

Organizational unit: UNFPA Nepal Country Office **Year of report:** 2017

Title of evaluation report: UNFPA Country Programme Evaluation: Nepal 2013-2016. Final Evaluation Report

Overall quality of report: **Very Good** **Date of assessment:** 12 September 2017

Overall comments: The report is structured according to the UNFPA guidelines. It is written in an accessible language appropriate for the intended audience. What is notable about this specific evaluation is its logical structure and references to the sources of data, however the findings are presented as a narrative, unstructured text, a minor issue. The report is too long: 97 pages, excluding the annexes. The tools for data collection are clearly explained in the section 1.3.2, which details the data analysis process generally, though not for all types of data. Methodological limitations and the sampling strategy are described. Data sources are well explained and presented in the annexes and data is well triangulated. There is evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations, for instance, interviewees were kept confidential. Cause and effect links between an intervention and its end results, as well as contextual factors are well described. The analysis elaborates on cross-cutting issues, for instance, different age groups, gender, castes, ethnicities and geographical locations. Conclusions are supported by evidence from the findings, with the evaluation often raising questions that help UNFPA to think about the underlying issues. Fifteen recommendations flow from the conclusions, but there are no references to specific conclusions. Timeframe for implementation is not clearly proposed and the recommendations are not prioritized. GEEW is included into the evaluation scope, but the evaluators do not describe any gender-responsive evaluation methods and tools, or data analysis techniques. Evaluation findings, conclusions and recommendations reflect a gender analysis: women, men, adolescent girls and boys.

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>	
I. 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 appropriate for the intended audience. What is notable about this specific evaluation is its particularly logical structure and references to the sources of data.
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)	No	The report is 97 pages, excluding the annexes. The evaluation is about 27 pages too long.
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 according to the UNFPA guidelines: there is a clear distinction made between Introduction, Country Context, UNFPA Programmatic Response, Findings, Conclusions, and Recommendations. The minor issue is that lessons learned are not clearly highlighted from the other findings.
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?	Yes	The annexes are complete.
<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 written as a stand-alone section. 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)?	Partial	Executive Summary follows the required structure. Purpose is not clearly written, but intended audience, objectives, brief description of intervention, methodology, conclusions and recommendations are presented as separate paragraphs. Programmatic recommendations are not mentioned in the executive summary. In addition, the evaluators briefly describe the main findings (pp. 14-16).
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is 5 pages, consistent with UNFPA norms.

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 specifies that "The intended primary users of the evaluation are UNFPA (at country, regional and global level) and the UNFPA Executive Board. Additionally, counterparts within the Government of Nepal (GoN) and other development partners are expected to benefit from evaluation results."
---	-----	---

2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	Chapter 2 "Country Context" explains development challenges and national responses, political and socio-economic context. Constraints are explained for all three program components (SRH, Population and development, GEEV): achievements, challenges, and government policies.
---	-----	--

3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Yes	There is a thorough discussion of the intervention logic for each mandate area of the country programme, although the report does not display an overall intervention logic diagram.
---	-----	--

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 text and in the evaluation matrix. The evaluation questions are based on the standard four OECD-DAC criteria and two additional criteria specific to UNFPA (p. 19). The evaluation questions are clear and correspond to the two components of the evaluation: 1) analysis of programmatic areas, and 2) analysis of strategic positioning. The table 2 describes relationship between EQs and evaluation criteria. The evaluation matrix is found in the annex 2. It is designed in accordance with the UNFPA Handbook for evaluation at UNFPA (pp. 109-110). The matrix has six columns: evaluation criteria, evaluation questions; assumptions to be assessed; indicators, sources of information, and data collection methods (Annex 2).
---	-----	--

5. Are the tools for data collection described and their choice justified?	Yes	The tools for data collection are clearly explained in the section 1.3.2. "Methods for data collection and analysis." The evaluators used different methods for data collection: documents review, semi-structures interviews, interviews and focus group discussions, and direct observation of targeted institutions and areas (p. 21).
--	-----	---

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	Annex 6 "Stakeholder Mapping" includes Stakeholder Map by Programme Components, SP Outcomes, CP Outcomes, and CP Outputs. Section 1.3.3 "Stakeholder and district sampling" provides information about persons consulted. Evaluation reference group was formed to ensure overall quality of the evaluation process (p. 23). It is unclear from the chapter 6 "Recommendations" and methodology section how the consultants discussed draft recommendations with the key stakeholders.
--	-----	--

7. Are the methods for analysis clearly described for all types of data?	Yes	The section 1.3.2. explains data analysis process generally, but not for all types of data: triangulation of information, document analysis and “direct observation along the lines of the assumptions for assessment and indicators” (p. 21).
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 in the section 1.3.4: time restrictions, limited coverage of earthquake response, absence of one of the evaluators, stakeholder unavailability, data gaps, and language constrains (pp. 22-23).
9. Is the sampling strategy described?	Partial	The sampling strategy is described in the section 1.3.3 “Stakeholder and district sampling.” The consultants refer to an illustrative sample rather than a statistically representative sample (p. 21). The evaluators cite the UNFPA handbook that states “the evaluators should not aim at obtaining a statistically representative sample, but rather an illustrative sample”. Based on this, they see a purposive sample and describe some of the factors used in selecting stakeholders for interviews and the selection of 3-6 districts to visit. However, the specific places visited and why they were more illustrative than others is not clear. They clearly explain criteria for selecting the districts for the interviews (p. 22) and a number of districts for the evaluation, but they do not specify how they chose the optimal size for individual and group interviews (pp. 21 - 22).
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The methodology clearly explains how the consultants disaggregated data: “To the extent possible, available and relevant, data were considered disaggregated at the level of gender, age and districts” (p. 21).
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The consultants covered wide range of stakeholders (UNFPA staff and implementing partners). The table 4 represents number of meetings along key topics (SRH, ASRH, GE, PD, and Emergency). Therefore, the design and methodology is appropriate for assessing the cross-cutting issues.

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 consultants triangulated data well. They explain this in the methodology (p. 21). It is also possible to find examples of the data triangulation in the text: “Non-representative interviews and FGDs, including with FCHVs, revealed that, overall” (p. 45), “Some indication of a change in attitudes can be gauged from 2015 pre- and post-training data: Interviews and discussions conducted gave the sense that men and boys who had benefited from UNFPA...were indeed more sensitive...” (p. 63).
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Data sources are well explained and presented in the Annex 3 “Documents consulted,” the Annex 5 “Interview Guides,” the Annex 7 “Persons consulted,” the Annex 8 “National policies and surveys” and the Annex 10 “Monitoring Data for CP Outcomes and CP Outputs.” Also, the evaluators comment on the quality of data in the text, for instance, they discuss that “The evaluation team was therefore unable to use the RRF to assess progress and UNFPA’s contribution” (p. 65).

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 in the section 1.3.4 and the mitigating measures for each are described.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	The evaluators explain that “Team members closely adhered to the UNEG Ethical Guidelines for Evaluation and the UN Code” (p. 21). While, the consultants do not provide more details, practices such as ensuring the confidentiality of interviews suggest that the Guidelines were followed.

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

To ensure sound analysis and credible findings

1. Are the findings substantiated by evidence?	Yes	That the findings are based on evidence is particularly noticeable from the text of this evaluation. The evaluators were objective in presenting and analyzing information and the sources of data for findings, like interviews or documents, are often mentioned. They were clearly based on the evaluation matrix.
2. Is the basis for interpretations carefully described?	Yes	In each finding, the evaluators thoroughly documented the basis for their interpretation, noting, for example, that in some cases no surveys that would show stakeholder perceptions were available and the finding is therefore contingent, but then interviews and observations in the field were used to suggest trends. There are also many examples of the basis for interpretations is carefully described throughout the analysis, for instance, the evaluators explain that “Data is available by sex and by disadvantaged groups: they shown that youth were less represented by young women than by young men” (p. 74).
3. Is the analysis presented against the evaluation questions?	Yes	Structure of the analysis chapter is logical. The analysis is presented according to the evaluation questions which correspond to the evaluation criteria. Within each evaluation question, assumptions are assessed based on the indicators, for instance, the section 4.2 “Effectiveness and sustainability” includes EQ 3, Assumption 3.1., and indicators 3.1.1.-3.1.18.
4. Is the analysis transparent about the sources and quality of data?	Yes	The consultants discuss the sources and quality of data throughout the analysis, for instance, they explain that “No assessments or perception surveys are available to document the satisfaction of such community mobilizers supported by UNFPA with trainings and sensitization workshops. Hence it is difficult to gauge opinions...” (p. 62).
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Section 4.2 “Effectiveness and Sustainability” explains cause and effect links between an intervention and its end results that is evident from the text of the chapter: UNFPA technical support for FHD influenced FP strategy 2020 (p. 44), trainings in Sindhuli “on insertion of IUCD and implant was considered particularly helpful for providing emergency FP services after the earthquake...” (p. 46), and other examples. Although, unintended outcomes are not highlighted in a way to easily find them. The report has well designed tables with data on outcomes, for instance, the tables 9 - 13. Explanations are clear and logical.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis shows different outcomes for different target groups: adolescent girls and youth (p. 42), women of different age (p. 46), educational sector (p. 53), and others.

7. Is the analysis presented against contextual factors?	Yes	The analysis is presented against contextual factors. Country context is well explained in the chapter 2. Also, contextual factors are described in other parts of the report, for instance, the consultant consider socio-cultural context (p. 77), Nepal earthquake (p. 80), and others.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The analysis elaborates on cross-cutting issues, for instance, the consultants discuss that “Disparities persist among different age groups, gender, castes, ethnicities and geographical locations. Issues related to discrimination, impunity, GBV and exclusion prevent the realization of the rights and potential of women” (p. 28).

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	There are strategic and programmatic conclusions in the report. The consultants present conclusions clearly and briefly (1 paragraph for one conclusion). Conclusions are supported by evidences from the findings, for instance, the consultants discuss that long-awaited reforms will influence UNFPA policy in the country; data gaps exist because of lack of available up-to-date and adequately disaggregated data (p. 93).
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	While Chapter 5 is entitled "conclusions and recommendations" in fact it only includes conclusions, with recommendations having a separate chapter. The consultants often raise questions to the UNFPA to think about the underlying issues like “The UNFPA CO has gone to great lengths to have a robust and workable RRF. However, an important challenge has been identified concerning the selection of and level at which to pitch RRF outcome indicators” (p. 94).
3. Do the conclusions appear to convey the evaluators’ unbiased judgement?	Yes	The evaluators succeeded in explanation of their conclusions and providing evidences to support the arguments. For instance, they discuss that “Besides reforming the policy framework and strengthening state systems and their capacities, UNFPA has rightly worked in different ways to promote social and behavior change and better utilization of services among women ... good experience has been gained through...” (p. 105).

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

To ensure the usefulness and clarity of recommendations

1. Do recommendations flow logically from conclusions?	Partial	Fifteen recommendations are divided into two categories: Strategic and Programmatic. The consultants state that the recommendations flow from the conclusions, but there are no 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 logical and clearly written. The intended user is the UNFPA Nepal office. Recommendations lack information on human, financial and technical implications.

3. Do recommendations appear balanced and impartial?	Yes	The recommendations look balanced and impartial. The consultants explain the reasons why they advise things to be done a certain way, for instance, they state that “In view of scarce resources, the CO ought to concentrate its Regular Resources on its original mandate where it has an added value” (p. 97).
4. Is a timeframe for implementation proposed?	Partial	Timeframe for implementation is not clearly proposed, but could be implied from the context (e.g. “during CP8” – p. 97).
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Partial	Recommendations are clearly presented. However, they are not prioritized, which does not facilitate 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 is included into the evaluation scope. The country program evaluation assessed gender equality program component (p. 18). Evaluation matrix has GEEW indicators, for instance: “Evidence that UNFPA consulted/surveyed/included (representatives of) marginalized/vulnerable women throughout programme/project design and implementation in all areas of work” (Annex 2).
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 are based on OECD-DAC criteria: relevance, effectiveness, efficiency and sustainability. Two criteria specific to UNFPA do not address GEEW (coordination, added value). The evaluation questions 3-5 include GEEW dimensions: vulnerable/marginalized women, GBV (pp. 19-20).
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	Evaluation consultants explain that “Team members closely adhered to the UNEG Ethical Guidelines for Evaluation and the UN Code (p. 21). For instance, all interviews were confidential. But, they do not provide details. Interview guides in the annex 5 do not include information about the UNEG guidelines. The methodology does not provide specific approaches for fostering participation and inclusiveness (recommended by the table 6.7. in the in the UNEG Guidance Integrating Human Rights and Gender Equality in Evaluations). The evaluators do not describe any gender-responsive evaluation methods and tools, and data analysis techniques. Desk review and data analysis stages included gender analysis (Annex 3: Documents Consulted). The evaluation consultants collected gender-related data with regards to the evaluation questions 3-5.
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	3	Evaluation findings reflect a gender analysis. The evaluators collected and analyzed gender-related data according to the evaluation indicators. For instance, the indicator 3.1.16. requires assessment of the “percentage of women aged 15 to 24 with unmet need for family planning...” (p. 49). The evaluators revealed that percentage increased to 43.3%. Evaluation conclusions 8 and 10 reflect a gender analysis: women, men, adolescent girls and boys (pp. 94-95). Evaluation recommendation 15 includes gender and GBV issues (p. 97).

(*) 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).

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	7	31	
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 main reason for the very good rating is that the analysis underpinning the findings is unusually good. This was despite the fact that some areas of the country were not available during field work.

Consideration of significant constraints

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

Yes No

If yes, please explain: