Organizational unit:	**Year of report:** 2019
Title of evaluation report:	**EVALUACIÓN DEL PROGRAMA DE PAÍS DEL UNFPA: COLOMBIA**
Overall quality of report:	**Very Good** **Date of assessment:** 3 April 2019

Overall comments: The evaluation looked closely at how results were delivered in a complex political context for UNFPA. The evaluation was helped by the fact that the UNFPA country office had an effective, results-based monitoring and evaluation system that provided much of the data on results, especially of what was called "productos". The evaluators were able to supplement the data collection with interviews and limited field visits, as a result of which they could effectively highlight the issues they confronted in measuring outcomes. The main weakness of the evaluation was that the stakeholder consultation process was not described.

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 assessment level followed by main comments . (use 'shading' function to give cells corresponding colour)*	
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	**Good**
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?	Yes	It is well written in Spanish.	
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 68 pages.	
3. Is the report structured in a logical way? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?	Yes	The report is structured in a logical way, with clear distinctions among 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	There is no information in the annexes 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 succinct, providing a review of the methodology, key findings and recommendations within three pages. Considering its use as a standalone section, it would be helpful for the evaluation team to spell-out acronyms initially, before their broader use through the report.	
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	It contains these sections.	
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The Summary is 3 pages.	

2. Design and Methodology	Yes No Partial	Assessment Level:	**Fair**
To ensure that the evaluation is put within its context			
1. Does the evaluation describe the target audience for the evaluation?	Yes	Yes, the target audiences are described: UNFPA country and regional offices (and HQ), government counterparts, civil society, and other UN organizations.	
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The report notes the particular importance of the peace process to end conflict and the need to address challenges in rural areas. There were some constraints on access to certain areas in which the programme was working due to the political changes. Additionally, the election in 2017 led to a change of counterparts in some areas.	
3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?	Yes	It notes that there was no formal theory of change, but it constructs one for the evaluation based on various documents. The evaluation briefly describes the reconstruction of the 'Results and Resources Framework' through the integration of more specific indicators related to gender and human rights.	
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 described in the text and the the evaluation matrix is developed and included in "ANEXO II. MATRIZ DE EVALUACIÓN". The evaluation team defines what is being measured, particularly defining the desired outputs and outcomes of UNFPA. The evaluation team specifically sought to differentiate between what is controlled by UNFPA (output) and what is not (outcome), which has been a common challenge within other UNFPA CPEs.
5. Are the tools for data collection described and their choice justified?	Yes		Yes, the tools for data collection are described (i.e. document review, semi-structured interviews, focus groups and field visits) and justified on the basis of who needs to be consulted and time available in the field (two weeks).
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)?	Partial		Stakeholder numbers are provided by stakeholder group (e.g. state, civil society, internal UNFPA) but with the exception of the UNFPA office and one ministry, the specific civil society organizations or ministries consulted are not described. The evaluation reference group was consulted in design and recommendations, however the overall sstakeholder consultation process/how other stakeholders were involved was not made clear.
7. Are the methods for analysis clearly described for all types of data?	Partial		The methods are described in general in the methodology section, although not in detail by type of approach. In the findings, the evaluation used a combination of quantitative tables and reporting on interviews.
8. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss how any bias has been overcome?	Yes		The main limitations are that not all stakeholders can be reached and there were political issues during the evaluation. How these were addressed is shown; however, broader discussion on limitations and mitigation strategies was insufficient. For example, how does purposive sampling or the inability to travel to specific sites affect the data collected; were remote interviews sufficient as a replacement to field visits?
9. Is the sampling strategy described?	Partial		The sampling straegy is a purposive sample intended to identify sources of data from counterparts and the sample is provided by the UNFPA Office. The evaluators note that it is not representative and is heavily weighted to state officials (Table 3). The basis for selecting places, however, for site visits is not explained.
10. Does the methodology enable the collection and analysis of disaggregated data?	Partial		While the methodology is theoretically sufficient to enable the collection and analysis of disaggregated data, it is not made clear within the text on how this was done in practice, though this was visible in documenty review (which included data disaggregated by gender, age and location)
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes		The evaluators made an effort to acquire data on gender and human rights, which allowed for the assessment on the extent to which the country programme integrated 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		Most of the data came from documents (especially country reports), but, in all cases, these were triangulated with other sources (interviews and observation, for example).
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes		Data was generated from an effective monitoring and evaluation system that was careful to define good indicators and collect data. The evaluators also used data from interviewes judiciously, ensuring its utility for the evaluation. The evaluators noted that they had a strategy to guarantee the quality of data (e.g. use of trustworthy secondary sources) and used triangulation when there were quality issues.
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 evaluators noted when data were not available or when it was not necessarily representative and either did not draw findings from that or sought alternative sources of data, mostly from interviews. The evaluation used data from secondary sources as well as focus group dicsussions and interviews. When gaps were identified, primarily in seconday sources, the evaluation triangulated against data collected from primary sources.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes		While the report does not explicitly discuss how the evaluators ensured that data was collected ethically and non-discriminatorily, data from interviews, for example, was acquired and presented anonymously and confidentiality was guaranteed, suggesting compliance with minimal ethical standards.

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		In each case, the finding was supported by evidence that could credibly substantiaate it.
2. Is the basis for interpretations carefully collected?	Yes		The findings and interpretations are clearly based on the data which was made available during the evaluation period.
3. Is the analysis presented against the evaluation questions?	Yes		The findings were systematically presented by criterion and question.

4. Is the analysis transparent about the sources and quality of data?	Yes	In each case, the source of data was clear.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	While the connection between what UNFPA did as an intervention and "outputs"/"productos" was often made, the connection to outcomes, as the evaluators note, was not always clear. The evaluation report seeks to draw cause and effect links within the discussion of 'achievments, progress and outstanding challenges' in each chapter on outcome areas. The evaluators provide a discussion of outcomes, where possible, however the report primarily focuses on outputs delivered by UNFPA.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The UNFPA programme had a diverse set of target groups and the outcomes for each were clearly shown.
7. Is the analysis presented against contextual factors?	Yes	In each case, the context was included as an explanation for the finding, either positive or negative. UNFPA faced challenges in reaching more rural areas due to protracted conflict within Colombia; in addition, because of the complex political environment surrounding the sensitive issues UNFPA addresses. The evaluation team are careful to include a discussion of contextual factors in the assessment of results.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The evaluators made a particular effort to address cross-cutting issues either in the specific question areas or through a separte section of the evaluation report.

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	The conclusions clearly indicate the findings from which they flow, organized by evaluation question.	
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	Yes, the conclusions build on and go beyond the findings, further contextualizing them (in the broader national context/underlying issues of the country programme). The conclusions offer a clear path to the recommendations.	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of bias.	

6. Recommendations	Yes No Partial	Assessment Level:	**Very good**
To ensure the usefulness and clarity of recommendations			
1. Do recommendations flow logically from conclusions?	Yes	The conclusions from which the recommendation is derived is shown.	
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 clear, their intended users are shown and in each case the specific steps to take are shown. When there are financial and technical implications, these are included.	
3. Do recommendations appear balanced and impartial?	Yes	There is no evidence of partiality.	
4. Is a timeframe for implementation proposed?	Yes	Most recommendations are for immediate action, while others are directed for design of the next country programme.	
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	The recommendations are prioritized (there are two levels of prioirty - high and medium) and they are presented in a way to facilitate management response and follow-up.	

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?	3	GEEW is integrated into the scope of the evaluation, and the indicators in the evaluation matrix are formulated in a way that allows the evaluation to collect data on the extent to which the country programme has integrated gender equality and human rights. gender and human rights. For example, under evaluation question 2, the following indicator is included: "Productos y estrategias que consideran explícitamente las desigualdades y la discriminación en razón del género, la condición étnica y la orientación sexual, y se proponen transformarlas."	:
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	GEEW is built into the evaluation questions, including a dedicated evaluation questions (and the indicators of others). For example, evaluation question 4 is: "Hasta dónde se ha avanzado en la integración transversal de los enfoques de género y derechos humanos en la implementación del Programa de País?"	
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	While the evaluation methodology employs a mixed-methods approach, appropriate to evaluating GEEW considerations and includes a diverse range of data sources (and processes, including triangulation for validation), the stakeholder list is not disaggregated (by gender or other identity), though a preliminary review of names suggests their sample was gender-inclusive.	

4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	3	There were both conclusions and recommendations about gender and how it should be addressed. There are specific conclusions and recommendations about GEEW, but this is also built into other conclusions and recommendations to show the gender aspect.

(*) 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

Quality assessment criteria (scoring points*)	Very good	Good	Fair	Unsatisfactory
	Assessment Levels (*)			
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	80	7	13	
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?
The fair rating is based on unevenness in the evaluation design and in its applications in terms of findings. While the weaknesses of the methods and findings need to be considered, the fact that the conclusions are strong and clearly expressed, suggests that the evaluation can be used to help design the next country program.

• 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: