

Organizational unit: UNFPA Pakistan Country Office **Year of report:** 2016

Title of evaluation report: UNFPA COUNTRY PROGRAMME EVALUATION: Pakistan (2013-2017)

Overall quality of report: **Good** **Date of assessment:** 7 September 2017

Overall comments: The report is structured according to the UNFPA guidelines and is easy to read. It has a somewhat complex structure with many sub-sections that make the report less focused on the evaluation questions and is about 120 pages. The evaluation framework is designed in accordance with the UNFPA requirements. The evaluation triangulated data collected as appropriate. In most cases the findings are supported by evidence, although the structure of the findings chapter is complex and it is unclear how sub-chapters relate to the evaluation questions. The consultants presented four strategic and two programmatic conclusions that flow clearly from the findings. Conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme. Recommendations flow from the conclusions, are clearly written but while some recommendations include financial and technical implications, others do not. Some recommendations have a timeframe but others general.

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		Insert <u>assessment level</u> followed by main <u>comments</u> . (use 'shading' function to give cells corresponding colour)	
I. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Fair
To ensure the report is comprehensive and user-friendly			
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?	Partial	The report is easy to read. In the meantime, it has a bit of a complex structure a lot of sub-sections which make the report not focused on the key ideas. It is mostly written well, but has punctuation and grammatical errors, which, at times, makes the language hard to follow.	
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 about 120 pages in total (96 pages without the annexes and 28 pages longer than the maximum length). The report could have been edited to the recommended size.	
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 UNFPA guidelines, lessons learned are mentioned clearly after a summary in each relevant section of the analysis.	
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 they do not contain 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 meets UNFPA requirements. The activity map on the very first page is a good primer for the Executive Summary demonstrating clearly that UNFPA is assisting through seven key strategies: FP Vouchers, Newly-wed counseling, fistula training centers, emergency response for IDPs, emergency preparedness, training and policy advocacy.	
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	There is a clear structure to the executive summary which includes all of the required elements.	

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

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

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 is clearly detailed in section 1.2 Scope and Audience of the Evaluation.
---	-----	--

2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The Table 1 presents Facts and Figures. The Chapter 3 explains the country context of reproductive health and population. Constraints are explained in the Section 1.3.6. "Limitations and Mitigation strategies" (p. 24).
--	-----	--

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 intervention logic is included in graphic form and described in section 3.2.2. The adequacy of the intervention logic is referenced within the evaluation findings (section 4.1.2).
--	-----	---

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	<p>The evaluation framework is designed in accordance with the UNFPA requirements: the evaluation questions are based on standard four OECD-DAC criteria and an additional criterion (Coordination/Strategic Positioning) was added to assess UNFPA's strategic positioning in Pakistan (p. 20).</p> <p>The evaluation questions are clear and correspond directly to the evaluation criteria. For instance, EQ 1 relates to the relevance, whereas EQ 2 addresses the responsiveness.</p> <p>The evaluation matrix is found in the annex 4, it is designed in accordance with the UNFPA Handbook for evaluation at UNFPA (pp. 109-110).</p>
--	-----	---

5. Are the tools for data collection described and their choice justified?	Yes	The tools for data collection include a desk review, semi-structured interviews of 112 key informants, eight focus group discussions and five site visits and observations. The rationale for selecting these methods is described, as well as their sampling strategy. The evaluators clearly explained their methodological choices, for instance, they justified that the focus group discussions were "a quick and effective approach to gathering information from a large number of programme beneficiaries..." (pp. 21-22).
--	-----	--

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	Section 1.3.5 describes stakeholder mapping during the Desk Review process (p. 23), although there is no formal map. Analysis of the stakeholders in the section 2.2.8. describes the three broad categories of stakeholders: government (public sector), NGOs including international INGOs (private sector implementers), and donors (p. 34). Table 4 presents stakeholders selected for interviews and focus group discussions. Stakeholder consultation process is described in the section 1.3.4. "Validity of Data" (p. 23).
--	-----	--

7. Are the methods for analysis clearly described for all types of data?	Yes	The methods for analysis are explained in the section 1.3.1. "Phase 3: Analysis and Synthesis Phase" and in the text, for instance, the consultants used a content analysis (p. 23), Programmatic and Strategic Positioning analyses (the table 3 "Association between Evaluation Questions and Analysis"). Analysis was primarily conducted in consultation with stakeholders and within the Evaluation Team.
--	-----	--

8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	Section 1.3.6 Limitations and Mitigation Strategies clearly details evaluation constraints and the evaluation's team strategy for addressing them. There are some constraints that beg for more detail; for example, how did the evaluation team assess key informant response bias which was described as "evidently biased in either direction (too positive or too negative)" in some cases.
9. Is the sampling strategy described?	Yes	Section 1.3.5. establishes sampling criteria and describes the sampling strategy "The final sample size for key informant interviews/FGDs/site visits included stakeholders and beneficiaries from all 4 components of the CP 8 programme activities, geographic representation of provinces, thematic areas and level of engagement. Field visits were selected on the basis of convenience, 1-2 sites for each component, and accessibility by air or ground transport and 5 sites were visited. A total 112 interviews and 8FGDs were conducted" (p. 23). In the meantime, the consultants do not provide a detailed justification for their sampling strategy.
10. Does the methodology enable the collection and analysis of disaggregated data?	Partial	The methodology enables the collection and analysis of disaggregated data. For example, the consultants collected data disaggregated by age (p. 30), and types of contraception (p. 32). However, there was a noted lack of availability of disaggregated data on a national level that affected the findings somewhat and was dealt with by the data that the consultants were able to acquire.
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, for instance, the evaluators mention "gender equality approach in reaching marginalised populations... The tool also asked on how targeting was done for gender and vulnerable populations in reaching populations and delivery of programming" (p. 22). Section on limitations discusses "Absence of Objective Baseline and Tracking Data Including Disaggregated Data" (p. 25).

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	The evaluation process and report clearly applied triangulation techniques and this was mentioned as a key strategy in ensuring data were valid (section 1.3.4 Validity of Data). The evaluation triangulated data collected as appropriate: primary data (semi-structured interviews/focus group discussions) were compared with desk review and analysis of secondary data (programmatic reports, M&E reports), and other independent data (p. 13).	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The sources of quantitative and qualitative data were clearly described in the report. Based on the description that can be considered reliable.	
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	This was included in section 1.3.6 Limitations and Mitigation Strategies. Data was not consistently disaggregated by gender in the context section (unless the data points were gender-specific, for example maternal deaths), however this was also mentioned as a key constraint in the available data from secondary sources.	

4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	Cross-cutting issues such as gender and marginalized populations are referenced. Section 1.3.7. "Ethical Considerations" explains how the evaluators ensured that data has been collected with a sensitivity to issues of discrimination and other ethical considerations (p. 26). Annexes 2 and 5 confirm that the responses were kept confidential.
--	-----	---

4. Analysis and Findings	Yes No Partial	Assessment Level:	Good
<i>To ensure sound analysis and credible findings</i>			
1. Are the findings substantiated by evidence?	Yes	The authors do well to substantiate their findings with evidence across data sources and collection methodologies. In most cases, the findings are supported by evidence: the consultants refer to the sources of data like interviews (references to the interviews) or documents review (references to the documents). But, in some cases the evaluators make general statements or make statements without a reference to a source of data, for instance, section "Common Observations and Findings" does not have any references on sources of data (p. 63). The text is a narrative description.	
2. Is the basis for interpretations carefully described?	Yes	Interpretations are carefully described throughout the analysis. For instance, the evaluators discuss that "Training and quality supervision of health care providers led to capacity building in public and private sector. However, despite being set up as a 3 arm trial, no comparisons are available between results of the three prongs of the project" (p. 66). Another example shows that the consultants analyze obtained data "Given the very low penetration of FP (15% of MWRA avail FP services in a given year) and MH (52%) services among the population suggests the need for engaging communities and particularly with local leaders..." (p. 79).	
3. Is the analysis presented against the evaluation questions?	Partial	The Structure of the Findings chapter is somewhat complex: there are the evaluation questions which correspond to the evaluation criteria. But, the consultants just mention them in the beginning of the chapter and create a lot of sub-chapters to present their findings. From the first look, it is unclear how these sub-chapters relate to the evaluation questions. The evaluation questions associated with the evaluation findings are listed at the start of their relevant section but sometimes the discussion is unrelated to the evaluation question.	
4. Is the analysis transparent about the sources and quality of data?	Yes	The analysis is transparent about the sources and quality of data. The evaluators paid attention to the quality of data, for instance, they explain that "A key issue has been that data from these surveys is not consistent" (p. 35). They also followed precise data validation process described in the sections 1.3.4. and 1.3.6.	

5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	<p>Section 4.2 “Effectiveness and Sustainability” has three evaluation questions to be covered by the evaluation. Unintended outcomes are mentioned in the text, for instance, they evaluators explain that “The unintended consequence of UNFPA CP8 support for extensive capacity building...” (p. 60). But, unintended outcomes are not specifically highlighted to easily find them (there are no tables or special paragraphs, for instance).</p> <p>The evaluation consultants explored causal links between outputs, outcomes, and impacts which is evident from the Figure 10 “Effects Diagram” and from the analysis. For instance, the consultants explain that “The “Pakistan Vision 2025” endorsed the linkage between population, development and population dynamics. Other examples of UNFPA support ... has resulted in various policy dialogues on population and development linkages, reproductive health and rights and youth. These contributed to supporting actions for ICPD beyond 2014 and post MDGs advocacy (2015); and provided evidence base for provincial policy documents ...” (p. 74). In the meantime, the structure of the Findings chapter (Summary, Key Results, Common Observations, and Findings) focuses on the presentation of general results rather than casual explanations: The evaluation report can benefit from more structural representation of casual explanations of Outputs-Outcomes-Impact pathways (for instance, tables).</p>
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	<p>The evaluators performed stakeholder mapping exercise and selected “stakeholders and beneficiaries from all four components of the CP 8 programme activities, geographic representation of provinces, thematic areas and level of engagement” (p. 23).</p> <p>The analysis shows different outcomes for different target groups. For instance, the consultants reveal the following program beneficiaries in the Effectiveness section: “governmental departments”, “public sector female and frontline service providers,” and “public sector institutions” (p. 60).</p>
7. Is the analysis presented against contextual factors?	Yes	The analysis is presented against contextual factors throughout.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	<p>Gender and youth topics are thoroughly covered in the analysis. For instance, the Population Development component contributed “technical and financial support to ... building capacity of public departments on ... gender equality” (p. 73). Vulnerability issues are addressed in the section Poverty and its Gendered Context (2.1.2). Human rights are discussed in the Law and Security section (2.1.5).</p>

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	<p>Conclusions are divided into two categories. The consultants presented four strategic and two programmatic conclusions.</p> <p>Conclusions flow clearly from the findings which is evident from justification provided to support each conclusion.</p>	
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	<p>Conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme, for instance, the evaluators explain that “What undermined the Country Programme from achieving its full potential of institutional and practice changes and outcomes ... can be directly attributed to limitations in the design and implementation approach” (p. 86); “While successful in some aspects, lack of measurement of results often meant that only token/ nominal endorsement by religious leaders or parliamentarians were obtained and construed as substantial successes” (p. 88).</p>	

3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	The conclusions were well-founded, though strongly worded, and often were drafted as though they were recommendations about what UNFPA 'should/could' or 'should not' do rather than as conclusions.
--	-----	--

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

To ensure the usefulness and clarity of recommendations

1. Do recommendations flow logically from conclusions?	Yes	Recommendations are divided into two categories: there are 6 Strategic and 2 Programmatic recommendations. They all flow from the conclusions with references to specific conclusions, except programmatic recommendations (there are no direct references to conclusions).
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations are clearly written. The consultants specify the intended users which are the UNFPA Country office, Government or Implementing partners, UN agencies, government, private sector, and others. Recommendations are action-oriented (they all have action points). Some recommendations include financial and technical implications, for instance, the consultants advise the UNFPA to "support facilitation of the private sector in production of cheaper local products (i.e. contraceptives, commodities), technical assistance for quality and standardisation," "develop a detailed strategy for maximizing programmatic inputs (technical, financial, logistics, and human) to deliver sustainable outcome level results" (p. 92). But, some recommendations are general like "UNFPA CO should establish mechanisms for formal engagement with provinces including inputs from the districts for defining the final shape of CP9 design and interventions" (p. 90).
3. Do recommendations appear balanced and impartial?	Yes	The recommendations look balanced and impartial. The consultants provide arguments to support their statements, for instance, they mention that "One limitation that was commonly identified during implementation of the MDGs and is often a recurring theme" (p. 91). But, in some cases the consultants are more general: "UNFPA should identify how gender and rights based equitable targeting will be done" (p. 91).
4. Is a timeframe for implementation proposed?	Partial	While some recommendations have a timeframe or it is implied in the text (e.g. "This is the right time for UNFPA to assist the government" – p. 91), some are general (e.g. "UNFPA should explore how to find cost-effective ways of engaging private sector partners" – p. 92).
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	Recommendations are prioritized (Moderate to High). Some Recommendations are clearly presented to facilitate appropriate management response, but it is unclear which recommendations are realistic to implement with available resources in the nearest future. All recommendations are High or Moderate to High.

7. Gender	0 1 2 3	Assessment Level:	Very 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?	2	The evaluation covered all four program components: i) Policy Advocacy, ii) Youth/Adolescent Sexual and Reproductive Health, iii) Family Planning and Maternal Health in both development and humanitarian settings, and iv) Population and Development (p. 18). GEEW was not included into the evaluation scope directly, but gender was assessed as a cross-cutting issue (EQ 5 and EQ 6). Evaluation matrix has GEEW indicators, for instance: • “No. of government staff trained in RH, PD, and gender issues” (p. 104). • “No. of national or provincial plans or documents with emphasis on RH, PD and gender issues during the CP8 period” (p. 105).
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?	3	The evaluation questions addressed how GEEW was integrated into the programme. Specifically, the questions addressing GEEW included: (EQ1) To what extent are the objectives of the UNFPA 8th CP adapted to the needs of the population (including vulnerable and marginalised groups)...; and (EQ 6) To what extent were the principles of equitable access, rights-based approach and gender-responsiveness integrated in UNFPA 8th country programme and its interventions/activities?
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	3	The tools specifically request the gender of the respondent and the questions are phrased openly so as to gather responses that may differ by gender. For example, a question in the focus group protocol refers to barriers to service access. Desk review and data analysis stages included gender analysis as gender has been mainstreamed into the program components.
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	3	Evaluation findings reflect a gender analysis. For instance, there are such sections as 2.2.3. “Nutrition of Women and Children” (p. 32), and 2.2.6. “Rights of Women and Youth” (p. 33). Evaluation conclusion 4 (EQ 1, 2, 8 and 9) reflect a gender analysis. Evaluation recommendations 2, 3, and 6, reflect a gender analysis.

(*) 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			
Total scoring points	29	64	7	
Overall assessment level of evaluation report		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

Consideration of significant constraints

The quality of this evaluation report has been hampered by exceptionally difficult circumstances: ☐ Yes ☒ No

If yes, please explain: