

Organizational unit:		Year of report:	2018
Title of evaluation report:	Evaluation du huitieme programme pays - UNFPA Benin 2014 - 2017		
Overall quality of report:	Fair	Date of assessment:	May 9 2018

Overall comments: The evaluation report is clear, concise and well-structured. It is written in an accessible language with minimal errors. The evaluation itself sufficiently elaborated on how GEEW was integrated in data collection, analysis and reporting. While the stakeholders have been mapped out, the consultation process was not wholly clear, including how, for example, stakeholders were (or were not) included in finalization of the recommendations. The report addresses the questions under each evaluation criteria and a mixed methods approach was used: the report made use of both qualitative data (from primary sources) and quantitative data (from secondary sources). While a results framework is included in the annexes, the evaluation does not describe the reconstruction of the intervention logic/theory of change nor does it assess its adequacy.

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: Fair
<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 is written in an accessible language and is easy to read. The language is appropriate for the intended audience and it comes with minimal grammatical spelling or punctuation errors.	
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 report is 70-page long, which is the recommended length for CPEs.	
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 in a logical way. There is a clear distinction between the analysis/findings section and the conclusions and recommendations sections.	
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 include the TORs, the list of interviewees, the data collection tools, as well as the evaluation matrix. A mapping of stakeholders is included, however the consultation process itself has not been detailed in the annex.	
<i>Executive summary</i>			
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 included in the main report. It is four-page long. It is a stand-alone section and does present the 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)?	Partial	There is a clear structure of the executive summary. The purpose of the evaluation, the intended audience, the methodology, the main conclusions and recommendations are all included. The executive summary, however, does not include the objectives or a description of the interventions.
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is concise and is 4-pages long.

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 target audience has been described in the report as being UNFPA Benin, the regional office and the programme stakeholders. Those stakeholders have been identified in the annexes under the section related to the mapping of the stakeholders.
---	-----	---

2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The development and institutional context of the evaluation has been clearly described. The constraints have also been identified.
---	-----	--

3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Partial	The annexes include a results framework for each component. Those results frameworks are referred to as being the reconstructed intervention logic but the evaluation does not describe the reconstruction of the intervention logic/theory of change nor did it assess its adequacy.
---	---------	---

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 clearly described in the report and in annex 4 which contains a thorough evaluation matrix structured in terms of the evaluation questions. The evaluation questions, data sources, methods for data collection have all been included in the report or in the annexes. Note however that information presented in the evaluation matrix is "cut off" due to the formatting of the evaluation report.
---	-----	---

5. Are the tools for data collection described and their choice justified?	Yes	The tools for data collection have been described in the report and their choice justified. The tools are included in the annexes.
--	-----	--

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 results of the mapping of the stakeholders are in the annexes. The stakeholder consultation process has not been described in detail (e.g. the fact that there is a reference group is given and that they were involved in reviewing the draft recommendations, but no further detail was provided on this).
--	-----	---

7. Are the methods for analysis clearly described for all types of data?	Yes	Yes, the evaluation describes the methods used to analyze the different types of data in the report.
--	-----	--

8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The methodological limitations have been acknowledged. The report discusses the challenges faced during the evaluation process and described the mitigation measures taken to reduce their impact on the quality of the report.
--	-----	---

9. Is the sampling strategy described?	Yes	The sampling strategy has been described in the report
--	-----	--

10. Does the methodology enable the collection and analysis of disaggregated data?	Partial	The methodology that has been used by the evaluators does not always enable the collection and analysis of gender-disaggregated data. For example the quantitative data that were collected through secondary sources did not come in gender-disaggregated form and there was no alternative adopted by the evaluation to collect the missing information.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	On the whole, the design and methodology are appropriate for assessing gender and cross cutting issues such as human rights. This included gender-based statistics such as a fistula, and qualitative data such as an analysis of national laws. Under EQ4, the evaluation team tried to assess the effectiveness of the program activities specifically vis-à-vis youth, women and other vulnerable groups.

3. Reliability of Data	Yes No Partial	Assessment Level:	Very good
-------------------------------	----------------------	-------------------	------------------

To ensure quality of data and robust data collection processes

1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluation did triangulate data collected as appropriate. The evaluation used different data sources and triangulated/compared these when substantiating the findings.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluation clearly identified and made use of reliable qualitative and quantitative data sources. The qualitative information was collected via primary sources and the quantitative information via credible secondary sources.
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	Yes, the evaluation made explicit the possible limitations. Some challenges were noted and the evaluators indicated how they worked to reduce/mitigate the impact of the limitations on the quality of the report.
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 the evaluators collected their information with adequate sensitivity to issues of discrimination and other ethical considerations. They were particularly concerned that gender would be well-treated. Data were collected according to the group of beneficiaries and included separate discussion with youth, women, and marginalized groups.

