**Title of report:** Evaluación del Programa País de Panamá - III Ciclo de cooperación 2016-2020  
**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 compared 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.

<table>
<thead>
<tr>
<th>Assessment Levels</th>
<th>Very Good</th>
<th>Good</th>
<th>Fair</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Description</td>
<td>strong, above average, exceptional</td>
<td>satisfactory, respectable</td>
<td>with some weaknesses, still acceptable</td>
<td>weak, does not meet minimal quality standards</td>
</tr>
</tbody>
</table>

### Quality Assessment Criteria

<table>
<thead>
<tr>
<th>1. Structure and Clarity of Reporting</th>
<th>Yes</th>
<th>No</th>
<th>Partial</th>
<th>Assessment Level</th>
<th>Very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>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?</td>
<td>Yes</td>
<td>This report is accessible and easy to understand.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Is the report of a reasonable length? (maximum pages for the main report, excluding annexes: 60 for institutional evaluations; 70 for COPs; 80 for thematic evaluations)</td>
<td>Yes</td>
<td>The executive summary is within the page limits of 70 pages for a country programme evaluation.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>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)?</td>
<td>Yes</td>
<td>The report structure is coherent and the sections are clearly delineated. The use of textboxes and diagrams contributed to the clarity of the report.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?</td>
<td>Yes</td>
<td>At three pages, the executive summary is reasonably concise and well within the required five-page limit.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### 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?  
6. Is there a clear structure of the executive summary: (i) Purpose; (ii) Objectives and brief description of intervention; (iii) Methodology; (iv) Main conclusions; (v) Recommendations?  
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?

### Design and Methodology

<table>
<thead>
<tr>
<th>1. Design and Methodology</th>
<th>Yes</th>
<th>No</th>
<th>Partial</th>
<th>Assessment Level</th>
<th>Very good</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Does the evaluation describe the target audience for the evaluation?</td>
<td>Yes</td>
<td>The target audience for the evaluation includes beneficiaries and affected populations.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Does the development and institutional context of the evaluation clearly described and constraints explained?</td>
<td>Yes</td>
<td>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.</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Does the evaluation report describe the reconstruction of the intervention logic and/or theory of change, and assess the adequacy of these?</td>
<td>Yes</td>
<td>The evaluation report describes reconstructing 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 of 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.</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### In terms of design and methodology

4. Is the evaluation framework clearly described in the text and in the evaluation matrix?  
5. Are the tools for data collection described and their choice justified?  
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)?

---

This report is accessible and easy to understand.
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?  
    Yes

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

4. Analysis and Findings  
   Yes

5. Conclusions  
   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?  
    Yes

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

4. Analysis and Findings  
   Yes

5. Conclusions  
   Yes
6. Recommendations

| Yes | No | Partial | Assessment Level | Good |

**To ensure the usefulness and clarity of recommendations**

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 timelines for implementation, though are intended to be implemented within the next programme cycle, so it is implied.

5. Are the recommendations prioritised?
   - Yes
   - The recommendations are prioritised as either high or medium-low of urgency to help facilitate management response.

7. Gender

| 0 | 1 | 2 | Assessment Level | Good |

**To assess the integration of Gender Equality and Empowerment of Women (GEEW)**

1. Is GEEW integrated in the evaluation scope of analysis and indicators designed in a way that ensures GEEW-related data to be collected?
   - Yes
   - GEEW is considered as a cross-cutting theme both within the evaluation scope, however it is not mainstreamed into the evaluation.
   - 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, however it is included as a cross-cutting theme to be addressed, and the assumptions and indicators for EQ5 assess how GEEW is implemented in line with the priorities set by the international and national policy frameworks and aligned with the UN Women policies and strategies. EQ5 looks at how cross-cutting issues of youth, gender, interculturality and rights are integrated into the programme design (Score = 3).

2. A gender-responsive methodology was used, including gender-responsive methods and tools, and data analysis techniques?
   - Yes
   - Data sources were diverse and triangulation and validation processes described. (Score = 3)

3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?
   - Yes
   - 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 affected by the intervention, particularly the most vulnerable, where appropriate! (Score: 0-3)

4. 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)
   - Data collection and analysis methods integrate gender considerations and ensure data collected is disaggregated by sex? (Score: 0-3)

5. Are the recommendations prioritised?
   - Yes
   - The recommendations are prioritised as either high or medium-low of urgency to help facilitate management response.

(*) This assessment criteria is fully based on the UN/DHAP Scoring Tool. Each sub-criteria shall be equally weighted (in correlation with the calculation in the tool and totaling the score: 11:12 = very good; 8:10 = good; 4:7 = Fair; 0:4= 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

<table>
<thead>
<tr>
<th>Quality assessment criteria (scoring points*)</th>
<th>Very good</th>
<th>Good</th>
<th>Fair</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Structure and clarity of reporting, including executive summary (7)</td>
<td>7</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Design and methodology (13)</td>
<td>13</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Reliability of data (11)</td>
<td></td>
<td>11</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Analysis and findings (40)</td>
<td></td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Conclusions (11)</td>
<td>11</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Recommendations (11)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>7. Integration of gender (7)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total scoring points</td>
<td>31</td>
<td>58</td>
<td>11</td>
<td></td>
</tr>
</tbody>
</table>

#### Overall assessment level of evaluation report

<table>
<thead>
<tr>
<th></th>
<th>Very good</th>
<th>Good</th>
<th>Fair</th>
<th>Unsatisfactory</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall assessment level of evaluation report</td>
<td>Good</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

(*) (a) Insert scoring points associated with criteria in corresponding column (e.g. - if 'Analysis and findings' has been assessed as 'Good', enter 40 into 'Good' column.
(b) Assessment level with highest 'total scoring points' determines 'Overall assessment level of evaluation report'. Write corresponding assessment level in cell (e.g. 'Fair').
(c) Use 'shading' function to give cells corresponding colour.

If the overall assessment is 'Fair', please explain:

- How it can be used:
- What aspects to be cautious about:

Where relevant, please explain the overall assessment Very good, Good or Unsatisfactory

The 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.

### Consideration of significant constraints

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

<table>
<thead>
<tr>
<th>Yes</th>
<th>No</th>
</tr>
</thead>
</table>

If yes, please explain: