

# Pilot Testing a Gender-Integrated Routine Data Quality Assessment Tool in Kenya

## Summary of the Results

### Background

Reducing the incidence and impact of HIV in Kenya is a significant priority for the Kenyan government. In addition to increasing access to HIV testing and treatment, addressing the needs of orphans and vulnerable children (OVC), and reducing the burden of gender-based violence (GBV) are critical pathways in HIV prevention efforts. Collecting age- and sex-disaggregated data and gender-sensitive indicators provides fundamental knowledge to assess the needs of diverse populations, their access to services, and the country's progress toward controlling the HIV epidemic.

MEASURE Evaluation, in collaboration with the United States Agency for International Development (USAID) and implementing partners (IPs), pilot-tested a

new tool to collect and analyze information from a gender perspective: Routine Data Quality Assessment, Plus Gender (RDQA+G). This report summarizes the results of the RDQA+G pilot test, conducted as part of a larger initiative to assess gender and HIV data quality, build capacity, and identify best practices for improving data quality in Kenya. Gender-specific results are emphasized here to illustrate the capacity and utility of the modified assessment tool.

### Methods

We used the RDQA+G tool to collect and analyze information from a gender perspective. To develop this tool, we modified the Routine Data Quality Assessment (RDQA) to incorporate sex- and age-disaggregated indicators and gender-sensitive indicators. For example, the RDQA+G examines not only the structure, functions, and capabilities of a monitoring and evaluation (M&E) unit, but also whether relevant M&E staff have received training on gender main-



Zaddy Kibao of MEASURE Evaluation and M&E Officer Priscah Kimenzu during the RDQA+G in Kenya. Photo: Carolina Mejia, MEASURE Evaluation

streaming in the past 12 months. The tool comprises a series of questions to determine if M&E units document or adopt guidelines on how to manage gender-sensitive data at each reporting level. It also determines whether M&E staff analyze data and develop visualizations of them disaggregated by age and sex. Sex- and age-disaggregated indicators and gender-sensitive indicators are incorporated across all levels of the data quality assessment.

The RDQA+G tool has two protocols: (1) data verification, which quantitatively measures the availability, accuracy, completeness, and timeliness of data; and (2) system assessment, which qualitatively evaluates the capacity of the reporting system to produce high-quality data. The data verification protocol assessed three program-level indicators: (1) the number of clients tested for HIV who received their results; (2) the number of clients who received GBV-related services; and (3) the number of OVC who received services. These indicators were investigated in terms of accuracy, timeliness, and completeness of reporting and the availability of reports. The system assessment protocol examined six functional areas: (1) M&E structure, functions, and capabilities; (2) indicator definitions and reporting guidelines; (3) data collection and reporting forms and tools; (4) data management processes; (5) evidence-based decision making; and (6) links with the national reporting system.

#### Systems assessment score ranges

0=Not applicable

1=No, not at all

2=Party

3=Mostly

4=Yes, completely

<sup>1</sup> Gender mainstreaming is a strategy for promoting gender equality. It involves ensuring that gender perspectives are central to all program or organizational activities.

The RDQA+G pilot test was conducted from November 29, 2016–December 8, 2016 in a sample of five

### Verification factor interpretation

Over 100 percent=underreporting

Under 100 percent=overreporting

service delivery points (SDPs) and two associated M&E units in Nairobi, Kenya. Sites were selected using convenience sampling based on their gender-related work with the two IPs (we refer to them as IP1 and IP2, to preserve their privacy and confidentiality). Implementing Partner 1 strengthens service delivery and institutional capacity of county health systems in Nairobi to increase access to and use of good-quality health services—including HIV; maternal, newborn, and child health; family planning (FP); and water, sanitation, and hygiene (WASH) services. Implementing Partner 2 provides community- and home-based care that improves the quality of life for people living with or affected by HIV/AIDS, through medical and nursing care, counseling and psychological support, food and nutrition, and infection prevention. The reporting period we reviewed was that of the semiannual performance reports (SAPRs) by IPs to USAID (SAPR 16: 1 April 2016–30 September 2016).

## Results

### IP1

#### Data Verification

Data verification factors at IP1 ranged from 22 percent to 101 percent; one SDP was found to have substantial overreporting. Although nearly all records were found to be complete at the M&E unit and at one of the two SDPs, the second SDP had a much lower percentage of data completeness (73%). The availability of data and the timeliness of reporting were 100 percent for both SDPs.

#### Systems Assessments: HIV\_TST Indicator

With respect to HIV services, the two SDPs associated with IP1 had an average score of 2.8 (range 2.7–2.9) on the overall systems assessment component of the RDQA+G; however, scores varied widely by functional area. Both sites scored highly in “linkages with the national reporting system.” The lowest scores were seen in “indicator definitions and reporting guidelines,” indicating a need for improvement. The highest score across functional areas was demonstrated by the M&E unit (average score: 3.7).

Scores on the gender-specific assessment for HIV services were similar to the overall systems assessment (average score: 2.9) of SDPs (Table 1). Although both SDPs had perfect scores (4.0) for “data collection and reporting forms and tools”; “M&E structure, functions, and capabilities”; and

“linkages with the national reporting system,” much lower scores were observed across the remaining functional areas. The M&E unit demonstrated a comparable performance (average score: 2.8).

#### Systems Assessments: GEND\_GBVI Indicator

For the overall systems assessment of GBV-related services, SDPs of IP1 had an average score of 2.8; the M&E unit had an average score of 3.7. Both SDPs and the M&E unit had perfect scores for “linkages with the national reporting system.” The lowest score was for “M&E structure, functions, and capabilities.”

SDP scores were similar on the gender-specific systems assessment of GBV-related services; whereas the scores for the M&E unit were comparatively lower than the overall system assessment (Table 2). The two IP1 SDPs (average score: 2.9) and the M&E unit (average score: 2.8) had comparable scores across functional areas. Although all sites had perfect scores for “indicator definitions and reporting guidelines” and “linkages with the national reporting system,” the scores for “data management processes” and “evidence-based decision making” were low, indicating a need for improvement.



Boys on Rusinga Island, Lake Victoria, Kenya. Photo: Ryan Harvey, Flickr Creative Commons

### IP2

#### Data Verification

Data verification factors of 100 percent were calculated for both M&E units, indicating accurate reporting both of HIV testing and counseling (HTC) and OVC data to USAID. However, verification factors varied greatly across SDPs, ranging from 73 percent to 107 percent for OVC data and 33 percent to 115 percent for HTC data. These findings suggest that both underreporting and overreporting occurs at IP2 SDPs. Completeness of data, similarly, varied from 35 percent to 100 percent. Availability of data and timeliness of reporting were found to be high across all sites (100%).

**Table 1. Gender system assessment summary, IP1 on HTC\_TST**

	I	II	III	IV	V	VI	
Assessment of data management and reporting systems	M&E structure, functions, and capabilities	Indicator definitions and reporting guidelines	Data collection and reporting forms and tools	Data management processes	Evidence-based decision making	Linkages with the national reporting system	Average (per site)
M&E unit	1.0	4.0	4.0	1.0	3.0	4.0	2.8
Service delivery point							
IP1_1	4.0	4.0	4.0	1.0	1.0	4.0	3.0
IP1_2	4.0	1.0	4.0	1.0	3.0	4.0	2.8
<b>Average</b> (per functional area)	4.0	2.5	4.0	1.0	2.0	4.0	2.9

Legend: 4.0–3.1; 3.0–1.5; <1.5

**Table 2. Gender system assessment summary, IP1 on GEND\_GBV**

	I	II	III	IV	V	VI	
Assessment of data management and reporting systems	M&E structure, functions, and capabilities	Indicator definitions and reporting guidelines	Data collection and reporting forms and tools	Data management processes	Evidence-based decision making	Linkages with the national reporting system	Average (per site)
M&E unit	1.0	4.0	4.0	1.0	3.0	4.0	2.8
Service delivery point							
IP1_1	4.0	4.0	4.0	1.0	1.0	4.0	3.0
IP1_2	4.0	4.0	1.0	1.0	3.0	4.0	2.8
<b>Average</b> (per functional area)	4.0	4.0	2.5	1.0	2.0	4.0	2.9

Legend: 4.0–3.1; 3.0–1.5; <1.5

**Table 3. Gender system assessment summary for IP2 on OVC\_SERV**

	I	II	III	IV	V	VI	
Assessment of data management and reporting systems	M&E structure, functions, and capabilities	Indicator definitions and reporting guidelines	Data collection and reporting forms and tools	Data management processes	Evidence-based decision making	Linkages with the national reporting system	Average (per site)
M&E Unit	1.0	4.0	4.0	1.0	3.0	4.0	2.8
Service delivery point							
IP2_1	1.0	4.0	4.0	1.0	1.0	2.5	2.25
IP2_2	1.0	4.0	4.0	1.0	1.0	4.0	2.5
IP2_3	1.0	4.0	3.0	1.0	1.0	1.0	1.8
<b>Average</b> (per functional area)	1.0	4.0	3.7	1.0	1.0	2.5	2.2

Legend: 4.0–3.1; 3.0–1.5; <1.5

**Table 4. Gender system assessment summary for IP2 on HTC\_TST**

	I	II	III	IV	V	VI	Average (per site)
Assessment of data management and reporting systems	M&E structure, functions, and capabilities	Indicator definitions and reporting guidelines	Data collection and reporting forms and tools	Data management processes	Evidence-based decision making	Linkages with the national reporting system	
M&E unit	1.0	4.0	4.0	1.0	3.0	4.0	2.8
Service delivery point							
IP2_1	4.0	4.0	4.0	1.0	1.0	4.0	3.0
IP2_2	2.5	4.0	4.0	1.0	1.0	4.0	2.8
IP2_3	1.0	4.0	4.0	1.0	1.0	4.0	2.5
<b>Average (per functional area)</b>	2.5	4.0	4.0	1.0	1.0	4.0	2.8

Legend: 4.0-3.1; 3.0-1.5; <1.5

## Systems Assessments: OVC\_SERV Indicator

The three SDPs associated with IP2 had an average score of 3.0 on the overall systems assessment of OVC services. The highest scores were for “indicator definitions and reporting guidelines” and the lowest scores were for “linkages with the national reporting system.” The M&E unit demonstrated the strongest system performance, with an average score of 3.7 across functional areas.

A gender-specific systems assessment for OVC services was performed at three IP2 SDPs (Table 3). Scores were low relative to the overall systems assessment (average score: 2.2), with the lowest scores observed for “M&E structure, functions, and capabilities” and “data management processes.” The SDPs did not appear to use evidence in their decision making. Notably, SDPs had perfect scores for “indicator definitions and reporting guidelines.” The M&E unit had the highest average score across functional areas (average score: 2.8).

## Systems Assessments: HTC\_TST Indicator

With respect to HIV services, the three SDPs associated with IP2 had an average score of 3.1. Perfect scores were achieved for “data collection and reporting forms and tools” and “linkages with the national reporting system”; the lowest scores were for “M&E structure, functions, and capabilities” and “evidence-based decision making.” Consistent with other systems assessments, the M&E unit demonstrated the strongest overall performance (average score: 3.7).

Scores on the gender-specific systems assessment for HIV services were low relative to the overall systems assessment at

SDPs (average score: 2.8) (Table 4). Perfect scores were achieved for “indicator definitions and reporting guidelines,” “linkages with the national reporting system,” and “data collection and reporting forms and tools.” The lowest scores observed were for “data management processes” and “evidence-based decision making.”

## Gender-Specific Data Systems: Strengths and Areas for Improvement across IPs

### Strengths

- Verified data were disaggregated by age and sex for all three indicators of interest.
- HIV, GBV, and most OVC data collection and reporting tools enabled disaggregation by sex and age.
- Some staff have already had gender training; all staff were willing to integrate gender in data collection, reporting, and decision making and to learn more about how to do this.

### Areas for Improvement

- Not all staff had received gender integration training.
- The quality of age and sex reporting was inconsistent across facilities.
- Gaps in availability of data existed in M&E structures, guidelines, and evidence-informed decision making around gender.
- The OVC database did not have a field for entering the sex of the active beneficiary. This information was in the paper-based logs, but it was missing in the database, and needed to be added.

<sup>2</sup> These gender-related questions and others are listed in MEASURE Evaluation’s “Gender-In” series of fact sheets: <https://www.measureevaluation.org/our-work/gender/gender-in-series>

- Instructions were needed for OVC data collection tools; OVC and HTC data analysis; and how to address late, incomplete, inaccurate, and missing OVC reports.
- Guidelines on how to securely store gender-sensitive data were not available; filing cabinets with gender-sensitive data needed to be locked; and confidentiality protocols were not in place.

## Discussion

The RDQA+G pilot test revealed multiple strengths of the IPs' data management systems. Results show that both IPs use a performance management plan as the primary reference for all M&E-related activities. At IP1, important strengths for SDPs are having documented formal indicator definitions and data collection tools and having documented data filing system guidelines. At IP2, strengths are the M&E unit providing feedback to SDPs, having documented data filing system guidelines, and SDPs reporting in a timely manner. The RDQA+G also indicated that both IPs need to improve in the areas of standardization, consistency, and completeness of data storage.

The RDQA+G pilot test investigated not only general data management systems, but also data systems specific to gender. Thus, several gender-specific strengths across sites in Kenya were identified. Sex- and age-disaggregated data are collected for all three indicators of interest, and data collection and reporting tools allow disaggregation by sex and age. Staff are willing to learn more about integrating gender in data collection, reporting, and evidenced-informed decisions; some staff have already had gender training. Gender-related scores were generally lower than overall scores for both IPs, revealing areas for improvement. Data quality gaps exist in M&E structures, guidelines, and evidence-informed decision making around gender. Additionally, the quality of age and sex reporting is inconsistent across facilities. Finally, instructions are needed for gender-related data analysis and the management of gender-sensitive data.

## Recommendations

Based on the RDQA+G pilot-test findings, we propose the following actions to improve the implementation of PEPFAR-funded programs by IPs and the USAID mission:

### For IPs

#### Overall

- The Ministry of Health (MOH) and IPs should provide technical assistance to the SDPs through continuous education and support the supervision of reporting to the districts, to improve data quality issues related to accuracy and completeness.
- The M&E units should provide systematic and regular feedback to all SDPs on the quality of their reporting (i.e., accuracy, completeness, and timeliness).
- The MOH should provide IPs with clear indicator definitions and develop job aids for documenting and collecting data at SDPs.
- How staff enter data in logbooks when the instructions on the national registries are unclear or do not include the necessary categories needs to be standardized.
- The M&E units need to ensure that the recording and reporting system avoids the double-counting of clients within and across health facilities (e.g., a beneficiary receiving the same service twice in a reporting period or a beneficiary registered as receiving the same service in two separate locations).

## Gender

- IPs could benefit from additional training in two areas: (1) capacity building on gender analysis for decision making and (2) basic gender training for staff with minimal gender experience (accompanied by a “refresher” course for more experienced staff) on why gender is important to consider for health programs.
- Gender guidance and mainstreaming documents should be developed and disseminated, where appropriate.
- Progress in reporting accuracy, timeliness, and completeness of gender data should be monitored over time using the RDQA+G tool at regular intervals.
- Project-specific databases should include data entry fields for sex and age.

## For USAID/Mission

- Inform IPs of the most current PEPFAR expectations for gender data integration and analysis.
- Where appropriate, facilitate IP training on M&E of gender and data use, and disseminate gender mainstreaming documents.
- Use RDQA+G instead of RDQA to facilitate improved gender-specific M&E.
- Provide gender data analysis tools to facilitate data use, reporting, and evidence-informed decision making for quarterly review meetings. For example, make available a tool that helps IPs answer the following gender-related questions:
  - Are there gender differences in who is accessing antiretroviral therapy?
  - Do women need permission to seek services for themselves or their children?
  - Is there provider bias toward clients based on sex, age, sexual identity, or gender identity?

- What is the number of OVC beneficiaries in OVC programs (by sex and age)?
- What are the numbers of beneficiaries served by PEPFAR OVC programs for children and families affected by HIV (by sex and age) in each of the following categories: active beneficiaries, graduated beneficiaries, transferred beneficiaries, and exited without graduation in the reporting period from the PEPFAR OVC program?
- Has the proportion of women ages 15–49 years who experienced physical violence from an intimate partner in the past 12 months changed over time?

## Conclusion

The RDQA+G pilot was completed in Kenya successfully. The results highlighted positive practices and strengths among IPs regarding M&E and gender-sensitivity. They also uncovered lessons to be learned and gaps in gender-related data quality. We recommend specific actions to improve the quality of the data provided by IPs: developing the capacity of M&E system staff to understand and collect sex- and age-disaggregated data, to conduct gender analysis, to use gender data to inform decision making, and to facilitate gender training and data use and provide gender guidance and mainstreaming documents.

Implementing partners would benefit from receiving documentation from USAID/Kenya on their expectations for gender integration and analysis. Overall, the results from this pilot test suggest that internal RDQA+G conducted at regular intervals and coupled with supervisory activities can improve the ability of HIV, GBV, and OVC programs to provide accurate, reliable data for effective and efficient planning, implementation, and decision making.