4. Analysis and Findings	Yes No Partial	Assessment Level:	Fair
---------------------------------	----------------------	-------------------	-------------

To ensure sound analysis and credible findings

1. Are the findings substantiated by evidence?	Yes	The findings are substantiated by evidence. The evaluators noted several improvements in the coverage of SR services, reduction of child and maternal death and were able to link them to the incremental investment that was done by UNFPA. Additionally, the programme was deemed flexible and several examples have been given to show how the programme reacted to a specific situation by diverting resources or adapting the planned activities. That was the case during the outbreak of the Lassa Fever, the TOMI fire, the cholera epidemic.
2. Is the basis for interpretations carefully described?	Yes	The structure of the report goes to the level of indicators within questions and the evaluators discuss how the data lead to the finding.

3. Is the analysis presented against the evaluation questions?	Yes	The analysis has been presented against the evaluation questions. All the evaluation criteria has been addressed appropriately
4. Is the analysis transparent about the sources and quality of data?	Yes	The analysis is transparent about the sources and quality of the data. The structure of the report goes to the level of indicators within questions and the evaluators discuss how the data lead to the finding.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The evaluation made an effort to show cause-effect links between the interventions and their results. For example, the improvement in the coverage rate with regard to Family planning materials is a direct result of UNFPA's contribution (among others); the reduction that has been noted in child and maternal death has also been directly linked to the investment made by UNFPA to train health workers and provide them with better and more adequate equipment. In a number of cases, data also show the delivery of services and the connection to UNFPA assistance). No unintended outcomes were noted.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The report does show different outcomes for different target groups, as relevant. The target population for each component as well as the results have been described.
7. Is the analysis presented against contextual factors?	Yes	The analysis was presented against contextual factors. Not only did the evaluators explained the context extensively but they tried to link them to the findings in most cases.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Partial	The report does not always sufficiently elaborate on how cross cutting issues such as equality and vulnerability or human right, have been integrated. The concept of discrimination has been referred to a number of times but the analysis related to that is mainly about GBV.

5. Conclusions	Yes No Partial	Assessment Level:	Very good
<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	
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	The conclusions go beyond the findings and provide an understanding of the underlying issues being evaluated.	
3. Do the conclusions appear to convey the evaluators' unbiased judgment?	Yes	The conclusions appear to convey the evaluators' unbiased judgment, they are grounded in the findings.	

6. Recommendations	Yes No Partial	Assessment Level:	Fair
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	The recommendations flow directly from the conclusions. The link has been established for each of the recommendations.	
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 clear, their intended users are identified but their financial implications were not determined.	
3. Do recommendations appear balanced and impartial?	Yes	The recommendations appear balanced and impartial. They come directly from the findings and conclusions.	

4. Is a timeframe for implementation proposed?	No	No timeframe for implementation has been proposed in the report.
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Partial	The recommendations have been prioritised to facilitate appropriate management response and their implementation. However, a breakdown of the actions needed for their implementation has not been done.

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?		GEEW was integrated in the evaluation scope of analysis. For example, question 2 in the evaluation asks how the country office has taken into account the needs of the target populations (adolescents/youth and women) in the continuous planning and in implementation and question 4 asks about effectiveness in providing knowledge to youth and women. 2 The evaluation matrix makes references to indicators that have been designed to capture GEEW data and there is a clear gendered approach to sexual and reproductive health.
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?		Two of the evaluation questions specifically address how GEEW has been integrated into the design, planning, and implementation of UNFPA support and the results achieved, especially in sexual and reproductive health as well as issues related to violence against women, The discussion of the methodology in the text notes that it is an important element. In the intervention logic described in Annex 9, the importance of gender equality as an expected result is emphasized. 2
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?		Gender-responsive tools have been used in the analysis. For example, focus groups discussions were used as a main data source and in defining how to use focus group discussions, the evaluators note that the focus groups included the targets of the support (women from 15-49 years as well as women treated for fistulas). In addition, considerable data was collected at the Ecole des Sages-femmes and from the Association of Sages-femmes. The data collection instruments for EQs 2 and 4 make reference to the specific effects on women. 2
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		The evaluation findings, conclusions and recommendations reflect the gender analysis that was done. There are specific gender findings relating to all of the questions addressed in the evaluation. As an example in conclusions, #3 states: "L'Agence s'est démarquée par l'effectivité d'apports stratégiques et opérationnels particuliers don't l'inclusion de l'approche genre." As another example, on recommendation 9, the evaluation states: "explorer et inclure des mécanismes de durabilité dans la conception des interventions Genre et Droits Humains". 2

(*) 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).

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	22	7	71	
Overall assessment level of evaluation report			Fair	
	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 analysis was designed and was thorough, meriting a good. It was less good in the recommendations which lacked time-horizons and priorities.

Consideration of significant constraints

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

- Yes No

If yes, please explain: