malawi’s Research Ethics Committee and the Institutional Review Board of Johns Hopkins University.

Consent
I am going to ask you to confirm that you agree to participate in this study. Do you consent to participate?
Introduction
I am [your name], here on behalf of the National Evaluation Platform, a collaboration between the National Statistical Office, Ministry of Health, and Johns Hopkins University. I am doing research on data collection and quality as part of the health information system. I am going to give you information and invite you to be part of this research. Before you decide, you can talk to anyone you feel comfortable with about the research.

This consent form may contain words that you do not understand. Please ask me to stop as we go through the information and I will take time to explain. If you have questions later, you can ask them of me or of another researcher.

Purpose of the Research
We are interested in collecting information on your experiences with submitting and using data as part of your work. The information that you share may be used to help provide recommendations to the Ministry of Health for improving the work that they do. We believe that your experiences are important in understanding how the process currently works and how it could be better.

Type of Research Intervention
You will be asked to participate in an interview which will involve answering questions about your experiences. This will take about one hour.

Participant Selection
You are being invited to participate in this study because of your work with the government. You have been selected based on your position title and the district in which you work.

Voluntary Participation
Your participation in this research is completely voluntary. You do not have to participate if you do not want to. The choice that you make will have not influence your job or on any work-related evaluations or reports. You may change your mind later and stop participating even if you agreed earlier to participate. You can leave the discussion or decline to answer any question.

Procedures
We are asking you to participate in an interview. If you agree to participate, we will ask you questions about your experiences collecting and submitting data. For instance, we may ask you to tell us about the process of submitting data to the national level each month. I can answer questions about the research that you might have. Then I will ask you questions about your experiences with data and give you time to share your knowledge. You do not have to answer any question that you are uncomfortable answering. The discussion will be tape-recorded, but you will not be identified by name on the tape. The tape will be kept on a password protected computer. The information recorded is confidential, and no one else except approved study staff will have access to the tapes.

Duration
It is expected that this will last for no more than 60 minutes. We may have a few follow-up questions in the weeks after the discussion, but you are under no obligation to answer them. All data collection should be concluded by October 2016.
Risks
There is a risk that you may share some personal information by chance, or that you may feel uncomfortable talking about some of the topics during our discussions. If you feel uncomfortable during our discussions, you can skip the question, stop answering questions, or leave. All information that you provide will be confidential.

Benefits
We believe that the information will help improve the health system’s data. Better data can be used for resource allocation, district planning, and policy development.

Confidentiality
We will keep your information confidential and only the research team will have access to it. Your name will not appear in our data. You will be assigned a code and we will use that to identify your contribution. When we present results, we will not identify you by name if we use a quote from you.

Sharing the Results
The knowledge that we get from this study will be shared with your district in a written report. We will publish the results so that other interested people may learn from the research.

Right to Refuse or Withdraw
It is your choice to participate in this study. You do not have to take part in this study if you do not wish to do so, and choosing to participate will not affect your job or job-related evaluations in any way. You may stop participating in the group discussion at any time that you wish without your job being affected.

Who to Contact
If you have any questions, you can ask them now or later. If you wish to ask questions later, you may contact Amos Misomali, Resident Advisor for Johns Hopkins University in Malawi at +265 888 209 405 or amos.misomali@yahoo.com. This proposal has been reviewed by Malawi’s Research Ethics Committee and the Institutional Review Board of Johns Hopkins University.

Consent
I am going to ask you to confirm that you agree to participate in this study. Do you consent to participate?
Appendix 2 - Data collection instruments

Focus group discussion guide for statistical clerks

Thank you for taking the time to speak with me today. As you may know, I am here as part of the National Evaluation Platform project, a collaboration between the Malawian National Statistical Office, the Ministry of Health, and Johns Hopkins University. Everything you say today will remain confidential. Today we will talk about the routine data collected by the health sector, this is sometimes referred to as the District Health Information System (DHIS). This refers to the data which are reported by health centers, village health clinics, and hospitals on their activities.

Before we begin, we should establish a few rules for our discussion so that everyone has a chance to speak and I am able to hear all of your answers. Does anyone have any suggestions for rules? (Wait for suggestions. If none, ask that everyone silence their phones, only one speak at a time, allow everyone the opportunity to speak, keep all things discussed confidential.) Is it acceptable to record our discussion? (If yes, turn on recorder.)

Background
1. To begin, could we start by stating your job title and how long you have been in your position.
   a. As part of your training at school, what instruction did you receive on mathematics or statistics?

2. When you first began in your position, did you receive any training?
   a. If yes: Tell me about it. Probe for training on reporting, data collection, data use.
   b. If no: Have you received any training since starting? If so, in what?

Understanding of routine data
3. Tell me about the data reporting process at your facility for the district.
   a. Probe: Is the process the same at all of your facilities?
   b. Probe: Who is responsible for aggregation?
   c. Probe: Is the report reviewed before submission? By who? For what?
   d. Probe: What happens if data are missing?
   e. Probe: What deadlines are in place?

4. What are some of the challenges in reporting?

5. How does your facility use the data? What examples can you provide for how the data is used by the in-charge or other facility staff?

6. What changes in data collection tools have taken place over time? How has reporting changed since you began in your position?
   a. Probe: Were you trained when these changes were made?

7. What feedback do you receive from the district or central level on the data which you submit?
   a. Probe: Who gives feedback? How often?

8. What has been your experience with the District Health Information System (DHIS-2)?
Data quality concerns
9. I want to ask you specifically about something from the outpatient register, number of ARI cases in children under-five. Can you tell me how you report on this each month?
   a. What diagnoses in the outpatient register are counted as an ARI case?
   b. What do you do if the age of the patient is missing?
   c. What do you do if you can’t read the diagnosis?
   d. What do you do if the diagnosis is pneumonia? Is it counted as an ARI case?
   e. Can you think of cases where children are diagnosed with ARI but not included in the register? Tell me about these examples.

10. Tell me about any changes over time with how you sum the ARI data.

11. We’re going to talk about data quality. How do you understand data quality?

12. What are the data quality issues with the ARI data?
   a. Probe: How do you know if the data are good or bad quality?

13. Overall, how well does the data that you report show what is going on in your facility?
   a. Which data are the most challenging to collect and report?
   b. Which data have the most problems?

14. What do you do when you think that the data that you report are incorrect?
   a. Are there ways to edit or correct the data?
   b. Probe: Tell me about a time that you had concerns about data quality. What were the problems? What did you do?

15. What feedback do you receive from district or central level on data quality?

16. What data collection or quality training activities have you participated in?

17. Are there trainings that you would like to receive?
   a. Probe: What trainings or skills could make your job easier?
   b. Probe: Are there things that you would like to do but need more training on?

Systems issues
18. What tools could make the reporting process easier?

19. What information does your district or facility need that you don’t currently collect?

20. We brought you here because of your responsibility for summing data to complete monthly health reports. What other tasks are you responsible for?
   a. What are the pressures of the job?

21. As we conclude, are there any other thoughts that you would like to share on data, data collection, reporting, or quality?
Interview guide for DHO HMIS officer

Thank you for taking the time to speak with me today. As you may know, I am here as part of the National Evaluation Platform project, a collaboration between the Malawian National Statistical Office, the Ministry of Health, and Johns Hopkins University. Everything you say today will remain confidential. Today we will talk about the routine data collected by the health sector, this is sometimes referred to as the District Health Information System (DHIS). This refers to the data which are reported by health centers, village health clinics, and hospitals on their activities.

Background
1. To begin, could you start by stating your job title and your educational qualifications?
   a. As part of your training at school, what instruction did you receive on mathematics or statistics?

2. What are the major responsibilities of your position?

3. When you first began in your position, did you receive any training?
   a. If yes: Tell me about it. Probe for training on reporting, data collection, data use.
   b. If no: Have you received any training since starting? If so, in what?

Understanding of routine data
4. Tell me about the data reporting process for your district.
   a. Probe: How many reports do you receive and what are they?
   b. Probe: What do you do if reports are not submitted by the facilities?
   c. Probe: How do you document that reports have been received on time? What deadlines are in place?
   d. Probe: What do you do when you receive the reports? (Probe: do you do any review?)
   e. Probe: What happens if data are missing?

5. How do you use the data that you report?
   a. How would you like to use the data?

DHIS-2
6. I’d like to ask you some questions about the DHIS-2. Are you the one who enters data into the system?
   a. If yes: Which programs or reports do you enter data for?
   b. About how long does it take you to enter data for the district and how many facilities are included in your catchment areas?

7. What has been your experience with the District Health Information System (DHIS-2)?
   a. What do you know about the data quality app that is part of the DHIS-2?
   b. How do you use the dashboard and data visualizer sections of DHIS-2?

8. I have been told that the DHIS-2 has data checks in place and you will get an error message if a value that you enter is invalid. What do you do if the value on the report triggers an error message?

9. What feedback do you receive from the central level or district leadership on the data which you submit?
   a. What do you do when you receive feedback?
Data quality concerns
10. I want to ask you specifically about information from the outpatient register and HMIS-15 report, number of ARI cases in children under-five.
   a. When you look at this information, do you think it shows the truth about ARI in your district? Why?
   b. Can you tell me about any problems that you see with this information?
   c. Who do you discuss the ARI data with? (Probe: Is there a clinical program officer at the district?
   d. How do you use the ARI data?

11. I’m going to ask you a bit about data quality. What does that term, data quality, mean to you?

12. What do you think about the quality of the ARI data?
   a. How do you know if the data are good or bad quality?
   b. What does good quality data look like?

13. What do you do when you think that the data that you report are incorrect or you are worried about the quality of the data?
   a. Are there ways to edit or correct the data?

14. Tell me about a time that you had concerns about data quality. What did you do?

15. What feedback do you receive from district or central level on data quality?

Systems issues
16. Overall, not just for ARI, how well does the data that you report show what is going on in your district?
   a. What changes could be made to reporting to better show what is going on in the district?

17. What information does your district need that you don’t currently collect?

18. Are there trainings that you could receive that would make your job easier or make you better at your job:
   a. Probe: What are they?

19. As we conclude, are there any other thoughts that you would like to share on data, data collection, reporting, or quality?
Jameson Ndawala  
National Statistical Office  
Box 333  
Zomba  

Dear Sir,  

RE: REQUEST TO CONDUCT A ROUTINE DATA QUALITY ASSESSMENT (DQA) AS ENDORSED BY THE M&E TWG TASK TEAM, ANALYSIS OF PNEUMONIA DATA, AND COLLECTION OF QUALITATIVE DATA ON FAMILY PLANNING AMONG YOUTH  

Thank for you for the above titled study that you submitted to the National Health Sciences Research Committee for review.  

The committee reviewed the study and exempted it from scientific and ethical review because it is an audit of an existing Ministry of Health (MOH) program. However, should the researchers wish to publish the results, they should share with the NHSRC.  

Kind regards from the Secretariat.  

FOR: CHAIRMAN, NATIONAL HEALTH SCIENCES RESEARCH COMMITTEE