

**Organizational unit:**  **Year of report:** 2017

**Title of evaluation report:** Independent Country Programme Evaluation - Liberia 2013-2017

**Overall quality of report:** Good **Date of assessment:** 5 October 2017

**Overall comments:** This evaluation is a thorough and clearly presented assessment of the country programme, and covers the organization's response to the 2014 ebola crisis. A comprehensive methodology was designed and mostly carried out including site visits to some of the most hard to reach counties in Liberia. The evaluators are transparent about the study's limitations. Causal connections are well established as the intervention logic for each of the four programme areas is presented as part of the context, and then throughout the effectiveness section there are tables that show the relevant output, indicators, baseline, targets and level of achievement for output areas. The Findings are supported by very clearly documented evidence. Shortcomings of the evaluation include lack of explanation of stakeholder consultation on the results of the study, and the Conclusions section which does not provide a comprehensive overview of the CP performance.

**Assessment Levels**

<b>Very Good</b>	strong, above average, best practice	<b>Good</b>	satisfactory, respectable
<b>Very Good</b>		<b>Good</b>	
<b>Fair</b>	with some weaknesses, still acceptable	<b>Unsatisfactory</b>	weak, does not meet minimal quality standards

<b>Quality Assessment Criteria</b>	<i>Insert <u>assessment level</u> followed by main <u>comments</u>. (use 'shading' function to give cells corresponding colour)</i>	
<b>1. Structure and Clarity of Reporting</b>	Yes No Partial	Assessment Level: <span style="background-color: #92d050; padding: 2px 10px;"><b>Good</b></span>
<i>To ensure the report is comprehensive and user-friendly</i>		
<b>1. Is the report easy to read and understand (i.e. written in an accessible language appropriate for the intended audience) with minimal grammatical, spelling or punctuation errors?</b>	Yes	The report is easy to read and very well structured. It is written in an accessible language with minimal grammatical and punctuation errors.
<b>2. Is the report of a reasonable length? (maximum pages for the main report, excluding annexes: 60 for institutional evaluations; 70 for CPEs; 80 for thematic evaluations)</b>	Yes	For a CPE, the report is of a reasonable length. It is 67 pages excluding the annexes.
<b>3. Is the report structured in a logical way? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?</b>	Yes	The structure of the report is logical. There is a clear distinction between analysis/findings, conclusions and recommendations. There are, however, no lessons learned.
<b>4. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological tools used (e.g. interview guides; focus group notes, outline of surveys) as well as information on the stakeholder consultation process?</b>	Partial	The annexes include the TORs, the list of interviewees and the evaluation matrix in addition to the tools. The bibliography is at the end of the report. However, the stakeholder consultation process is missing.
<i>Executive summary</i>		
<b>5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?</b>	Yes	The Executive Summary is written as a stand-alone section and presents the main results of the evaluation.
<b>6. Is there a clear structure of the executive summary, (i.e. i) Purpose, including intended audience(s); ii) Objectives and brief description of intervention; iii) Methodology; iv) Main conclusions; v) Recommendations)?</b>	Yes	The executive summary follows the proposed structure. It includes the purpose, intended audience, objectives and description of interventions, methodology used, main conclusions and recommendations.

7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is of a reasonable length: 5 pages
---	-----	--

<b>2. Design and Methodology</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
----------------------------------	----------------------	-------------------	-------------

*To ensure that the evaluation is put within its context*

1. Does the evaluation describe the target audience for the evaluation?	Yes	The evaluation describes the target audience for the evaluation. The primary users of this CPE are UNFPA Liberia, the UNFPA West and Central Africa Regional Office, UNFPA Head Quarters and donors. Other key users include the Government of Liberia, the United Nations Country Team (UNCT) and Civil Society Organizations.
---	-----	---

<b>2. Is the development and institutional context of the evaluation clearly described and constraints explained?</b>	Yes	The country context and constraints in respect to UNFPA's programmatic areas is described well and succinctly. There is also a brief section on ODA disbursements that includes comparison to neighboring countries, and compares UNFPA's budget to other UN agencies. There are 4+ pages dedicated to presenting a clear overview of the financial structure of the program. Although lengthy, it is useful for explaining the surge of funding after the ebola outbreak.
---	-----	--

<b>3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?</b>	Partial	The report does reconstruct the program logic for each thematic area. There is also a good description of how UNFPA's work is situated in the UNDAF. However, the assumptions and underlying factors that may impede/facilitate the achievement of results are not described though they are important in the description of a theory of change.
---	---------	--

*To ensure a rigorous design and methodology*

<b>4. Is the evaluation framework clearly described in the text and in the evaluation matrix? Does the evaluation matrix establish the evaluation questions, assumptions, indicators, data sources and methods for data collection?</b>	Yes	The evaluation framework is clearly described in the text and in Annex 2. The evaluation matrix does establish the evaluation questions, assumptions, indicators, data sources and methods for data collection.
---	-----	---

5. Are the tools for data collection described and their choice justified?	Yes	Multiple tools were used for data collection and justified. It is noted that the planned interviews of beneficiaries were not able to be completed due to time constraints.
--	-----	---

6. Is there a comprehensive stakeholder map? Is the stakeholder consultation process clearly described (in particular, does it include the consultation of key stakeholders on draft recommendations)?	Partial	There is a stakeholder map in Annex 3b and there is information regarding the identity of the stakeholders but as it has been noted the sample that was drawn may not fully represent the full range of stakeholders of the country programme activities. The purpose and types of groups/institutions that comprised the Evaluation Reference Group are described, and it is noted that it was not possible for them to meet before or after the field work (p.2). There is no further reference to what consultation there might have been on any part of the evaluation.
--	---------	---

7. Are the methods for analysis clearly described for all types of data?	Yes	The methods for data analysis are described in the text for all types of data.
--	-----	--

8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Partial	Methodological limitations are described but the report does not indicate what steps were taken to minimize their impacts on the quality of the findings. The study noted several limitations. There were limited numbers of interviews of beneficiaries. The reason for not interviewing beneficiaries is not completely clear - it seems that there was not time to ensure a sample in all counties - however, it seems that even a few interviews would have strengthened the methodology.
9. Is the sampling strategy described?	Yes	The evaluators explain how they aimed to achieve an illustrative sample but that, in the end, the sites and stakeholders selected were a purposive sample that might not have been illustrative of differences.
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The methodology does enable collection and analysis of disaggregated data according to age and gender. While the methodology does not specifically describe how disaggregated data are collected and analyzed, the stakeholders are gender disaggregated.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The design is appropriate for assessing cross cutting issues such as vulnerability, gender equality and human rights, especially reproductive rights. The consultants used the design to analyze these themes well in the report.

<b>3. Reliability of Data</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
<i>To ensure quality of data and robust data collection processes</i>			
<b>1. Did the evaluation triangulate data collected as appropriate?</b>	Yes	The methodology does mention that the data has been triangulated and the consultants have made reference to several sources to justify a finding in their report. There were multiple methods used to collect data, and it is noted that the analysis was developed based on triangulating information from the various sources, although there was not always specific mention of how this was done (p.2)	
<b>2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?</b>	Yes	The report relies primarily on qualitative data and primary data collected was mainly qualitative. The data are generally reliable although the proposed methodology included a survey instrument which was not used, with quantitative data being obtained from review of documents that were comprehensive and trustworthy.	
<b>3. Did the evaluation make explicit any possible limitations (bias, data gaps etc.) in primary and secondary data sources and if relevant, explained what was done to minimize such issues?</b>	Partial	The data limitations are clearly indicated but the mitigating measures were not described.	
<b>4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?</b>	Yes	There is evidence that data was collected with sensitivity to ethical considerations. The report makes reference to the use of UNEG standards several times. As an example, it was noted that focus groups were conducted with consideration of informed consent and confidentiality.	

<b>4. Analysis and Findings</b>	Yes No Partial	Assessment Level:	<b>Good</b>
<i>To ensure sound analysis and credible findings</i>			

1. Are the findings substantiated by evidence?	Yes	The evidence for the findings is clearly documented and well supported by examples. Exemplary examples are frequently provided. There is a good examination of efficiency factors within each program area. The report gives an analysis of the qualitative data and these are backed with quantitative data collected from secondary data sources.
2. Is the basis for interpretations carefully described?	Yes	The basis for interpretation is provided. For example, throughout the Effectiveness section there are charts that show the relevant output, indicators, baseline, targets and level of achievement.
3. Is the analysis presented against the evaluation questions?	Yes	The analysis is presented against the evaluation questions. The consultant did an extra step and introduced the findings under each criteria with an introduction and a short summary.
4. Is the analysis transparent about the sources and quality of data?	Yes	Sources of data are cited throughout the Findings. The evaluators were careful to explain where existing data sources were weak and the impact of this on the analysis. For example, as there was no nationally representative baseline data for ASRH indicators, the evaluators noted that the extent of improvement in young people's SRH could not be assessed (p.36). In several cases the evaluators pointed out flaws in data from documents reviewed.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Cause and effects links between the interventions and their end results have been explained. The evaluators thoroughly explain linkages between the different programs and the progress (or lack of) towards intended results. An unintended outcome highlighted was that families of some young survivors of sexual abuse are aiming to make financial gains from reporting abuse within the GE&RR program.
6. Does the analysis show different outcomes for different target groups, as relevant?	Partial	There is little differentiation of results by different groups or geographic areas within each program.
7. Is the analysis presented against contextual factors?	Yes	The analysis has been presented against contextual factors particularly the impact of the Ebola Virus outbreak and its impacts on the programme. Additionally, for several of the outputs under Efficiency there are subsections on "Constraining and facilitating factors".
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The report examines cross-cutting issues in the analysis of all program components, Gender equality, vulnerability and human rights issues have extensively been discussed. For example, the discussion about the ASRH program examines the increased risks that vulnerable groups face (p. 36).

5. Conclusions	Yes No Partial	Assessment Level:	<b>Good</b>
<i>To assess the validity of conclusions</i>			
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions flow clearly from the findings and for each conclusion the relevant criteria and program area is specified. The consultant did a good job linking the conclusion to the findings of each evaluation criteria.	

2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Partial	The conclusions are very specific and as such do not provide the bigger picture of the CP's many accomplishments, as well the challenges and underlying issues. It would have been useful if the summaries presented for each criteria in Findings section were drawn upon when formulating the conclusions.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	The conclusions seem to convey the consultants unbiased judgement. They are rooted in the findings and justified.

<b>6. Recommendations</b>	Yes No Partial	Assessment Level:	<b>Fair</b>
---------------------------	----------------------	-------------------	-------------

*To ensure the usefulness and clarity of recommendations*

<b>1. Do recommendations flow logically from conclusions?</b>	Yes	All the recommendations flow from the conclusions with references to specific conclusions. The recommendation are even mentioned under the conclusions
<b>2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?</b>	Partial	The recommendations are clearly written. They are targeted at intended users but are not action oriented. Also their resource implications have not been determined.
3. Do recommendations appear balanced and impartial?	Yes	The recommendations appear balanced and in line with the conclusions.
4. Is a timeframe for implementation proposed?	Yes	A timeframe is specified for most recommendations; usually they are to be considered for the next CP.
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	Priority levels are clear for each recommendation (Medium or High). The recommendations are designed appropriately for management response and follow up.

<b>7. Gender</b>	0 1 2 3	Assessment Level:	<b>Fair</b>
------------------	------------------	-------------------	-------------

*To assess the integration of Gender Equality and Empowerment of Women (GEEW) (\*)*

1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?	2	As GEEW is a key component of the CP it is part of the scope of analysis, however issues pertaining to empowerment of women are less visible. Examples of relevant indicators include: Number of gender-based violence survivors accessing support services in 'safe homes/one stop centres'.
2. Do evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved?	2	The evaluation criteria and questions do include GEEW issues. They do try to assess how GEEW has been streamlined during program design, implementation and monitoring. There is no criterion specific to GEEW but one of the questions is: "To what extent has UNFPA support in the area of gender equality contributed to women's empowerment and reduction of gender based violence especially in rural and difficult-to-reach communities?"
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	1	Mixed methods are used but otherwise the description of the methodology is not specific about how gender responsiveness is achieved. And, other than reporting on the number of male and female evaluation participants, there is minimal use of gender-disaggregated data.

4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	2	Gender equality and reproductive rights are addressed in Findings, Conclusions and Recommendations. However, there isn't a standalone/ specific section that provides a gender analysis of the overall CP.
---	---	--

(\*) This assessment criteria is fully based on the UN-SWAP Scoring Tool, see Annex 7. Each sub-criteria shall be equally weighted (in correlation with the calculation in the tool and totalling the scores 11-12 = very good, 8-10 = good, 4-7 = Fair, 0-3=unsatisfactory).

### Overall Evaluation Quality Assessment

	Assessment Levels (*)			
Quality assessment criteria (scoring points*)	Very good	Good	Fair	Unsatisfactory
1. Structure and clarity of reporting, including executive summary (7)		7		
2. Design and methodology (13)			13	
3. Reliability of data (11)			11	
4. Analysis and findings (40)		40		
5. Conclusions (11)		11		
6. Recommendations (11)			11	
7. Integration of gender (7)			7	
<b>Total scoring points</b>		58	42	
<b>Overall assessment level of evaluation report</b>		Good		
	Very good very confident to use	Good confident to use	Fair use with caution	Unsatisfactory not confident to use

- (\*) (a) Insert scoring points associated with criteria in corresponding column (e.g. - if 'Analysis and findings' has been assessed as 'Good', enter 40 into 'Good' column.  
 (b) Assessment level with highest 'total scoring points' determines 'Overall assessment level of evaluation report'. Write corresponding assessment level in cell (e.g. 'Fair').  
 (c) Use 'shading' function to give cells corresponding colour.

If the overall assessment is 'Fair', please explain

- How it can be used?

- What aspects to be cautious about?

Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory

The approach to obtaining and analyzing data on results is at a high standard.

**Consideration of significant constraints**

The quality of this evaluation report has been hampered by exceptionally difficult circumstances:

Yes

No

If yes, please explain: