

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

Title of evaluation report: End Evaluation of GoK/UNFPA 8th Country Programme 2014-2018. Final Submission to UNFPA Kenya Country Office August 2017.

Overall quality of report: Very Good **Date of assessment:** 1 October 2017

Overall comments: Overall, this report was excellent but with some limitations in the methodology and construction of recommendations. Mainly, the sampling methodology and method of analysis required more detail and the recommendations could have better assigned timelines and priority status in order to make them more clear and action-oriented. The evaluation was gender-responsive. The report has a clear structure and includes a specific section on lessons learned which makes this report different from the other UNFPA reports (Chapter 6.3.).The findings examine context in detail but the final report briefly assesses the results chain logic and does not reconstruct it. The methodology is appropriate for assessing the cross-cutting issues. Qualitative and quantitative data sources are justified and the data limitations are clearly indicated. The analysis is transparent about the sources and quality of data and the consultants review cause and effect links. The results for different target groups are shown. The analysis is presented against contextual factors. The report examines cross-cutting issues. This report is different from other UNFPA reports in terms of very brief and clear conclusions (1 paragraph). The recommendations are action-oriented, but some recommendations lack of information on human, financial or technical implications. GEEW was included into the evaluation scope. Evaluation matrix has GEEW indicators, the methodology is appropriate for addressing GEEW, but there are no gender-responsive data analysis techniques. Evaluation conclusions and recommendations reflect a gender analysis.

Assessment Levels

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

Quality Assessment Criteria	<i>Insert <u>assessment level</u> followed by main <u>comments</u>. (use 'shading' function to give cells corresponding colour)</i>		
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Good
<i>To ensure the report is comprehensive and user-friendly</i>			
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?	Yes	The report has a clear structure, but content navigation is not automatic. Also, the chapter "Findings" has narrative description of the Outputs-Outcomes pathways, but does not present this in table view. That makes the findings more time consuming to review.	
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)	Yes	The text of the main report meets the length for CPE, excluding the Executive Summary, it is 63 pages. The annexes are included in the report (50 pages).	
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)?	Yes	The report is structured logically. It does include a specific section on lessons learned which differentiates this report from the majority of other UNFPA reports (Chapter 6.3.).	
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?	Partial	The annexes are complete, but do not include information on the stakeholder consultation process.	

Executive summary

5. Is an executive summary included in the report, written as a stand-alone section and presenting the main results of the evaluation?	Yes	The Executive Summary is written as a stand-alone document and it presents the main results of the evaluation.
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)?	Yes	The Executive Summary follows the required structure, but it does not provide a brief description of intervention. The evaluators have included seven lessons learned (half of the page) into the Executive summary. This helps to attract attention to the report.
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is 4 pages.

2. Design and Methodology	Yes No Partial	Assessment Level:	Fair
----------------------------------	----------------------	-------------------	-------------

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 target audience is UNFPA from country to headquarters level, the evaluation reference group and network, and key stakeholders” (p. x).
---	-----	---

2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The context is well explained in the chapter 2 “Country Context” and chapter 3 “UNFPA Response and Programme Strategies,” including the analysis of the previous and current Country Programme. The findings examine context in detail, such as for instance, the consultants assess how the approach of 8th CP output on gender relates to the country context: “Approach appreciates the Kenyan context that has adopted a robust normative legal and policy framework promoting human rights and fundamental freedoms, with specific prohibition of harmful cultural practices and gender based violence” (p. 36).
--	-----	---

3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Yes	The final report assesses the results chain logic for each program component (p. 26, p. 32, p. 36, p. 39). It is reflected in the results matrix in Annex 4. There is also a note on the p. 12 of the methodology section that “The reconstructed ToC and the assumptions therein will be tested during the conduct of the evaluation.” This was done.
--	-----	--

To ensure a rigorous design and methodology

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?	Yes	The evaluation framework is included in Annex 4 for reference and broken down in each thematic section. The evaluation matrix included in the Annex presents the required items, including evaluation questions, assumptions, indicators, data sources, and methods of data collection.
--	-----	---

5. Are the tools for data collection described and their choice justified?	Yes	The tools used are described in the evaluation matrix and in the text (Section 1.3.) and justified by annex 2 (Persons Met and Institutions) and annex 3 (Documents Consulted). Key data collection tools and approaches for each target group are explained in the Table 1.1.
--	-----	--

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)?	Yes	The stakeholder mapping exercise is mentioned on p. 4, but there is no output from this exercise in the report (map, table, or diagram). Section 1.3.1. mentions stakeholder consultations (“The CPE was a participatory process”), but it is unclear if the consultants engaged stakeholders in the program logic review. Inputs from the Evaluation Reference Group are mentioned in drafting the methodology for the evaluation (p. 12). The consultants presented the draft evaluation report to the CO and to the ERG before presentation to a wider range of stakeholders.
--	-----	--

7. Are the methods for analysis clearly described for all types of data?	Partial	There is a brief section on the methods for data analysis stating that “The linked trend analysis explored the change in results over time in quantitative indicators ... and, where possible in qualitative results” (p. 3). But, the methods for analysis are not clearly described 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?)	Yes	Methodological limitations and bias are clearly explained, and mitigation measures are described. For instance, the evaluators mention the following limitation: “Team leader on board over a week after the national consultants, making it challenging to meet the original timelines for the design report” (p. 4). A review of the timeline was done to adjust.
9. Is the sampling strategy described?	No	The process for determining the sample was not adequately described. The evaluators state that “Stakeholder selection followed the UNFPA Handbook guidelines for sexual and reproductive health and for adolescents and youth” (p. 3), but do not describe how they applied the guidelines.
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The TOR requires the data to be “disaggregated by sex, age and location, where possible (p. 12). Annex 2 “List of Persons/Institutions Met” disaggregates data by organization and title of informant only. In the meantime, the findings include disaggregated data, for instance: “The first is on factors affecting early sexual debut in young women aged 15-24” (p. 40).
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The methodology is appropriate for assessing the cross-cutting issues “...of human rights, gender mainstreaming within the work of UNFPA, and synergies between programme areas were also relevant” (p. 2). The methodology also specifies that “Regarding gender equality and women’s empowerment and population dynamics, the number of IPs was small enough to allow the evaluators to include all” (p. 3).

3. Reliability of Data	Yes No Partial	Assessment Level:	Very good
<i>To ensure quality of data and robust data collection processes</i>			
1. Did the evaluation triangulate data collected as appropriate?	Yes	A section on data collection and analysis (p. 3) explains the methods used for data triangulation: document review; focus group discussions with primary and secondary beneficiaries; field visits and observation.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Qualitative and quantitative data sources are justified in the evaluation (Annex 2 “Persons Met and Institutions” and the Annex 3 “Documents Consulted”). The Limitations section (1.3.2) explains that “the team is confident that a wide, sufficiently representative range of stakeholders was reached at national and county levels” (p. 4). As there was no clear explanation of the selection of primary data sources, there is a question about the illustrativeness of the sample.	
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?	Yes	The data limitations are clearly indicated and the mitigating measures for each described (Section 1.3.2).	
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	There is evidence that data were collected with sensitivity to ethical considerations. As an example, in Annex 5 it is clear that UNEG standards of ethics were applied to ensure confidentiality of interviewees data.	

4. Analysis and Findings	Yes No Partial	Assessment Level:	Very good
<i>To ensure sound analysis and credible findings</i>			
1. Are the findings substantiated by evidence?	Yes	The evidence for the findings are indicated in footnotes. For example, the evaluators state that “By the end of 2016 the project had not met obstetric fistula repair targets...” and provide a reference to the UNFPA Kenya report (p. 30). Sources of information are explained more fully in the Annex 2 “Persons Met and Institutions,” the Annex 3 “Documents Consulted,” and the Annex 4 “Evaluation Matrix.”	
2. Is the basis for interpretations carefully described?	Yes	The basis for interpretation is presented clearly. The consultants collected quantitative data according to the indicators and presented information in terms of tables that are easy to review and interpret, such as for example, RMNCAH Project Performance on Key Indicators. There is an example from the text: “Field observation and beneficiary and IP feedback found little evidence of UNFPA-produced IEC materials, including in the drop in centres for youth and for female sex workers. Also, UNFPA-supported infrastructure and equipment was observed rarely to be UNFPA-branded...” (p. 28).	
3. Is the analysis presented against the evaluation questions?	Yes	The analysis is presented according to the evaluation questions which are mentioned in the title of each sub-chapter. For example, section 4.2.1.3 “Integration of gender and a human rights based approach” corresponds to the evaluation question 2b “To what extent has the programme integrated the cross-cutting issues of gender and human rights based approaches?”	
4. Is the analysis transparent about the sources and quality of data?	Yes	The analysis is transparent about the sources and quality of data. For example, the consultants explain that “Performance against targets has in most part been highly or fully achieved, ... although achievements could not be clearly quantified where the team could not access county baseline data...” (p. 26).	
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The evaluation report reviews cause and effect links between the outputs and outcomes of the program. However, the evaluators sometimes describe their results in the form of narrative text without clear distinction of outputs and outcomes pathways. For instance, the consultants assessed connection between the output “capacity to generate data through improved systems” and outcome “utilization of these in policy formulation and programming.” They found that UNFPA will unlikely “create the desired impact, especially at county level” (p. 41). They clarify that “the dissemination of the Maternal Mortality Survey was weak and has not achieved the desired outcomes” (p. 41). Unintended outcomes are assessed throughout the analysis, such as for instance, the report has a separate section called “Unforeseen consequences.”	
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The results for different target groups are shown. The evaluators look at the results on a county-by-county basis as well as per program beneficiaries (midwives, mothers, children, adolescents, and youth.). For example, the Sexual and Reproductive Health program has been successful in the achievement of the Output 1 in the six high maternal mortality burden counties of Isiolo, Lamu, Mandera, Marsabit, Migori, and Wajir (p. 27).	

7. Is the analysis presented against contextual factors?	Yes	The analysis is presented against contextual factors such as private sector support (US\$3 million for the SRH Output 1), conservative religious attitudes (p. 27), “robust normative legal and policy framework promoting human rights and fundamental freedoms, with specific prohibition of harmful cultural practices and gender based violence” (p. 36).
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 (pp. 31, 34, 38, and 41) as there is a specific evaluation question which addresses integration of “the cross-cutting issues of gender and human rights based approaches” (p. 24).

5. Conclusions	Yes No Partial	Assessment Level:	Very good
-----------------------	----------------------	-------------------	------------------

To assess the validity of conclusions

1. Do the conclusions flow clearly from the findings?	Yes	This report is different from other UNFPA reports by presenting very brief and clear conclusions (1 paragraph). The conclusions include reference to the specific evaluation question, and, though there are no direct links to the findings in the text, the conclusions clearly flow from them.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Yes	Conclusions describe underlying issues of the program. For example, the evaluators explain that UNFPA needs to reconsider their approach in the country, because “with declining core resources this approach is increasingly unsustainable and cutbacks have already been made” (p. 53).
3. Do the conclusions appear to convey the evaluators’ unbiased judgment?	Yes	Conclusions appear to convey the evaluators’ unbiased judgment. The evaluators provide objective arguments, such as, for instance, when they explain that “UNFPA leveraged extensive funding for the project and secured strong international and national buy in” to “increase access to and improve quality of services, create demand, build institutional capacity... .” (p. 54).

6. Recommendations	Yes No Partial	Assessment Level:	Fair
---------------------------	----------------------	-------------------	-------------

To ensure the usefulness and clarity of recommendations

1. Do recommendations flow logically from conclusions?	Yes	The recommendations all flow from the conclusions with references to specific conclusions.
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	The recommendations are clearly written. The consultants specify the intended user in the introduction of the chapter “All recommendations are addressed to the Kenya country office. There are no recommendations to the regional office of UNFPA or headquarters” (p. 58). In general, recommendations are action-oriented, but some recommendations lack information on human, financial or technical implications. For instance, recommendation 3 includes financial implications (CO should leverage resources for greater investment), but does not describe human and technical resources (p. 61).
3. Do recommendations appear balanced and impartial?	Yes	The recommendations (as with the conclusions) appear unbiased and transparent according to the findings and analysis. The consultants describe the issues and provide their comments neutrally and objectively.

4. Is a timeframe for implementation proposed?	Yes	It is said in the introduction to the recommendations that a time horizon is “planning for the next country programme” (p. 58). No more details are provided.
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.

7. Gender	0 1 2 3	Assessment Level:	Good
------------------	------------------	-------------------	-------------

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?	3	GEEW was included into the evaluation scope. One of the four thematic areas was gender equality and women’s empowerment (p. 1). Evaluation matrix has GEEW indicators. In addition, one of the three effectiveness questions deals with gender. As noted in the methodology section, “Regarding gender equality and women’s empowerment and population dynamics, the number of IPs [implementing partners] was small enough to allow the evaluators to include all.”
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	Evaluation criteria (which are based on OECD-DAC criteria) do not have a GEEW dimension: relevance, effectiveness, efficiency, and sustainability. UNFPA specific criteria do not address GEEW as well: added value, corporate strategic alignment, and responsiveness. In the meantime, the evaluation assessed the cross-cutting themes of human rights and gender mainstreaming within the work of UNFPA (p. 2). The evaluation question 4 includes gender component “To what extent has the programme integrated the cross-cutting issues of gender and human rights based approaches?” (p. 34, the Annex 4). Other examples include: for instance: “Evidence of the integration of gender and a rights based approach within the planning, programme and project documents of UNFPA; Evidence of the integration of gender and a rights based approach provided by KIs and beneficiaries” (p. 34, the Annex 4).
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	The methodology is appropriate for addressing GEEW and gender-responsive data analysis techniques were used to determine findings. Gender differences in statistics and in document review were noted throughout. The methodology explains that focus group discussions were conducted with adolescent males and females (p. 4). Desk review and data analysis stages included gender analysis (Annex 3 “Documents Consulted”). It is also said that “Regarding gender equality and women’s empowerment and population dynamics, the number of IPs was small enough to allow the evaluators to include all” (p. 3). However, there is no gender tabulation of the respondents in the Annex 2 “Persons Met and Institutions”.

4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	3	<p>Evaluation findings reflect a gender analysis as the cross-cutting themes incorporated gender and human rights, including for GBV. There are evidences of gender-responsive analysis in the text, for instance, the evaluators investigated that “Gender and human rights, including for GBV, were insufficiently integrated, and insecurity in some counties and conservative religious attitudes impeded development in some cases” (p. 27).</p> <p>Evaluation conclusions reflect a gender analysis, for instance: Strategic Conclusions # 2 (gender) and # 6 (gender mainstreaming). Program conclusions have the following examples: Sexual and Reproductive Health (#2 - GBV), Adolescents and Youth (# 3 - gender friendliness), Gender Equality and Women’s Empowerment – all four conclusions, and Population Dynamics – no GEEW issues. Evaluation recommendations reflect a gender analysis: strategic recommendations # 2 (gender integration and human rights), # 3 (gender), and # 6 (gender issues). Program recommendations include such examples: Sexual and Reproductive Health (# 2 – GBV, # 5 - adolescent girls), Adolescents and Youth (# 2 - gender equality, # 3 - gender friendly services); Gender Equality and Women’s Empowerment - all four recommendations; Population Dynamics (# 2 – gender).</p>
---	---	--

(*) 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 totaling the scores 11-12 = very good, 8-10 = good, 4-7 = Fair, 0-3=unsatisfactory).

Overall Evaluation Quality Assessment

Quality assessment criteria (scoring points*)	Assessment Levels (*)			
	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		
Total scoring points	62	14	24	
Overall assessment level of evaluation report	Very 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: