

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

Title of evaluation report: The UNFPA 3rd Country Program 2012-2017 The Kyrgyz Republic

Overall quality of report: **Good** **Date of assessment:** 26 September 2016

Overall comments: The evaluation report is structured around UNFPA standards. The report is comprehensive but some sections exceed the recommended length. The evaluation design and methodology are clearly outlined. Triangulation was applied throughout the evaluation, qualitative and quantitative data sources are explained and data is gender disaggregated by gender. Data limitations are described and respective mitigation measures are developed and implemented. Data is analyzed and interpreted systematically but in some cases the analysis does not reference the data source. Conclusions and recommendations are presented together in a well-structured and succinct manner. However, conclusions are not systematically related to findings and they could be formulated more specifically. Recommendations are prioritized, targeted but not always operationally-feasible. The evaluation methodology, methods, and data analysis techniques are not specifically gender-responsive. However, gender equality and the empowerment of women is included in the evaluation scope because it is an area of the country programme.

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 <i>assessment level</i> followed by main <i>comments</i> . (use 'shading' function to give cells corresponding colour)	
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	Fair
<i>To ensure the report is comprehensive and user-friendly</i>		Comment: The report includes all of the required components for the body and annexes with the exception of a section on Lessons Learned (which was one of 4 objectives of the study). The report is comprehensive and user-friendly, but is somewhat long. The main part of the report is 84 pages, the overall length is 239 pages. There are 13 tables, 18 figures, and 9 annexes. The report is structured in a logical way: Acronyms; Executive Summary; Introduction; Methodology including Approach and Limitations; Context; Findings/Analysis. However, Conclusions and Recommendations are presented in one chapter. There is a clear distinction made between analysis/findings, conclusions, recommendations.	
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	There are no annexes such as Methodological tools used and List of interviewees. But there are annexes "Site Visit Schedule and Stakeholder Listing" and "Training Listings." There is a column in the Evaluation Matrix (Annex 2) called "Data Collection Methods." However, some sections are overly lengthy (the Executive Summary is 7 pages and the Country Context is 12 pages), and there are some densely worded paragraphs that exceed 1/2 page. The heavy use of acronyms also detracts from the reports readability (the list of acronyms is 2.5 pages although some do not appear in the body of the report, i.e. EA for Evaluation Assistant and HII for Health Insurance Institute).	
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 Executive Summary provides details of the content of the report that serves as a stand-alone section. However, it is densely written and at 7 pages, it is double the recommended length. Fewer details (particularly in description of program and findings), the use of sub-section headers and bullet points, and shorter paragraphs would increase readability and utility.	
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		
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		
<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		
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		
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	No		

2. Design and Methodology	Yes No Partial	Assessment Level:	Good
<i>To ensure that the evaluation is put within its context</i>		Comment: The description of the methodology is clearly stated: desk review, visits to programme sites in 4 regions, and interviews with a range of stakeholders including program beneficiaries (using purposive and non-random sampling).	
1. Does the evaluation describe the target audience for the evaluation?	Yes	The report describes that the evaluation is for accountability and learning purposes. The target audience for the evaluation is "UNFPA Kyrgyzstan country office, the Government agencies, national institutions and local NGO." The development and institutional context of the evaluation is clearly described in the chapter Chapter 2 "Country Context." The evaluation report presents the intervention logic in the Figure 7 "Simplified Logic Model for UNFPA Kyrgyzstan 2014 Aligned CP Framework," but the model has very small font size to read the text easily. The explanation above the Figure 7 does not provide the detailed discussion around theory of change and SMART-ness of the outputs and outcomes. Design and Methodology are clearly explained in the section 1.3. "Methodology and Process." The chapter provides clear and detailed explanation of the evaluation approach and framework (Desk review, Site visits, Semi-structured group and individual interviews with stakeholders, Follow-up interviews with trainees, Focus group discussions and client exit interviews with stakeholders and client/beneficiaries). The methods chosen are appropriate for addressing the evaluation questions. The evaluation explains constraints and limitations in the section 1.3.3. "Availability assessment, limitations and risks." The evaluation questions, assumptions, indicators, techniques and Tools for data collection are provided in a detailed manner in the Annex 2 "UNFPA Kyrgyzstan CPE Design Report Evaluation Matrix (Draft 0.6) 25 July 2016". Details of participatory stakeholders' consultation process are provided in the section "Stakeholder Involvement" (p. 2). The evaluation consultants worked closely with the Evaluation Reference Group. ERG ensured stakeholders' participation in the design and conduct of the evaluation. There is a comprehensive and credible stakeholder map in the Annex 7. The evaluation questions incorporate cross-cutting issues as the program focused on Reproductive Health and Rights, Youth, Gender Equality and Population and Development. The design and methodology do not include specific instruments or approaches to assess the cross-cutting issues, but the evaluators tried to address these issues during the evaluation. For instance, the consultants mention the gender of the respondents in their methodology "All these respondents were female" (p. 4). The study relied heavily on qualitative data (as proposed in the ToR) but no information is given on how the analysis was undertaken other than "the analysis is based on triangulation of information obtained from various stakeholders' views, as well as with secondary data and documentation reviewed by the team" (p. 2). If a some sort of database or coding system was used to organize the data, it was not explained. Evaluators noted that the TOR did not have explicitly stated objectives to address gender, human rights and vulnerability. The evaluators did take steps to incorporate perspectives of women and men, they presented gender disaggregated data where possible, and used a gender continuum framework to assess the extent to which gender was incorporated within each program area. The stated it was beyond the scope of the assignment to elaborate on human rights and vulnerability	
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes		
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Yes		
<i>To ensure a rigorous design and methodology</i>			
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		
5. Are the tools for data collection described and their choice justified?	Yes		
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		
7. Are the methods for analysis clearly described for all types of data?	Yes		
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes		
9. Is the sampling strategy described?	Yes		
10. Does the methodology enable the collection and analysis of disaggregated data?	Partial		
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes		

3. Reliability of Data	Yes No Partial	Assessment Level:	Very good
To ensure quality of data and robust data collection processes		Comment: Triangulation was applied throughout the evaluation that is evident from the text: "A "triangulation" of the two approaches helps validate the findings" (p. 119), "The analysis is based on a synthesis and triangulation" (p. iv).	
1. Did the evaluation triangulate data collected as appropriate?	Yes	Sources of qualitative and quantitative data are explained in the methodology: desk review, site visits, semi structured group and individual interviews, group and individual follow-up interviews with former trainees in UNFPA-supported training events, focus group discussions and exit interviews with stakeholders and client/beneficiaries. The report includes references, 13 tables and 18 figures to present the analysis and support the findings. The Annex 2 has well described sources of data collection.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Credibility of data is ensured by References chapter in the end of the report; Names of the documents and dates such as "National Action Plan for Gender Equality (2015-2017);" Examples from the interviews "There were several concrete examples that clearly demonstrated how stakeholders were	
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	ready to carry out project activities" (p. 60), related annexes such as Annex 3 "Site Visit Schedule and Stakeholder Listing," Annex 6 "Copies of data collection instruments," Annex 7 "Map of Program Activities by Area," Annex 8 "Training Listings," and Annex 9 "Description of Gender Continuum."	
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	Limitations are explained in the section "1.3.3. Availability assessment, limitations and risks." Mitigation measures are developed and implemented. Details of participatory stakeholders' consultation process are provided in the section "Stakeholder Involvement" (p. 2). The report says that the evaluation was conducted "in accordance with the UNEG's Ethical Guidelines for Evaluation, at www.unevaluation.org/ethicalguidelines ." Also, the evaluators discuss how carefully they prepared and conducted the interviews "The SAQ was developed with a consistent set of precautions for informed consent and confidentiality with questions" (p. 3).	

4. Analysis and Findings		Yes No Partial	Assessment Level: Good
<i>To ensure sound analysis and credible findings</i>			Comment: The findings are structured according to the evaluation questions, and each of the four program areas is assessed according to the standard UNEG criteria. There is a table for each program that shows outcomes, related outputs, indicators, baselines, targets, and yearly achievements for those indicators. It is also helpful that a summary of findings is provided for each program and criteria, however inconsistent formatting for sub-sections creates some confusion. The information is analysed and interpreted systematically, but not very logically. Sometimes paragraphs are too long (up to 25 lines). It is not easy to follow the logic behind the discussion in a such long paragraph (p. 47, p. 49, p. 62). Assumptions are presented clear in the Annex 2. The analysis is presented against the evaluation questions, which are mentioned in the beginning of each section. But, often there is no clear connection between the evaluation question and the text. In some cases, the analysis does not have references on the sources of data. For instance, there are too general references in this case "Based on review of financial documents, stakeholder interviews, and a review of Annual Work Plans, the UNFPA Gender and Gender Based Violence Area has made excellent use of its human, financial and technical resources" (p. 58). There is no specific source of data in this statement "Interviews with stakeholder and analysis of documents demonstrated that UNFPA CO used a variety of innovative approaches (p. 59)." It is not possible to easily check these statements.
1. Are the findings substantiated by evidence?	Yes		
2. Is the basis for interpretations carefully described?	Yes		
3. Is the analysis presented against the evaluation questions?	Yes		
4. Is the analysis transparent about the sources and quality of data?	Yes		
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes		
6. Does the analysis show different outcomes for different target groups, as relevant?	Partial		
7. Is the analysis presented against contextual factors?	Yes		

8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes		
5. Conclusions	Yes No Partial	Assessment Level:	Good
<i>To assess the validity of conclusions</i>			
1. Do the conclusions flow clearly from the findings?	Yes	<p>Comment: Conclusions and recommendations are combined in one chapter and are subdivided into (a) strategic conclusions and recommendations, and (b) program conclusions and recommendations. The section is highly structured and succinctly presented; conclusions are organized by criteria and program area. There are no direct links and cross-references between Conclusions and Findings. In some cases the evaluators provide clear examples to support their arguments. For instance, the following conclusion could be considered as specific "Conclusion 3... For example, to have a target for the total number of staff to be trained to provide FP services, there should be an estimate of the number of staff currently employed within the MoH who provide FP services." Conclusion 5 sounds specific as well "Given the current context of a growing trend toward religious conservatism, UNFPA KR's long-term experience in working collaboratively with religious leaders on RH, Gender and GBV and youth will become increasingly important." Some conclusions are formulated as broad statements, they are not very SMART, for instance, "A nation-wide multisectoral approach to family planning is urgently needed to reverse negative trends in contraception use...." Such words as "multisectoral approach" and "negative trends" could be more specific. In the Conclusion 2 it is not explained how the evaluators came to such conclusion that UNFPA KR has "has a highly developed effective policy advocacy capacity." It would be appropriate in this case to clarify the "capacity" (for example, enough money or professional staff, etc.). Also, it is unclear what "A wide range" means in the statement "A wide range of stakeholders is the best insurance for successful advocacy."</p> <p>The conclusions demonstrate some level of analytical abstraction as the evaluators made their generalizations based on interdependences and relations between processes, but not provide the references to the findings to support their arguments. For instance, Conclusion 6 states that "UNFPA KR has developed strategies and policies on the basis of high quality assessments and nationally representative data, which have contributed to their relevance and effectiveness." There is no reference to the text to support the argument that UNFPA assessments and data were of high quality. It is unclear how these assessments and data contributed to relevance and effectiveness.</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?	Partial		
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes		

6. Recommendations	Yes No Partial	Assessment Level:	Good
<i>To ensure the usefulness and clarity of recommendations</i>		Comment: Recommendations flow logically from conclusions (there is a link to a specific conclusion in the title of each recommendation).	
1. Do recommendations flow logically from conclusions?	Yes	The recommendations are sufficiently clear and targeted at the intended users. Some recommendations are operationally-feasible like "...to finalize a plan for an uninterrupted supply for specific methods, with or without GKR funds, as soon as possible and not later than December 2016" and "UNFPA RH should proactively assess the feasibility of replicating the next KR DHS by 2022 in the absence of USAID support." However, many recommendations are not SMART enough to be operationally-feasible. Such recommendations are formulated as a course of action (what to do) rather than specific solutions (how to do), for instance, "I.1 UNFPA KR needs to focus on FP advocacy...to strengthen partnerships... to provide assistance to the MOH... to ensure that existing mechanisms for ensuring access to contraceptives... should pay more attention to practical aspects of the FP training... should contribute to the capacity development of Midwifery Association... needs work with the MoH to invest in improved monitoring and developing" (p. 78). Another example: "UNFPA KR needs to invest resources to consolidate and expand its policy advocacy efforts." It would be useful to know the amount of money the evaluators suggest to relocate for investments. The SRH Recommendation Number 2 also could be more specific in terms of kinds of support that UNFPA KR should continue to provide to "key populations as part of its HIV prevention programs" (support could be financial, expert, or training). It is mentioned in the section 1.3. "Methodology and Process" that an evaluation reference group was formed to "provide feedback on the content and quality of the CPE report". However, nothing said how the recommendations were discussed with the stakeholders. There are 6 strategic and 12 program recommendations. The number of recommendations is manageable, but it is not easy to say if the recommendations are realistic. Strategic Recommendations are all high priority but the Program Recommendations are prioritised.	
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes		
3. Do recommendations appear balanced and impartial?	Yes		
4. Is a timeframe for implementation proposed?	Partial		
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes		

7. Gender	0 1 2 3 (**)	Assessment Level:	Good
To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)		Comment: GEEW integrated in the evaluation scope of analysis as	
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	"Gender Equality and Empowerment of Women" was one of the areas of the UNFPA Country Program 2012 – 2016. Performance indicators for the evaluation questions are presented in the Annex 2. Indicators are designed in a way that ensures GEEW-related data to be collected, for instance, "Evidence of use of...assessments for the development, implementation and updating of the outputs...of the for program areas: SRH, Youth, Gender and PD."
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	"Degree of concurrence of CP outputs and activities...of the four program areas: SRH, Youth, Gender and PD." Evaluation criteria and evaluation questions do not specifically address how GEEW has been integrated into design, planning, and implementation of the Program and the results achieved. There are no specific gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected. It is said in the Methodology section that "The design report for this evaluation did not have explicitly stated objectives to address gender, human rights and vulnerability. These methodological considerations were nonetheless addressed to some extent ... the sampling of stakeholders was developed to achieve a balance representation of women and men... Gender disaggregated data are presented where possible... a separate sub-section was added to discuss gender as a cross cutting issue... the evaluation addressed rights issues as part of the assessment of UNFPA CO advocacy for RH laws and policy...The evaluation site visit schedule and stakeholder interviews, FGDs and client exit interviews attempted to reach vulnerable beneficiaries from marginalized groups e.g. injecting drug users, sex workers with limited success."
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?		2	It is possible to find such notes in the report as "Over 80% of these respondents were women." All of the respondents of the Client/Beneficiary Interviews and Focus Group Discussions were female as well. Thus, in the future it may be useful to elaborate more on gender-responsive evaluation methodology.
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		2	The evaluation findings, conclusions and recommendations reflect a gender analysis to the extent that Gender was a key Program component.

(*) 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).

(**) Scoring uses a four point scale (0-3).
0 = Not at all integrated. Applies when none of the elements under a criterion are met.
1 = Partially integrated. Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.
2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met but still improvement could be done.
3 = Fully integrated. Applies when all of the elements under a criterion are met, used and fully integrated in the evaluation and no remedial action is required.

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	11	82	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
<p>(*) (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.</p> <p>If the overall assessment is 'Fair', please explain</p> <p>• How it can be used?</p> <p>• What aspects to be cautious about?</p> <p>Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory</p> <p>Consideration of significant constraints</p> <p>The quality of this evaluation report has been hampered by exceptionally difficult circumstances: <input type="checkbox"/> Yes <input checked="" type="checkbox"/> No</p> <p>If yes, please explain:</p>				