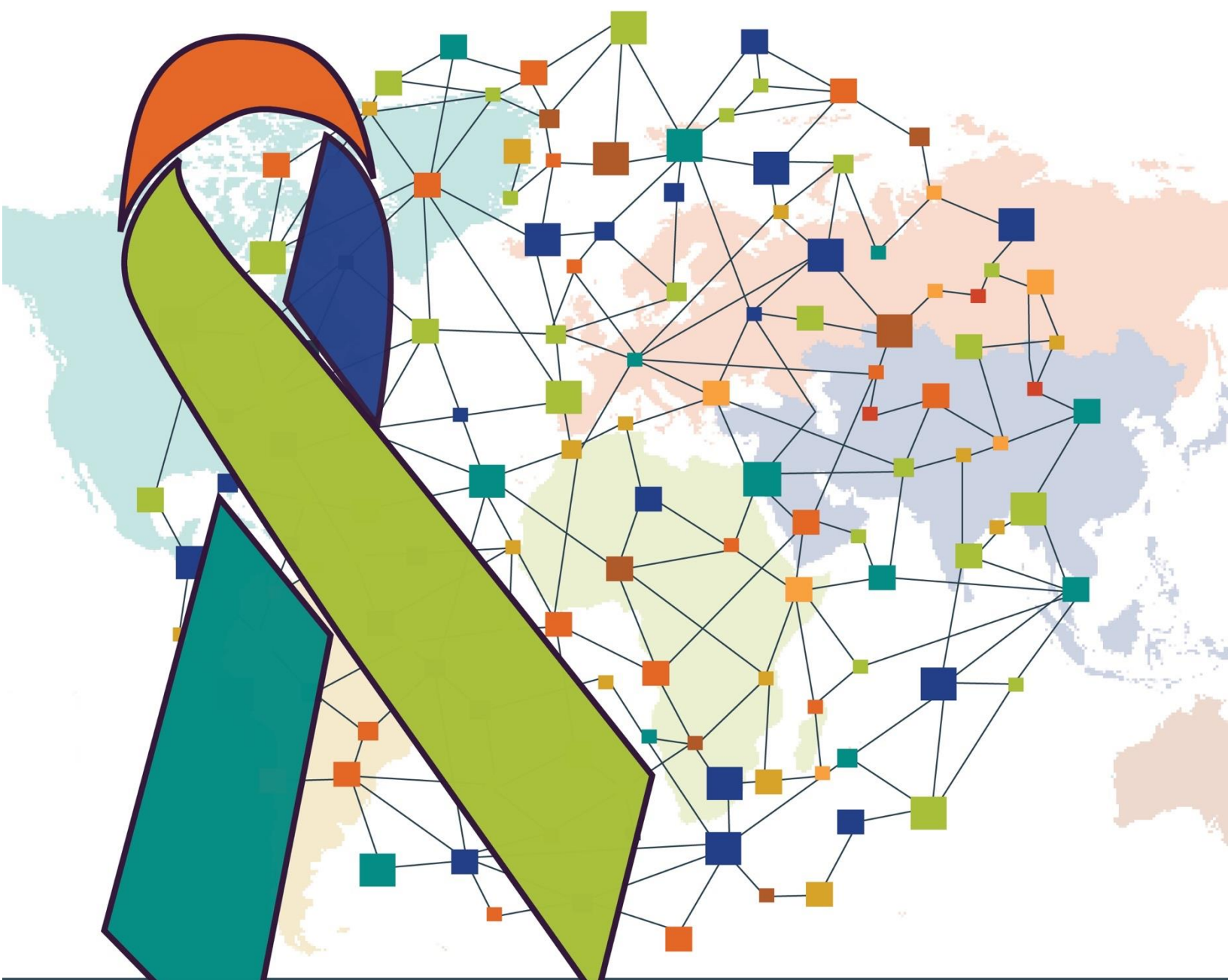


Measuring the Quality of HIV/AIDS Client-Level Data Using Lot Quality Assurance Sampling (LQAS)

September 2019

USAID
FROM THE AMERICAN PEOPLE

PEPFAR

The Global Fund
To Fight AIDS, Tuberculosis and Malaria

MEASURE Evaluation

Measuring the Quality of HIV/AIDS Client-Level Data Using Lot Quality Assurance Sampling (LQAS)

David Boone, PhD
Suzanne Cloutier, MSPH
Sergio Lins, BSc

September 2019

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350
measure@unc.edu
www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of the MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. MS-19-176

ISBN: 978-1-64232-188-3

USAID FROM THE AMERICAN PEOPLE PEPFAR The Global Fund MEASURE Evaluation

ACKNOWLEDGMENTS

We thank the United States Agency for International Development (USAID) and the United States President's Emergency Plan for AIDS Relief for their support of this work.

MEASURE Evaluation would like to thank the Bureau for Global Health/Office of HIV/AIDS at the U.S.Agency for International Development (USAID), in particular Ana Scholl, Kristen Wares, Noah Bartlett, and Webert Jose for identifying the need for rapid assessment of HIV source document completeness and providing valuable input into the tool and method.

We would also like to thank the USAID Mission in Burundi for its support in conducting the pilot test—in particular, Miriam Bassi, the United States President's Emergency Plan for AIDS Relief (PEPFAR) team leader, and Apollinaire Kavungerwa, Monitoring and Evaluation (M&E) specialist. In addition, thanks are due the MEASURE Evaluation team in Burundi for conducting the HIV/AIDS data quality assessment with local partners. Thanks especially to Serge Bisore, chief of party, Hypax Mbane, statistician, and Jean-Pierre Rwantabagu, M&E specialist.

Thanks also to the Burundi National AIDS Control Program (Programme National de la Lutte contre le Sida [PNLS]), and FHI360 and Chemonics for their help in conducting the assessment, which served as a pilot test for the tool.

This guide was prepared by David Boone, PhD, Suzanne Cloutier, MSPH, and Sergio Lins—all of MEASURE Evaluation, John Snow, Inc.

We thank the knowledge management team at MEASURE Evaluation, University of North Carolina at Chapel Hill, for editorial, design, and production services.

Suggested citation:
Boone, D., Cloutier, S., & Lins, S. (2019). Measuring the Quality of HIV/AIDS Client-Level Data Using Lot Quality Assurance Sampling (LQAS). Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina at Chapel Hill

CONTENTS

FIGURES

TABLES

ABBREVIATIONS

ART	antiretroviral therapy
CHW	community health worker
LQAS	lot quality assurance sampling
NGO	nongovernmental organization
PEPFAR	United States President's Emergency Plan for AIDS Relief
SI	sampling interval
USAID	United States Agency for International Development

GLOSSARY

Beneficiary source document	Community-level document that stores beneficiary records, such as household/family members, etc.
Comparison source document	Source document that is compared to the main source document when conducting source document concordance assessments.
Data completeness scenario	Type of data completeness assessment by either using a single-source document or comparing two source documents
Data quality level	Subjective categorization of sites (e.g., good or needs improvement) based on program implementers/managers' perception of general data quality with a focus on accuracy, timeliness, and completeness
Decision rule	Criterion for acceptability of a lot based on the completeness of its records
Lot	The data source documents at health facilities or at community level for an HIV/AIDS program
Lot quality assurance sampling	A classification method described in this guide to define acceptable and unacceptable levels of data completeness
Main source document	The source document that is sampled to obtain the records that will be assessed for completeness
Client source document	Health facility document that stores client records, such as client cards, registers, etc.
Risk	The level of sampling error that is acceptable
Sample size	The number of client/beneficiary records that will be selected for assessment from a lot
Source document	A generic term that refers to client documents at the facility level or beneficiary documents at the community level
Source document completeness	Verification that key data elements are correctly filled out in a single source document
Source document concordance	Comparison and verification of key data elements between two different source documents

INTRODUCTION

Tools and methods for assessing data quality have significantly advanced over the past 10 years, driven in part by the need for good HIV/AIDS data to inform programs. Most of the existing tools, however, focus on aggregate data at subnational levels. Very few tools measure the quality of data at the primary source–individual client documents at health facilities and beneficiary documents for community-based programs. Reviewing the quality of data in these types of documents is time consuming and resource intensive. A triage system using lot quality assurance sampling (LQAS), a rapid survey method, can be implemented to identify acceptable or unacceptable source documents using a small sample of records.

This guide was developed to describe how to sample HIV/AIDS client or beneficiary records and classify them according to quality, with a quantifiable level of confidence. A companion Excel tool—the LQAS Triage System Data Collection and Analysis Tool—is available at https://www.measureevaluation.org/resources/publications/tl-19-51.

Using the classification method of LQAS, a small set of records is sampled to evaluate and compare to a predetermined threshold of quality to identify the data quality of the lot as acceptable or unacceptable. Unacceptable lots can then be targeted for more in-depth reviews, while acceptable lots can be skipped until the next round of monitoring, where again they will be assessed via a sample. This method saves time, effort, and resources while yielding statistically sound results with quantifiable confidence and error.

The intended audience for this guide is supervisory staff. This approach may be implemented by supervisory teams already conducting data quality assurance at the health facility and community levels. When used as part of a routine system of data quality assurance, it will improve HIV/AIDS data in source documents, allowing for improved client and beneficiary management. In addition, since data quality issues will be identified and resolved at the source, aggregate data that are reported to national programs will be more accurate.

LQAS CONCEPTS

LQAS Definition

LQAS is a method of classification originating from the work developed for statistical quality control. LQAS has transitioned into the health sciences and has been used by health program implementers/managers at the level of a "supervision area" to identify priority areas or indicators that are not reaching an established benchmark. The method can provide an accurate measure of health system quality at a more aggregate level (e.g., health program area).

LQAS is considered to be a relatively rapid and inexpensive data collection approach that allows health program implementers to use small sample sizes and more frequent sampling to categorize and prioritize areas by their performance on key indicators. The smaller sample size, and thus, lower expense, is due to LQAS's primary purpose of classification rather than the derivation of a point estimate. Since its introduction in the mid-1980s (Valadez, 1991), it has been used to assess immunization coverage; post-disaster assessment of health status; women's health, such as family planning and antenatal care; growth and nutrition monitoring; diarrheal disease control; and quality management in urban zones, rural areas, and on a national scale in an increasing number of countries (Robertson & Valadez, 1982).

Lots

An ideal lot is the smallest unit that can provide meaningful information to a health implementer/manager when monitoring a health program. Commonly in LQAS, a program area is divided into supervision areas or "lots," which may consist of villages, urban zones, or health facility catchment areas. These lots are defined based on programmatic or administrative boundaries, forming programmatically relevant clusters of sample elements.

Lots can be classified as meeting or not meeting a predetermined target level of performance and the performance of different lots can be compared. This predetermined target is a crucial step in LQAS, where the main idea is to separate lots into "good performing" and "poor performing" groups to identify areas where resources can be targeted to reach a more acceptable level of performance.

To assess and measure the quality of the HIV/AIDS client/beneficiary records using LQAS method, this guide defines a lot as the collection of client or beneficiary records contained in a source document (e.g., client register) for a specific HIV/AIDS program at either the facility level (registers, client cards) or the community level (household/family member data).

Data Quality Dimensions and Levels

Data quality is a multi-dimensional concept used to define the quality of collected and reported data through a data management and reporting system. When a system performs well across these dimensions, program implementers/managers can place trust in and base decisions on the data. Systematically assessing each dimension is a necessary requirement in improving data quality. Programs, projects, and organizations need to achieve excellence across the various dimensions of data quality for program implementers/managers to have sound information on which to base program decisions and to evaluate progress toward established goals.

Among the various dimensions of data quality, the following dimensions are often the focus of data quality assessments:

- **Accuracy**: Data faithfully describe the quality and quantity of service delivery.
- **Data completeness**: The expected data elements are available in the expected place. This refers to data fields with no missing values in the source documents and reports, as well as no data consistency errors, e.g., age = 8 in an adult register.
- **Report completeness**: The availability of reports/data at all levels of the reporting system
- **Timeliness**: Data are available when needed. Data are timely when they are up-to-date (current) and when the information is available to inform decisions.
- **Reliability**: Data for the same client/process/service delivered are consistent across multiple data sources.

Since those dimensions are essential for routine monitoring of the quality of data collected and reported, program implementers/managers should have a sense of the overall level of data quality associated with each of their sites in a facility-based system and/or with each of their community health workers (CHWs) in a community-based system. For simplicity, health facilities can be classified prior to an assessment according to two classifications:

1. **Good data quality**, which is a strong combination of good results for the following:
   - Source documents (e.g., registers, and reports with no missing or inconsistent values)
   - Reports are submitted on time from sites to districts (facility-based system) and from CHWs to a community health supervisor at the health facility or a nongovernmental organization (NGO)/implementing partner (community-based system).
   - Source documents and reports are available at sites (facility-based system), or at NGOs or with community health supervisors (community-based system).
   - Data are reported with good accuracy (i.e., agreement between data recorded and data reported).
2. **Data quality needing improvement**, which is a moderate to weak combination of the results described at the good data quality level:
   - Some source documents or reports have missing or inconsistent values.
   - There may be delays in data compilation or report submissions.
   - Reports are sometimes unavailable at sites, at NGOs, or with community health supervisors.
   - Significant variation is encountered in data accuracy measurements.

Sites (facility-based system) and CHWs (community-based system) should be categorized for supervision purposes by using the two data quality levels as a reference. This categorization is based on the experience of program implementers/managers to identify sites and CHWs having good data quality or needing improvement for a particular HIV/AIDS program. In theory, mature programs would be expected to have better data quality than newer programs that have not had the benefit of experience.

The appropriate classification of sites or CHWs before the assessment will help with the accurate identification of those needing intervention for data quality. This categorization will also help supervisors (e.g., district staff, program implementers/managers, program coordinators, monitoring and evaluation staff) from facility-based

and community-based programs to define the sample size of client/beneficiary records to be selected for the data completeness assessment. If the data quality level is unknown, assume it to be "in need of improvement" for the first assessment.

Sampling

In the traditional LQAS method, simple random sampling is used to select a small sample from each lot. This guidance will use systematic random sampling. Systematic random sampling is a type of probability sampling method in which the sample records are selected based on a random starting point and a fixed interval, called the sampling interval, resulting in eligible records (i.e., those with service delivery results in the selected period) having a similar chance of being selected. When done correctly, this method will approximate the results of simple random sampling. This type of sampling is very simple to implement, as well as cost and time efficient.

Since the primary purpose of this LQAS analysis is to classify HIV/AIDS source documents as either having acceptable or unacceptable data completeness, a point estimate is not a priority. This allows the use of smaller samples than what would be needed to derive an estimate representative of all the records for a specific source document within an HIV/AIDS program. If a precise estimation is desired, it can be calculated by combining results from the sampled lots across facilities.

Sample Size

The parameters that are used to define the sample size (number of client/beneficiary records) are quality thresholds, the approximate size of the sample population, and the minimum acceptable probability of misclassification. The quality threshold is a reference point to determine the acceptability of the client/ beneficiary records with regard to completeness and consistency. There is an upper and lower limit and, for the purposes of this guidance, it is defined for each of the data quality levels described above.

- **Upper threshold (P_U):** Benchmark for quality established equal to or above which data quality is deemed acceptable.
- **Lower threshold (P_L):** Benchmark for quality below which service quality is deemed very unacceptable.

Data Quality	LT	UT
Good	85%	95%
Needing improvement	75%	90%

Risk

An important consideration of LQAS is the amount of sampling error that is acceptable. All sampling methods (and thus all surveys that are not conducted using a census) have some level of error. In LQAS, there are two types of error, alpha (α) and beta (β), also respectively referred to as "consumer" and "provider" risk.

The provider error (β) measures the risk, or probability, that an acceptable lot will be classified as unacceptable, while the consumer error (α) measures the risk that an unacceptable lot will be classified as acceptable. In the healthcare setting, the provider is the organization providing the healthcare interventions (most often the

government), while the consumer is the intended recipient of the healthcare intervention (e.g., the general population, pregnant women, children less than five years of age).

Provider and consumer error are also respectively known as sensitivity and specificity. A low provider error typically results in a higher consumer error, and vice versa. In general, these should be kept as small and as equal as possible. The risk to the consumers is that a healthcare intervention will be judged to be effective when it is not, and potentially, resources will not be invested to solve the problems the program needs to improve effectiveness. The risk to providers is that an acceptable level of quality will be judged unacceptable for a given intervention, and resources will be applied to strengthen the intervention that does not need it. The provider and consumer risks should be weighed and the consequences of both forms of error understood when determining the sample size.

In this guidance, the recommendations for acceptable error are the same for both data quality levels (i.e., generally good data quality and data quality needing improvement).

- **α (consumer) error**: The risk/probability of misclassifying a lot with unacceptable data quality as acceptable. Recommended – 5%.
- **β (provider) error:** The risk/probability of misclassifying a lot with acceptable data quality as unacceptable. Recommended – 10%.

Note that the α error is lower than the β error since, for our purposes, it is more important to minimize the risk of classifying poor data quality lots as acceptable than the risk of misclassifying good data quality lots as unacceptable. This will ensure that most or all lots with poor data quality are identified accurately for remedial action.

Decision Rule

The decision rule is based on desired standards for the lot and depends on all the sampling parameters described above. As a reminder, a lot is defined (for this particular application of LQAS) as the collection of client/beneficiary records recorded in a source document for a program area at a health facility or community-level catchment area. For this method, the decision rule is the number of sampled records within a particular source document that must be deemed acceptable in order for the entire lot to be deemed acceptable. If the number of acceptable sampled records is not reached, the lot must be rejected as unacceptable.

Data Completeness Scenarios

Data completeness and consistency are the data quality dimensions assessed with this method. By selecting a number of key data elements per program area at either the facility or in the community, the data completeness and consistency can be verified in two scenarios, as follows:

- **Source document completeness** (using a single source): verification of how many selected key data elements are correctly filled out in a program source document (register, client card) for each selected client record at the facility level, or how many are correctly filled out in a CHW form for each selected beneficiary record at the community level.
- **Source document concordance** (comparing sources): verification of how many selected key data elements correspond (i.e., have the same value) between two different data sources for each client

record at the facility level (e.g., register vs. client card) or for each beneficiary record at the community level (e.g., beneficiary form vs. household/family form).

LQAS Strengths and Limitations

Strengths:

- LQAS is an easy and inexpensive method to assess data quality using a small sample of records.
- The method is effective at identifying substandard lots, as well as those that meet the data completeness standards.[1]
- LQAS involves paper/pencil analyses rather than requiring computer analyses (though an automated tool has been developed to help users analyze findings).
- The method produces information that can be rapidly interpreted by program implementers and managers.
- Identifying and resolving data quality in source documents will help improve aggregate reporting from public health programs.

Limitations:

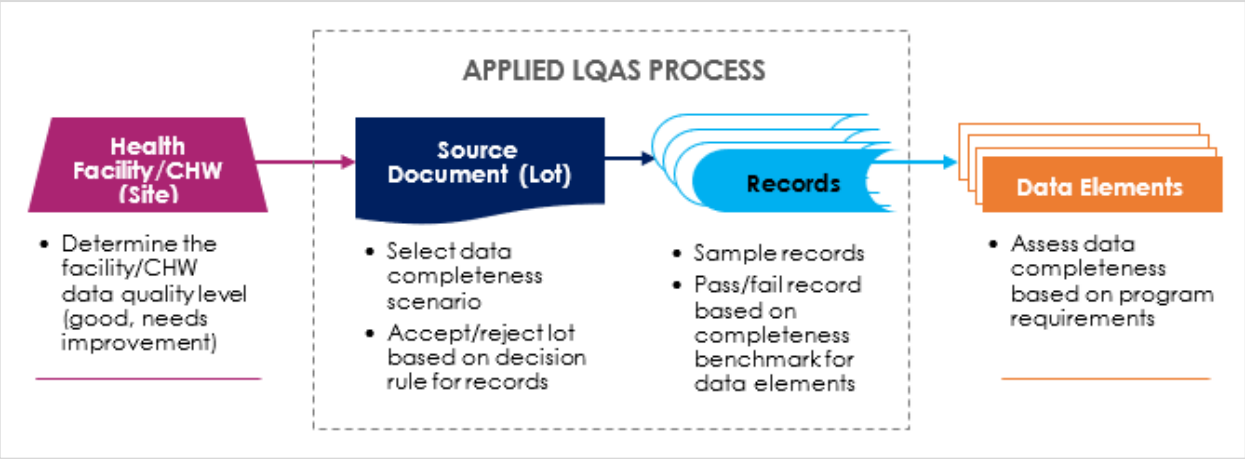
- LQAS requires updated source documents and access to all the records within those documents.
- Point estimates of data quality parameters are not possible at the facility level, but they can be calculated by combining lots to get information on source document completeness and concordance within a specific program or geographical area.
- The evaluation of consistency, or concordance, is sometimes hindered by non-standard recording of data elements. That is, the values can have different names or codes, but mean the same thing. For example, ART regimens can be recorded based on codes for constituent molecules (e.g. AZT, 3TC, etc.), or by trade names (such as "Combovir").

[1] Boone, D., Bisore, S., Hypax, M. Rwantabagu, J.P., & Ly, M. (2019). Data Quality Assessment (DQA) for HIV Program Indicators in Burundi. Chapel Hill, NC, USA: MEASURE Evaluation, University of North Carolina

DATA COMPLETENESS VERIFICATION PROCESS USING LQAS

The objective of the LQAS Triage System is to evaluate data completeness and consistency of source documents during routine supervisory visits. Given the concepts described above, a data completeness verification process can be defined using the LQAS method, as shown in the process flow in Figure 1.

Figure 1. Process for assessing data completeness in HIV/AIDS records using the LQAS method



Below are the steps required to carry out an assessment to classify sites as having client/beneficiary source documents with either acceptable or unacceptable data completeness.

Before the Site Visit

Step 1. Select the health program (e.g., HIV/AIDS, antiretroviral therapy [ART]).

Several factors should be considered when selecting health programs, such as:

- How problematic a health program is in terms of its data completeness
- The level of investment in a health program
- The complexity of the data
- The availability of guidelines to fill out the data sources
- Knowledge of personnel in the use of the data sources

Note that if more than one health program is selected for assessment, each one will require a separate LQAS analysis to determine whether or not its source documents have acceptable data completeness. For that matter, different data elements within the same health program can require different sampling schemes; a data element with reasonably good suspected completeness would require a higher quality threshold than a data element with poor suspected completeness in order to identify poor performing lots. (The LQAS Triage System Data Collection and Analysis Tool [available in MS Excel at www.measureevaluation.org/resources/publications/tl-19-51] requires the selection of one data quality range corresponding to either "good" or "needing improvement." Data elements with different suspected data quality should be analyzed in different instances of

the tool—i.e., one for data elements in the "good" range, and one for data elements in the "needs improvement" range.)

Step 2. Determine the source document(s) and data elements to be assessed.

Once the health program is selected, the data quality scenario must be defined in order to select which source document(s) will be assessed. If a source document completeness assessment is being conducted, choose the pertinent source document, which this guidance will refer to as the main source document. If a source document concordance assessment is being conducted, choose at least two, and at most three, related documents, and designate one of the documents as the main source and the other(s) as the comparison source(s).

The selection of data elements is dependent on the indicator(s) of interest. Indicators and their associated data elements have various levels of complexity based on the health program. For instance, the data elements associated with counseling and testing (number counseled, tested, and received results) are more straightforward than the data elements associated with ART (number of adults and children currently receiving ART) because of the ART program's more complicated reporting requirements.

The key data elements can be determined based on the following criteria:

- How problematic an indicator is in terms of its data completeness
- The relationship of a data element to the indicator(s) of interest
- How essential the data element is to program monitoring, i.e., the cost of incomplete data

Step 3. Define the assessment period.

After the main source document has been chosen, determine the period for which the data completeness assessment will be performed. For longitudinal indicators, i.e., those monitored over time (e.g., current on ART), a longer assessment period may be required.

Otherwise, if supervisory visits occur frequently, start the assessment period from the date of the last supervisory visit and end on the date of the upcoming site visit. If supervisory visits are irregular or infrequent, it may make sense to pick a recent period of time, e.g., the last quarter.

Consider also the volume of data and relatedly, the client volume. If a facility only sees records of 100 clients served for the quarter, it does not make sense to sample records to save time and resources, since it is little trouble to review 100 records.

The assessment period will determine the total number of source document records from which the sample will be selected. For example, if the assessment period is 12 months, all the client/beneficiary data recorded in the main source document during that period will be eligible for sampling and must be available.

Step 4. Determine the sample size and decision rule to apply to the records within a lot.

As noted in the sample size section above, there are two data quality levels in which a health program could be categorized: (1) good data quality (e.g., mature programs, or those supported by an implementing partner) and (2) data quality needing improvement (e.g., new or problematic programs). The sample size depends on which data quality level is applicable and the size of the facility in terms of client volume (i.e., the size of the "population" from which we are sampling). Thus, each facility will have a different sample size and decision rule, depending on the "size" of the facility. The data quality thresholds should be the same for all facilities. If the data quality level is unknown, assume it to be "in need of improvement" for the first assessment.

Table 1 shows sample sizes and their associated decision rules for the recommended sampling parameters by data quality level. The LQAS decision rule is based on the sample size, as well as the quality thresholds, and acceptable error rates.

Table 1. Sample size and decision rules, by data quality level

	LQAS Sample Sizes and Decision Rules (for $\alpha = 0.05$ and $\beta = 0.10$)			
	85%–95%		75%–90%	
Facility Size (Patient Volume)	Sample Size	Decision Rule	Sample Size	Decision Rule
50	all	--	all	--
75	44	40	32	27
100	46	42	37	31
125	46	42	38	32
150	49	45	43	36
175	56	51	38	32
200	56	51	43	36
225	56	51	44	37
250	58	53	44	37
275	65	59	44	37
300	58	53	48	40
325	57	52	44	37
350	58	53	49	41
375	66	60	49	41
400	66	60	49	41
425	66	60	49	41
450	67	61	49	41
475	67	61	49	41
500	66	60	49	41
550	67	61	49	41
600	67	61	49	41

650	67	61	49	41
700	67	61	49	41
750	67	61	49	41
800	67	61	49	41
850	68	62	49	41
900	67	61	49	41
950	68	62	49	41
1000	67	61	49	41
1050	68	62	54	45
1100	68	62	49	41
1150	68	62	54	45
1200	68	62	49	41
1250	68	62	54	45
1300	68	62	54	45
1350	68	62	54	45
1400	68	62	54	45
1450	68	62	54	45
1500	68	62	54	45
1550	68	62	54	45
1600	68	62	54	45
1650	68	62	54	45
1700	68	62	54	45
1750	68	62	54	45
1800	68	62	54	45
1850	68	62	54	45
1900	68	62	54	45
1950	68	62	54	45
2000	68	62	54	45

A more extensive table of sample sizes and associated decision rules appears in the LQAS Triage Tool. If circumstances dictate that different parameters are needed for the LQAS analysis (e.g., different quality thresholds, different error rates), the sample size and decision rule can be recalculated using a sample size calculator available on the Internet. One such calculator can be found at http://lqas.spectraanalytics.com/.

Step 5. Determine the data elements to be assessed within a record.

For example, the following data elements could be assessed as part of a source document completeness assessment of an ART register or a source document concordance assessment between the client medical record (ART client card) and the ART register:

(1) Date of last ART
(2) Regimen at last ART
(3) Date of last viral load test
(4) Result of last viral load test
(5) Clinical stage at diagnosis
(6) Date medically eligible for ART
(7) ART start date
(8) Client functional status at six months
(9) Adherence to treatment regimens
(10) Client treatment status (i.e., alive and on treatment or not)

Completeness will be calculated for all data elements selected, and the number complete will determine whether the data element meets the pre-established standard or not (data element is recorded = pass; data element not recorded = fail). The same applies for a test of concordance across data sources (the values are there and match = pass; the values are not there or do not match = fail). If the number of successes (or "passes") is equal to, or greater than, the decision rule established by the sampling scheme, the lot "passes" and is deemed good quality for the whole collection of records (the "lot") for this particular assessment.

(The LQAS Triage System Data Collection and Analysis Tool can accommodate up to five data elements. If more than five are to be assessed in a given assessment, another copy of the tool can be used.)

During the Site Visit

Step 6. Determine the total number of records to be assessed.

Based on the data assessment period (Step 3), count the total number of records within this period (e.g., the last quarter, the last 12 months).

Step 7. Sample the records.

Obtain the source document(s) that contain the data elements that were chosen in Step 2. If conducting a source document concordance assessment, the records in the main source document will be sampled, and then the sampled records will be located in the comparison source document(s). If a sampled record cannot be found in the comparison source, it is recorded as a non-match.

As previously stated, systematic random sampling may be used to sample the records and requires a sampling interval (SI). The sampling interval is calculated by dividing the total number of records to be assessed (Step 6) by the sample size (Step 4). The value of the sampling interval determines the pace of the sampling.

Sampling Example for a Hypothetical Situation

An LQAS assessment will be performed for the ART program at site X for the period of January–March 2019. Site X has 325 clients active on treatment as noted in the March 2019 monthly report. It is generally considered to have a high level of data quality based on past reports, which means that per the sampling parameters in Table 1, the sample size is 57 client records (P_U = 95%, P_L = 85%; α = 0.05, β = 0.10).

To select the 57 client records in the ART register, it is necessary to calculate the sampling interval. The sampling criteria are as follows:

N = total number of ART clients active on treatment at the end of the selected assessment period

n = sampled number of ART client records, i.e., 57

SI = N / n

If N = 325, then SI = 325 / 57 = 5.7

Use systematic random sampling to sample client records following the steps below:

(1) The sampling interval is 5.7, which is not a whole number. We cannot exactly sample each 5.7th record. We need a whole number for selecting every i^{th} record, where i is the sampling interval. We also want to ensure that all client records have a chance to be sampled. If we choose 5 as the sampling interval, we will finish sampling before we reach the end of the collection of records and a small number will not have had the opportunity to be sampled. Rather, choose six, and when you reach the end of the collection, start over from the beginning. (Some records will have an increased probability of selection, but that is preferable to leaving some out entirely.)

(2) Now randomly select a starting point within the first six client records. If the records are folders in a filing cabinet (e.g., client medical records) begin with the first drawer. They are likely in order of client treatment number, which corresponds to the treatment start date. If you are using the ART register, start with the first page from the beginning of treatment. If the results are within a specific timeframe, begin on the page corresponding to the beginning of that timeframe.

(3) To select the random starting point, you can write the digits 1–6 on slips of paper and randomly select one from a concealed place (or any other method to ensure randomization of the selection of the starting record).

(4) The client number that is drawn will be the first sampled client. Assume client record two is the first sampled client.

(5) From the first sampled client (number two in the order), add six (the sampling interval), and select the next client number. In this example, the next sampled client will be the eighth client record, i.e., two + six.

(6) Continue adding six to each sampled client record until 57 client records have been selected to be assessed. If you reach the end of the records and 57 have not yet been selected, go back to the beginning of the records to continue selection. See Table 4.

(7) Note that there are three client records with a higher probability of selection (cells shaded dark purple). If you have to start over at the beginning of the list and the record is already selected, choose the next record in order and continue sampling until you have sampled your target number.

Table 2. 57 sampled records using SI=6, starting with the second client record

Random starting point

1	26	51	76	101	126	151	176	201	226	251	276	301
2	27	52	77	102	127	152	177	202	227	252	277	302
3	28	53	78	103	128	153	178	203	228	253	278	303
4	29	54	79	104	129	154	179	204	229	254	279	304
5	30	55	80	105	130	155	180	205	230	255	280	305
6	31	56	81	106	131	156	181	206	231	256	281	306
7	32	57	82	107	132	157	182	207	232	257	282	307
8	33	58	83	108	133	158	183	208	233	258	283	308
9	34	59	84	109	134	159	184	209	234	259	284	309
10	35	60	85	110	135	160	185	210	235	260	285	310
11	36	61	86	111	136	161	186	211	236	261	286	311
12	37	62	87	112	137	162	187	212	237	262	287	312
13	38	63	88	113	138	163	188	213	238	263	288	313
14	39	64	89	114	139	164	189	214	239	264	289	314
15	40	65	90	115	140	165	190	215	240	265	290	315
16	41	66	91	116	141	166	191	216	241	266	291	316
17	42	67	92	117	142	167	192	217	242	267	292	317
18	43	68	93	118	143	168	193	218	243	268	293	318
19	44	69	94	119	144	169	194	219	244	269	294	319
20	45	70	95	120	145	170	195	220	245	270	295	320
21	46	71	96	121	146	171	196	221	246	271	296	321
22	47	72	97	122	147	172	197	222	247	272	297	322
23	48	73	98	123	148	173	198	223	248	273	298	323
24	49	74	99	124	149	174	199	224	249	274	299	324
25	50	75	100	125	150	175	200	225	250	275	300	325

Step 8. Assess the completeness of the data elements.

A desk review will be conducted to inspect the sampled records within the source document(s) to identify missing or inconsistent data. Use the decision rules in Table 1 to determine whether a lot is acceptable. If N = 50 or fewer, review all the records and calculate the actual percentage of missing and discordant data.

A tool has been developed in Excel to assist in assessing the completeness and concordance of data elements and records. See below for instructions on how to use the tool.

LQAS Triage System Data Collection and Analysis Tool

A tool to facilitate the collection and analysis of data has been developed in MS Excel. The tool is generic and can be used with any health program, data source, or data elements. It can accommodate data for up to 40 health facilities at once. If more sites are to be evaluated, multiple copies of the tool can be employed.

Using the Tool

The Excel workbook contains macros to help configure the tool for use. When launching Excel, be sure to click on "Enable content" when prompted.

After selecting health facilities to evaluate for source document data quality, enter the information for each site on the Facility Info tab. The Facility Info tab has three fields that describe all sites, and seven fields specific to each site (Figure 2).

Assessment Information

- Period for review
- Quality thresholds
- Number of facilities to be reviewed

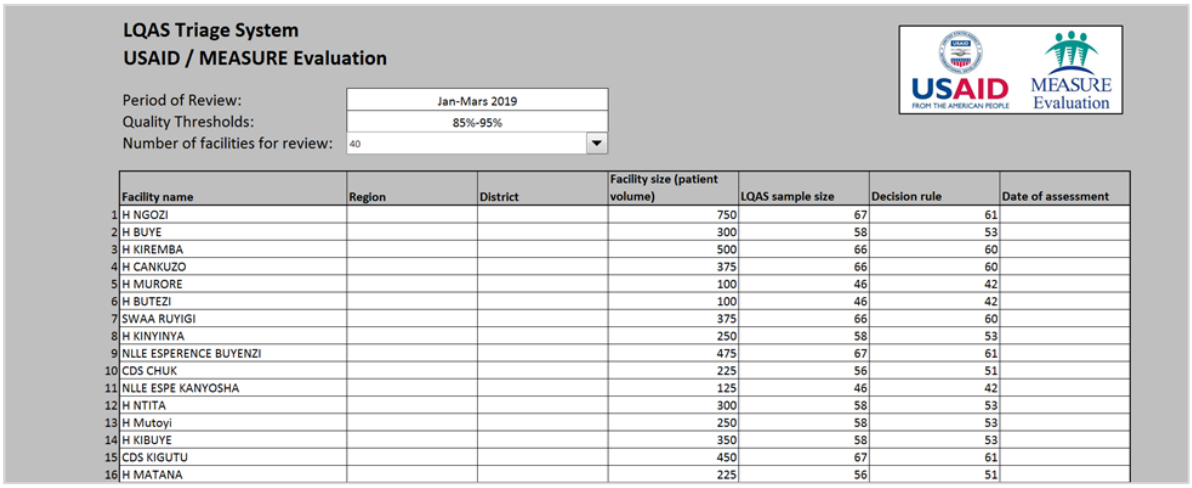
Health Facility Information

- Facility name
- Region
- District
- Facility size (client volume)
- LQAS sample size
- Decision rule
- Date of assessment (at the site)

Type the "period for review" in cell D5 and use the drop-down list to select the quality thresholds and the number of health facilities to be included in this copy of the tool. Selecting the number of sites will reveal the data collection/analysis pages for each of the sites.

Now enter the facility-specific information, with each facility on its own line. The information will automatically populate the site-specific pages in the tool. The sample sizes and decision rules are provided automatically based on the size of the facility entered in the facility size column. Use the drop-down list provided to select the nearest value to the actual number for the period selected.

Figure 2. Facility Info tab



LQAS Triage System
USAID / MEASURE Evaluation

Period of Review: Jan-Mars 2019
Quality Thresholds: 85%-95%
Number of facilities for review: 40

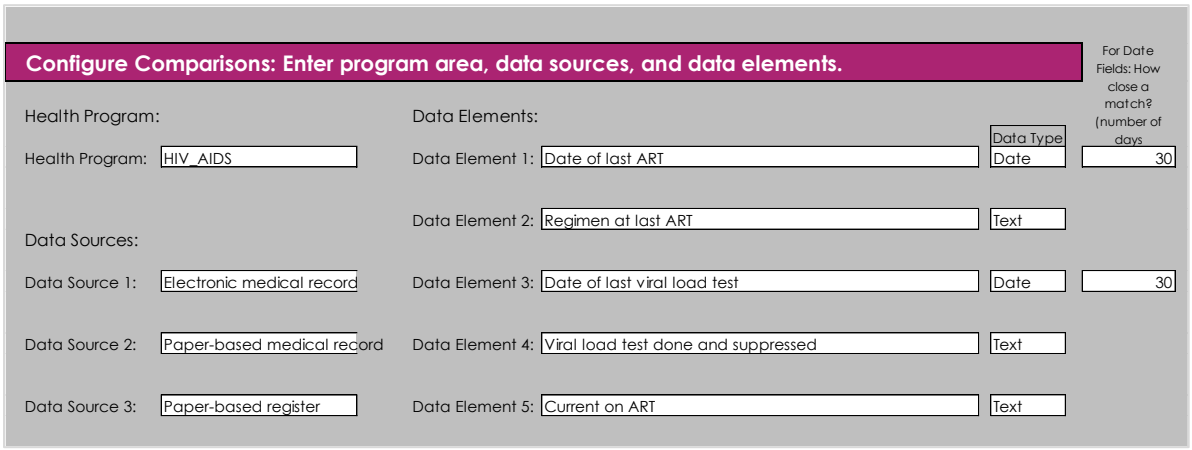
	Facility name	Region	District	Facility size (patient volume)	LQAS sample size	Decision rule	Date of assessment
1	H NGOZI			750	67	61	
2	H BUYE			300	58	53	
3	H KIREMBA			500	66	60	
4	H CANKUZO			375	66	60	
5	H MURORE			100	46	42	
6	H BUTEZI			100	46	42	
7	SWAA RUYIGI			375	66	60	
8	H KINYINYA			250	58	53	
9	NLLE ESPERENCE BUYENZI			475	67	61	
10	CDS CHUK			225	56	51	
11	NLLE ESPE KANYOSHA			125	46	42	
12	H NTITA			300	58	53	
13	H Mutoyi			250	58	53	
14	H KIBUYE			350	58	53	
15	CDS KIGUTU			450	67	61	
16	H MATANA			225	56	51	

Parameters Tab

Now select the Parameters tab and configure the tool for your specific assessment. You can select the Health Program in cell E6 using the drop-down list provided. If the health program you are evaluating does not appear on the list, you can enter a user-defined value by selecting "other (specify)" from the drop-down list and enter the health program in the space provided.

You can select up to three data sources to review and compare. Enter the data sources in cells E12, E15, and E18. Again, if the data source does not appear on the list, select "other (specify)" and enter a user-defined data source.

Figure 3. Parameters tab



Configure Comparisons: Enter program area, data sources, and data elements.

For Date Fields: How close a match? (number of days

Health Program:

Health Program: HIV_AIDS

Data Sources:

Data Source 1: Electronic medical record

Data Source 2: Paper-based medical record

Data Source 3: Paper-based register

Data Elements:

		Data Type	For Date Fields: How close a match? (number of days
Data Element 1:	Date of last ART	Date	30
Data Element 2:	Regimen at last ART	Text	
Data Element 3:	Date of last viral load test	Date	30
Data Element 4:	Viral load test done and suppressed	Text	
Data Element 5:	Current on ART	Text	

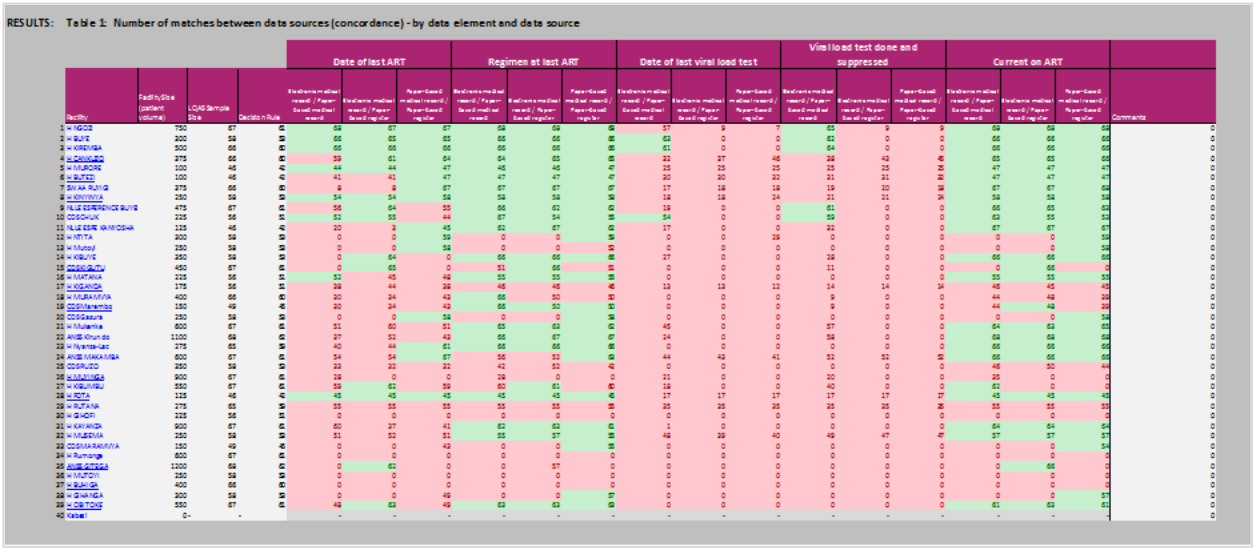
Now enter the data elements you want to assess in cells J6, J9, J12, J15, and J18 (enter user-defined values as necessary, as shown in Figure 3). For each data element, specify the data type (either text, number, or date) in the spaces provided in column L. If "date" data type is selected, you will be prompted to enter a value which represents the number of days between the two date values being compared for which a "match" is declared. For exact matching for date values, enter "0" in the fields in column N.

Analysis Tab

The Analysis tab contains four tables that each present results from the review of data elements.

- **Table 1: Number of matches by data element and data source** depicts for each data element the results of comparisons across up to three data sources (Figure 4). If the number of "matches" equals or exceeds the decision rule, the cell is colored green (the comparison meets the pre-defined standard). If the number of matches is below the threshold value for the data element, the cell is colored red. If the comparison is "not done" for whatever reason (e.g., missing data in source documents), the cell is colored grey.
- **Table 2: Concordance of data elements across data sources** shows same results as Table 1, but instead the values "yes" and "no" are depicted depending on whether the comparison for the data element meets the pre-defined standard. Further, the percentage of facilities in the sample of sites that meet the standard is calculated, as well as the percentage not meeting the standard, and the percentage "not done." Table 2 is also color-coded by result.
- **Table 3: Completeness of data elements** shows the number of complete data elements in each facility for each data element and in each data source.
- **Table 4: Completeness of data elements** displays the percentage complete of each data element for each data source and each facility. In addition, the percentage complete across facilities for each data element is provided, as well as the number and percentage of facilities having zero percent completeness.

Figure 4. Analysis tab, Table 1: Number of matches by data element and data source


RESULTS: Table 1: Number of matches between data sources (concordance) - by data element and data source

					Date of last ART			Regimen at last ART			Date of last viral load test			Viral load test done and suppressed			Current on ART			
	Facility	Facility/Site (patient volume)	LQAS Sample Size	Decision Rule	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based medical register	Paper-based medical record / Paper-based register	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based medical register	Paper-based medical record / Paper-based register	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based medical register	Paper-based medical record / Paper-based register	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based medical register	Paper-based medical record / Paper-based register	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based medical register	Paper-based medical record / Paper-based register	Comments
1	H NGOD	750	67	61	64	67	67	64	64	64	57	9	7	65	9	9	64	64	64	0
2	H BUYE	300	58	53	65	65	65	66	66	66	63	0	0	61	0	0	66	66	66	0
3	H KIREMBA	500	66	60	66	66	66	66	66	66	61	0	0	64	0	0	66	66	66	0
4	H CAHUZEZI	375	66	60	59	61	64	64	65	65	32	27	46	28	43	46	64	65	65	0
5	H MURORE	100	46	42	44	44	47	46	46	47	25	25	25	25	25	25	47	47	47	0
6	H BUTEZI	100	46	42	41	41	47	47	47	47	30	30	31	31	31	31	47	47	47	0
7	SHAA BUYIGI	375	66	60	9	9	67	67	67	67	17	16	18	19	20	18	67	67	66	0
8	H KINYINYA	350	58	53	54	54	58	58	58	58	18	18	24	21	21	24	58	58	58	0
9	HLIE EMERGENCE BUYE	475	67	61	66	64	55	66	62	62	19	0	0	61	0	0	66	65	62	0
10	CDSCHUK	225	56	51	52	55	44	67	54	55	34	0	0	59	0	0	63	55	54	0
11	HLIE ESPE VAMOSHA	125	46	42	20	3	45	63	67	63	17	0	0	22	0	0	67	67	67	0
12	H MITTA	300	58	53	0	0	59	0	0	58	0	0	29	0	0	0	0	0	59	0
13	H Mutoyi	250	58	53	0	0	58	0	0	58	0	0	0	0	0	0	0	0	58	0
14	H KIBUYE	350	58	53	0	64	0	66	66	66	27	0	0	28	0	0	66	66	66	0
15	CDSGISUTU	450	67	61	0	65	0	51	66	51	0	0	0	11	0	0	0	66	0	0
16	H MATANA	225	56	51	52	45	48	55	55	55	0	0	0	0	0	0	55	55	55	0
17	H KIGANDA	175	56	51	28	44	29	46	46	46	13	13	13	14	14	14	46	46	46	0
18	H MURAMVYA	400	66	60	30	34	43	66	50	50	0	0	0	9	0	0	44	49	39	0
19	CDSMaramba	150	49	45	20	34	43	66	50	50	0	0	0	9	0	0	44	49	39	0
20	CDSGisuru	250	58	53	0	0	58	0	0	58	0	0	0	0	0	0	0	0	58	0
21	H Mukenke	600	67	61	51	60	51	65	63	65	45	0	0	57	0	0	64	63	63	0
22	ANSE Khun do	1100	68	62	37	52	43	66	67	67	24	0	0	58	0	0	68	68	68	0
23	H Nyanse-lac	275	65	59	40	44	61	66	66	66	0	0	0	0	0	0	66	66	66	0
24	ANSE MAKAMBA	600	67	61	54	54	67	56	52	62	44	42	41	32	32	32	66	66	66	0
25	CDSRUZO	350	58	53	20	22	22	42	52	42	0	0	0	0	0	0	48	50	48	0
26	H MURINDA	900	67	61	19	0	0	19	0	0	11	0	0	10	0	0	25	0	0	0
27	H KIBUMBU	550	67	61	59	62	59	60	61	60	18	0	0	40	0	0	61	0	0	0
28	H ROTA	125	46	42	45	45	45	45	45	45	17	17	17	17	17	17	45	45	45	0
29	H BUTANA	275	65	59	55	55	55	55	55	55	25	25	25	25	25	25	55	55	55	0
30	H GIHORI	225	56	51	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
31	H KAYANZA	900	67	61	60	37	41	61	63	61	1	0	0	0	0	0	64	64	64	0
32	H MUSEMA	350	58	53	51	53	51	55	57	55	49	39	40	48	47	47	57	57	57	0
33	CDSMARAMVYA	150	49	45	0	0	43	0	0	55	0	0	0	0	0	0	0	0	54	0
34	H Rumonge	600	67	61	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
35	ANSE STEGA	1100	68	62	0	62	0	0	0	57	0	0	0	0	0	0	0	66	0	0
36	H MUTOY	250	58	53	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
37	H BUHIGA	400	66	60	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
38	H GIHOGA	300	58	53	0	0	49	0	0	57	0	0	0	0	0	0	0	0	57	0
39	H GIKITORE	550	67	61	49	63	49	63	63	63	0	0	0	0	0	0	61	63	61	0
40	Kabeti	0-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	-	1

Figure 5. Analysis tab, Table 1: Closer look

				Date of last ART			Regimen at last ART		
Facility	Facility Size (patient volume)	LQAS Sample Size	Decision Rule	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based register	Paper-based medical record / Paper-based register	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based register	Paper-based medical record / Paper-based register
H NGOZI	750	67	61	68	67	67	68	68	68
H BUYE	300	58	53	66	65	65	66	66	66
H KIREMBA	500	66	60	66	66	66	66	66	66
H CANKUZO	375	66	60	59	61	64	64	65	65
H MURORE	100	46	42	44	44	47	46	46	47
H BUTEZI	100	46	42	41	41	47	47	47	47
SWAA RUYIGI	375	66	60	8	8	67	67	67	67
H KINYINYA	250	58	53	54	54	58	58	58	58
NLLE ESPERENCE BUYEI	475	67	61	56	64	55	66	62	62
CDS CHUK	225	56	51	52	55	44	67	54	55
NLLE ESPE KANYOSHA	125	46	42	20	3	45	62	67	62
H NTITA	300	58	53	0	0	59	0	0	59
H Mutoyi	250	58	53	0	0	58	0	0	52
H KIBUYE	350	58	53	0	64	0	66	66	66
CDS KIGUTU	450	67	61	0	65	0	51	66	51
H MATANA	225	56	51	52	45	48	55	55	55
H KIGANDA	175	56	51	38	44	38	46	46	46
H MURAMVYA	400	66	60	30	34	43	66	50	50
CDS Marembo	150	49	45	30	34	43	66	50	50
CDS Gasura	250	58	53	0	0	58	0	0	58

Note the columns to the left in Table 1 (Figure 5). Facility names, size, sample size, and decision rule automatically appear in the analysis tab (and Facility Data tabs) once you have entered these data in the Facility Info tab. The names of the facilities are hyperlinked to the facility data tabs corresponding to the facility results displayed. Click on the facility name to jump to the facility data page for that facility to review results in more detail.

Health Facility Data Tabs

Each health facility in the assessment has its own page for results. Results can be typed into the appropriate columns while in the field, or data can be "cut and pasted" from copies of the tool used at health facilities to collect data. Remember to only "paste values" when pasting data into the tool. That is, do not paste formulas and formatting, but only the value of the cells being cut and pasted. (To paste only values, type alt-E-A after copying data into memory and select "values.") The cells where data are to be entered are colored white, whereas all other cells where data should not be pasted are gray.

(The cells where data should not be pasted are protected, i.e., "locked," but there is no password, and it is simple to "unlock" the cells. Please be mindful about pasting over formulas in cells that are not intended for data entry.)

Figure 6. Health Facility Data tabs

Health Facility Name:	H NGOZI	Region:	-	District:	719	Health Program:	HIV_AIDS
Facility size (patient volume)	750	Sample size:	67	Decision rule:	61	Date of assessment:	-
Period of review:	Q1 2019						

Retun to Analysis page

ACTUAL SAMPLE SIZE = 68

	Patient Number	Patient Name	Date of Birth	Sex	Electronic medical record					
					Date of last ART	Regimen at last ART	Date of last viral load test	Viral load test done and suppressed	Current on ART	Comments
1	003367		15-Jul-75	F	20-03-19	TDF/3TC/EFV	15-06-18	ind	Active	
2	003318		15-Jul-85	F	07-03-19	TDF/3TC/EFV			Active	
3	003304		01-Jan-10	F	30-03-19	ABC/3TC+EFV	07-06-18	det	Active	
4	003279		15-Jun-77	F	20-03-19	TDF/3TC/EFV	20-06-18	ind	Active	
5	003274		15-Aug-87	F	22-03-19	TDF/3TC/EFV	14-09-18	ind	Active	
6	003213		15-Jul-80	F	10-03-19	TDF/3TC/EFV	14-03-19	ind	Active	
7	003200		01-Jan-58	F	25-03-19	TDF/3TC/EFV	16-04-18	ind	Active	
8	003174		01-Jan-75	F	07-03-19	TDF/3TC/EFV	13-12-18	ind	Active	
9	003164		01-Jan-92	F	22-02-19	TDF/3TC/EFV	07-12-17	ind	Active	
10	003112		01-Jan-80	M	04-03-19	TDF/3TC/EFV	07-01-19	ind	Active	
11	003043		02-Jun-94	F	11-03-19	TDF/3TC/EFV	13-06-18	ind	Active	
12	003000		01-Jan-86	M	13-03-19	TDF/3TC/EFV	07-12-16	ind	Active	

The fields at the top of the page are facility identifiers, as well as information specific to the facilities for the assessment (e.g., the client record sample size for the facility, the decision rule, the date of the assessment) (Figure 6). Note also the hyperlink in cell K2 which permits you to jump back to the analysis page from anywhere in the workbook to review aggregate results.

The Facility Data tabs can accommodate samples of up to 70 client records for five data elements across three data sources.

Assessing Completeness and Concordance

To the right of the data entry fields are auto-populated fields that assess the completeness of the data elements and the concordance between data in different data sources (see Figures 7 and 8). These fields do not require any data entry but it is a good practice to scrutinize the results to make sure that data elements are being compared appropriately and that matches are being declared according to the established criteria.

Below the grids to match data elements are smaller tables to summarize the results (rows 78–82 from column AA to BH on all Facility Data tabs). The Analysis Tables pull results from these tables for each facility.

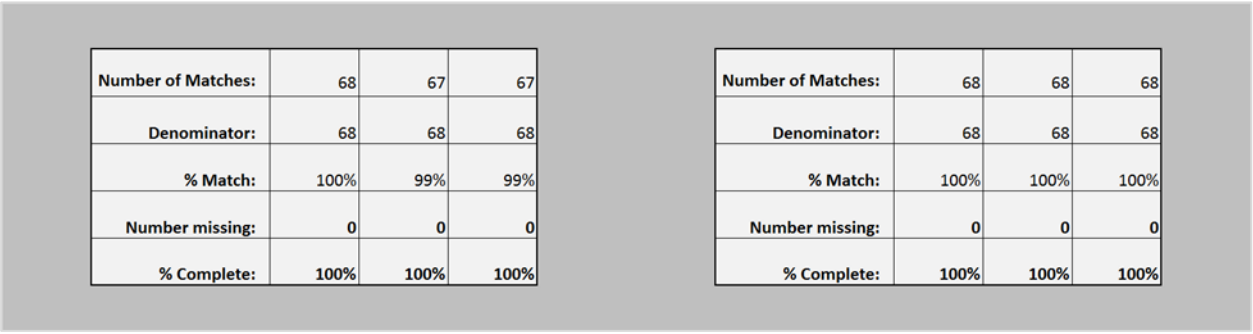
Figure 7. Health Facility data tabs, completeness and concordance of data elements

Paper-based register					
Date of last ART	Regimen at last ART	Date of last viral load test	Viral load test done and suppressed	Current on ART	Comments
20-03-19	TDF/3TC/EFV	15-06-18	ind	Active	
26-02-19	TDF/3TC/EFV			Active	
30-03-19	ABC/3TC+EFV	07-06-18	det	Active	
20-03-19	TDF/3TC/EFV	20-06-18	ind	Active	
22-03-19	TDF/3TC/EFV	14-09-18	ind	Active	
09-03-19	TDF/3TC/EFV	14-03-19	ind	Active	
25-03-19	TDF/3TC/EFV	16-04-18	ind	Active	
07-03-19	TDF/3TC/EFV	13-12-18	ind	Active	
22-02-19	TDF/3TC/EFV	07-12-17	ind	Active	
04-03-19	TDF/3TC/EFV	07-01-19	ind	Active	
11-03-19	TDF/3TC/EFV			Active	
13-03-19	TDF/3TC/EFV			Active	
19-03-19	TDF/3TC/EFV			Active	
20-03-19	AZT/3TC/NVP			Active	
15-03-19	TDF/3TC/EFV			Active	

Concordance between data sources for 'Date of last ART'					
Electronic medical record	Paper-based medical record	Paper-based register	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based register	Paper-based medical record / Paper-based register
20-03-19	20-03-19	20-03-19	0	0	0
07-03-19	07-03-19	26-02-19	0	11	11
30-03-19	30-03-19	30-03-19	0	0	0
20-03-19	20-03-19	20-03-19	0	0	0
22-03-19	29-03-19	22-03-19	7	0	7
10-03-19	10-03-19	09-03-19	0	1	1
25-03-19	25-03-19	25-03-19	0	0	0
07-03-19	07-03-19	07-03-19	0	0	0
22-02-19	22-02-19	22-02-19	0	0	0
04-03-19	04-03-19	04-03-19	0	0	0
11-03-19	13-03-19	11-03-19	2	0	2
13-03-19	11-03-19	13-03-19	2	0	2
19-03-19	19-03-19	19-03-19	0	0	0
20-03-19	20-03-19	20-03-19	0	0	0
15-03-19	15-03-19	15-03-19	0	0	0

Concordance between data sources for 'Regimen at last ART'					
Electronic medical record	Paper-based medical record	Paper-based register	Electronic medical record / Paper-based medical record	Electronic medical record / Paper-based register	Paper-based medical record / Paper-based register
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
ABC/3TC+EFV	:/3TC+EFV	:/3TC+EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1
AZT/3TC/NVP	/3TC/NVP	/3TC/NVP	1	1	1
TDF/3TC/EFV	F/3TC/EFV	F/3TC/EFV	1	1	1

Figure 8. Health Facility data tabs, calculations for completeness and concordance



Number of Matches:	68	67	67
Denominator:	68	68	68
% Match:	100%	99%	99%
Number missing:	0	0	0
% Complete:	100%	100%	100%

Number of Matches:	68	68	68
Denominator:	68	68	68
% Match:	100%	100%	100%
Number missing:	0	0	0
% Complete:	100%	100%	100%

Print Versions of the Data Collection Tools

The data collection tool can be readily printed. Print versions of the data collection tools specific to each data source are available in the final three tabs of the workbook (at the extreme right). The data collection tools are formatted to be printed in landscape orientation and on both front and back of the paper. There are 70 rows, one each for up to 70 client records. The column headers are filled automatically when you enter data in the Parameters tab.

The tabs are named:

- Data Source 1_print
- Data Source 2_print
- Data Source 3_print

REFERENCES

Robertson, S. E., & Valadez, J. J. (2006). Global review of health care surveys using lot quality assurance sampling (LQAS), 1984-2004. *Social Science & Medicine,* 63 (6), 1648–1660. Retrieved from http://doi.org/10.1016/j.socscimed.2006.04.011.

Valadez, J. J. (1991). *Assessing child survival programs in developing countries: Testing lot quality assurance sampling.* Cambridge, MA, USA: Harvard University Press.

APPENDIX. LQAS DATA COMPLETENESS WORKSHEET: HEALTH FACILITY

Best Practices for Using the LQAS Data Completeness Worksheet: Health Facility

1. The supervisory team can use the LQAS Data Completeness Worksheet to monitor data completeness at the health facility under its geographic area of management. This is part of the data quality assurance process in the supportive supervision activity.
2. Standardize the coding conventions that will be used in the LQAS Triage System Data Completeness and Analysis Tool prior to the assessment so that results are consistent across teams. For example, if recording treatment regimens, all data collectors should use trade names or generic names, but not both. Also standardize how dates are recorded (e.g., two digits for day, month, year; decide whether to follow European or American date formats).
3. If you are conducting a larger data quality assessment concurrently (where the LQAS Triage System Data Completeness and Analysis Tool is being used to conduct cross-checks) and you are sampling records that also need to be reviewed for the DQA, ensure that the sampled records are kept apart from the larger recount to avoid double-counting the records.
4. It can take a few hours to review the data sources at the facility. Sometimes these records can be in use, especially if you arrive during clinic hours. Be mindful not to disrupt the normal operations of the facility—perhaps by visiting in the afternoon when there are fewer clients, or by being flexible and reviewing what happens to be available and waiting for a source that is in use.

   If sampling from an electronic source (e.g., electronic medical records), it may be possible to export the required data in an electronic format (e.g., an Excel or .CSV data file). This can save considerable time and also avoid tying up the electronic source while data are being abstracted.
5. Remember to paste only "values" in the Excel tool. If you cut and paste formulas and formatting in the data entry pages, the results will not necessarily reflect accurately the data quality at the facility.

LQAS Data Completeness Worksheet: Health Facility

Facility Name: _____ Region: _____, District: _____ Sample size: _____

Date of review: _____ Period of review: _____ Decision Rule: _____

					Paper-based medical record					
	Patient Number	Patient Name	Date of Birth	Sex	Date of last ART	Regimen at last ART	Date of last viral load test	Viral load test done and suppressed	Current on ART	Comments
1										
2										
3										
4										
5										
6										
7										
8										
9										
10										
11										
12										
13										
14										
15										
16										

MEASURE Evaluation
University of North Carolina at Chapel Hill
123 West Franklin Street, Suite 330
Chapel Hill, NC 27516 USA
Phone: +1 919-445-9350
measure@unc.edu
www.measureevaluation.org

This publication was produced with the support of the United States Agency for International Development (USAID) under the terms of the MEASURE Evaluation cooperative agreement AID-OAA-L-14-00004. MEASURE Evaluation is implemented by the Carolina Population Center, University of North Carolina at Chapel Hill in partnership with ICF International; John Snow, Inc.; Management Sciences for Health; Palladium; and Tulane University. Views expressed are not necessarily those of USAID or the United States government. MS-19-176

ISBN: 978-1-64232-188-3

USAID FROM THE AMERICAN PEOPLE

PEPFAR

The Global Fund To Fight AIDS, Tuberculosis and Malaria

MEASURE Evaluation