



Organizational unit: UNFPA Evaluation Office		Year of report: 2020	
Title of evaluation report: Evaluación del Programa País de Panamá - III Ciclo de cooperación 2016-2020			
Overall quality of report: Good		Date of assessment: Nov-20	
Overall comments: This report provides an evaluation of the UNFPA Panama Country Programme. The strengths of the evaluation include the clear integration and analysis of gender and vulnerability into the evaluation questions and matrix, sampling frame, findings and analysis, and recommendations and well-organized conclusions. The evaluation uses a comprehensive and mixed-methods approach, which is appropriate for assessing cross-cutting issues, incorporating indicators and evaluation tools requiring the engagement of diverse groups, including the consultation of beneficiaries. It is limited by incomplete descriptions of the ethical protocols employed in data collection and limited assessment of potential bias as a result of a primarily purposive sampling strategy. The findings reflect a reasonable analysis of the intervention's strengths and weaknesses contrasted against contextual factors, and a clear gender and vulnerability analysis. And, while the cause and effect links between the intervention and end results, including discussion of unintended outcomes, are present, they sometimes are not clearly validated/triangulated by multiple sources of evidence across both the output and outcomes levels. The conclusions provide a balanced perspective and are grounded on the evaluation findings with very good integration of contextual factors. Overall, more attention could be given to clarifying the evaluation design and methodology, as well as incorporating more analysis of qualitative data collected through interviews, focus groups and site visits.			
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 assessment level followed by main comments. (use 'shading' function to give cells corresponding colour)</i>	
1. Structure and Clarity of Reporting		Yes No Partial	Assessment Level: Very good
<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	This report is accessible and easy to understand.
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	This report is within the page limits of 70 pages for a country programme evaluation.
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 structure is coherent and the sections are clearly delineated. The use of textboxes and diagrams contributed to the clarity of the report.
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 contain all the required information.
<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 standalone section, presenting the main findings of the evaluation. However, the recommendations are described quite broadly, limiting the utility of this section as a stand-alone resource for decision-makers.
6. Is there a clear structure of the executive summary, (i.e. i) Purpose; ii) Objectives and brief description of intervention; iii) Methodology; iv) Main conclusions; v) Recommendations)?		Yes	The executive summary is clearly structured and it includes most the required sections, though more substantive details on the country programme interventions could have been provided.
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?		Yes	At three pages, the executive summary is reasonably concise and well within the required five-page limit.
2. Design and Methodology		Yes No Partial	Assessment Level: Very good
<i>To ensure that the evaluation is put within its context</i>			
1. Does the evaluation describe the target audience for the evaluation?		Yes	The evaluation had the dual purpose of accountability and learning through an independent assessment of results achieved in the four programme components and extracting of key lessons for the next programme cycle. As such, the target audience for this evaluation is the UNFPA Country Office in Panama, the LACRO regional office, UNFPA headquarters and national counterparts/stakeholders.
2. Is the development and institutional context of the evaluation clearly described and constraints explained?		Yes	The development and institutional context of the country, and how it relates to the evaluation (such as an assessment of the accuracy of national data), is clearly described in Chapter 2 on the national context.
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 evaluation report describes constructing a Theory of Change in section 3.2.2 as it notes the programme did not have one, minimally meeting this criterion. This section could have been strengthened with a visual representation of the reconstructed ToC as well as a clear description on how the ToC will be used in the evaluation, especially considering a notable limitation related to the gaps in the intervention logic and in available data from the countries results framework. It does, however, present and analyse the country office results framework / UNDAF and includes it as an annex.
<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	The evaluation framework is briefly described in the text through the presentation of the evaluation objectives, questions and a detailed table on methods and tools (Table 2). The full evaluation matrix is presented in annex 2 and includes most necessary components, except assumptions.
5. Are the tools for data collection described and their choice justified?		Yes	The tools for data collection are described in detail in Table 2, including a summary of the types of quantitative and qualitative data gathered from each tool (document review, budgetary analysis, semi-structured interviews, focus groups, and observations). The rationale for their selection is briefly reviewed in the final column, titled 'comments'.
6. Is there a comprehensive stakeholder map? Is the stakeholder consultation process clearly described (in particular, does it include the consultation of key stakeholders on draft recommendations)?		Yes	The report describes the process for stakeholder consultation in multiple sections of the evaluation (methodological process and mapping of actors), noting the validation of findings, conclusions and recommendations with the evaluation reference group and technical teams of the country office on two separate occasions: after fieldwork and in a second presentation and workshop upon drafting of the report. The workshop agenda is included as an annex, and tables and charts clarifying the distribution of stakeholders by type are folded into the report as a form of stakeholder map. The full list of stakeholders provided by the country office is also included as an annex.

7. Are the methods for analysis clearly described for all types of data?	Yes	The report describes the process for the management and analysis of data and there is some detail on budgetary analysis provided in the table of methods used.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	Three key methodological limitations and their mitigation strategies are described.
9. Is the sampling strategy described?	Yes	The sample of stakeholders included in Annex 8 for this evaluation was initially based on a list provided by the evaluation office. The sample of stakeholders was further analysed and prioritized after the document review using the following criteria: (1) the level of actor (e.g. government, private sector, UN agencies, etc.); (2) relationship to each of the country office programmes / priority areas; (3) leadership and engagement with country programmes; (4) gender balance. The areas for assessment were also selected purposively based on the concentration of programmes in those areas (i.e. comprehensive sex education, and areas of programming for this, were not included in the evaluation).
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The use of a mixed methods approach enabled the collection and analysis of gender disaggregated data. Ultimately, 75% of the 111 respondents were female.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The mixed methods approach used is appropriate for assessing cross-cutting issues, incorporating indicators and evaluation tools requiring the engagement of diverse groups, including the consultation of beneficiaries.
3. Reliability of Data	Yes No Partial	Assessment Level: Fair
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	The triangulation of data collected is demonstrated by the detailed sources of evidence in both the evaluation matrix and footnotes throughout the report.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluators clearly identify, and make use of, reliable qualitative and quantitative data sources. In addition, the evaluation notes the lack of consistent information on indicators in the results framework across years and reports, which was noted early and mitigated through the collection and verification of key data points through interviews.
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?	Partial	The evaluation makes explicit some limitations in the secondary data, as mentioned above, though analysis of potential limitations of primary data, either as a result of the sampling process or the country context, is not provided in any detail.
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	The report provides some description on how evaluators' approach to data collection was sensitive to ethical considerations, though it could be more detailed. The report only noted that the transcripts were anonymized to protect the privacy of the respondents, and while the team were careful to engage a broad range of stakeholders, there was not information on the process for ensuring that discussions/interviews were in accessible locations, times, settings suitable for ensuring participation of the most vulnerable.
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 report provides sources of evidence in the text and footnotes, as well as the annex.
2. Is the basis for interpretations carefully described?	Partial	The evaluators describe the basis for interpretations in most cases. However, the findings primarily rely on the document review with minimal analysis of the qualitative data collected. Quotes are used and occasionally provide additional context, but are not analyzed in any substantive manner. For example, it was mentioned how informants from the maternal homes visited cited difficulties with lack of personnel and food and that some users receive 'inadequate treatment' from staff, however the relationships and implications of this to UNFPA programming are not unpacked.
3. Is the analysis presented against the evaluation questions?	Yes	The findings are presented against the evaluation questions.
4. Is the analysis transparent about the sources and quality of data?	Yes	The evaluators are transparent about the sources of data in text, footnotes, and summaries in the evaluation matrix. They are also explicit about the quality of secondary data.
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The cause and effect links between the intervention and end results, including discussion of unintended outcomes, are present though sometimes are not clearly validated/triangulated by multiple sources of evidence across both the output and outcomes levels. The presentation of good practices through text boxes were useful in presenting more clear cause and effect links between the intervention and its results.
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The evaluation findings analyze and show differences across relevant outcomes for diverse target groups.
7. Is the analysis presented against contextual factors?	Yes	The findings are presented against relevant contextual factors at the national and global levels showing both the policy and programmatic factors, as well as social norms and ideologies, affecting results achievement. In addition, the distinct section on unintended outcomes assisted in contextualizing some of the factors facilitating and/or hindering results achievement.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	Analysis of cross-cutting issues is provided both within the description of the national context and in the section on findings. Both sections present disaggregated data and an analysis of issues affecting different vulnerable populations, such as indigenous groups and the LGBT+ population.
5. Conclusions	Yes No Partial	Assessment Level: Very good
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions flow clearly from the findings and are defined as strategic or programmatic. The programmatic conclusions are organized around the same intervention areas which the findings on effectiveness were presented on. The report references the associated evaluation question and recommendation number(s) for each conclusion.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Yes	The evaluation conclusions are well-rounded on both the strengths and weaknesses of UNFPA interventions demonstrating the evaluators' understanding of the underlying contextual and programmatic issues.
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:	Good
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	The recommendations flow logically from the conclusions, and each recommendation notes the conclusions on which it is based.	
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations clearly specify to whom they are directed and provide clear steps for operationalizing each of them.	
3. Do recommendations appear balanced and impartial?	Yes	As noted above, the recommendations flow clearly from the findings and do not show any evidence of bias.	
4. Is a timeframe for implementation proposed?	Partial	The recommendations do not have timeframes for implementation, though are intended to be implemented within the next programme cycle, so it is implied.	
5. Are the recommendations prioritized?	Yes	The recommendations are prioritized as either high or medium-levels of urgency to help facilitate management response.	
7. Gender	0 1 2 3 (**)	Assessment Level:	Good
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>			
1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?		<p>a. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? (Score: 0-3) GEEW is considered as a cross-cutting theme both within the evaluation scope, however is not mainstreamed into the evaluation objectives. (Score = 2).</p> <p>b. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? (Score: 0-3) There is no standalone criterion on GEEW, but as mentioned, it is included as a cross-cutting theme to be addressed, and the assumptions and indicators for EQ1 assesses how GEEW is implemented in line with the priorities set by the international and national policy frameworks and aligned with the UNFPA policies and strategies. EQ5 looks at how cross-cutting issues of youth, gender, interculturality and rights are integrated into the programme design (Score = 3).</p> <p>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3) GEEW was integrated into questions under the Relevance criterion. (Score = 2).</p> <p>d. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results? (Score: 0-3) Assessment of the sufficiency of information collected during the implementation period on specific results on GEEW is not specifically addressed, though there is some analysis provided within the relevance section. (Score = 1).</p>	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		<p>a. Does the evaluation specify how gender issues are addressed in the methodology, including: how data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? (Score: 0-3) The evaluation methodology briefly addresses how gender issues were addressed in the methodology, primarily through the collection and analysis of sex disaggregated data. (Score = 2).</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analysing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3) The methods and data collection tools include key informant interviews, focus group discussions and observations, as well as information on how the sample size was drawn, which specifically included the consideration of gender and vulnerability. The sample size was decent at 111 stakeholders. (Score = 3).</p> <p>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3) Data sources were diverse and triangulation and validation processes described. (Score = 3)</p> <p>d. Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? (Score: 0-3) The evaluation methods and sampling frame address the diversity of stakeholders, but it is not certain to which extent the most vulnerable were reached, despite having beneficiaries included in the sampling frame. (Score = 2)</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3) The evaluation methodology and findings suggest that ethical standards were considered and adhered to, however there is minimal description of the process followed in the report, with the acceptance of cleaning data of identifying information to ensure confidentiality. (Score = 2)</p>	
3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		<p>a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality? (Score: 0-3) Gender analysis is evident in the background section, clearly showing an intersectional analysis of how different social groups are affected by policies, practices, and social norms. (Score = 3).</p> <p>b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable? (Score: 0-3) While the findings include data analysis that explicitly triangulates the sources of data, there is minimal use of quotes or stories to highlight the voices of different groups. Most quotes are from document sources (Score = 2).</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) There is some evidence provided on both the anticipated and unanticipated effects of the interventions on human rights and gender equality, including challenges with reach and equal access to services for vulnerable groups. Unexpected effects are presented under its associated evaluation question. (Score = 3).</p> <p>d. Does the evaluation report provide specific recommendations addressing GEEW issues, and priorities for action to improve GEEW or the intervention or future initiatives in this area? (Score: 0-3) There are multiple recommendations which explicitly address issues of GEEW and excluded groups. (Score = 3).</p>	
(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. 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				
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	31	58	11	
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> <ul style="list-style-type: none"> • How it can be used? • What aspects to be cautious about? <p>Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory The report was good. The methodology was strong but it needed further detail on processes for stakeholder consultation and ethics, as well as potential bias inherent in the design. The report sometimes drew findings off of primarily secondary data sources which weren't adequately triangulated against primary data or data at multiple levels (output/outcome). However, the conclusions clearly brought together the most relevant and evidence-based findings.</p> <p>Consideration of significant constraints 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>				