
Overall Assessment: The CP evaluation lacks focus considering that the MTR took place a year earlier, and does not display any attempt by the evaluation team in the methodology to build on the detailed ToRs and identify an appropriate focus for the evaluation. The evaluation relies heavily on the CPAP tracking tool as a data source. The findings section is poorly presented, utilizing a confusing system of sub-sections that is more descriptive than analytical. The chosen structure of arranging findings by mandate fails to present a cohesive picture. Recommendations are not presented separately but listed at the end of each programme activity and often lack sufficient detail or practicality as well as being too numerous and unprioritised. There is an absence of any detailed conclusions. Overall this report fails to meet the requirements of the ToRs, mainly through the evaluation team’s focus on description rather than analysis.

Quality Assessment criteria

<table>
<thead>
<tr>
<th>Assessment Levels</th>
<th>Very Good</th>
<th>Good</th>
<th>Poor</th>
<th>Unsatisfactory</th>
</tr>
</thead>
</table>

1. Structure and Clarity of Reporting
To ensure report is user-friendly, comprehensive, logically structured and drafted in accordance with international standards.

Checklist of minimum content and sequence required for structure:
- i) Acronyms; ii) Exec Summary; iii) Introduction; iv) Methodology including Approach and Limitations; v) Context; vi) Findings/Analysis; vii) Conclusions; viii) Recommendations; ix) Transferable Lessons Learned (where applicable)
- Minimum requirements for Annexes: ToRs; Bibliography List of interviewees; Methodological instruments used.

Poor
The structure is comprehensive and the report is reasonably clearly structured although a few elements would benefit from extra detail, such as Context. A detailed list of acronyms is provided and the ToRs are mostly complete, particularly with regard to methodology. However there is no separate recommendations section, with the evaluators instead listing recommendations at the end of each output, which is a structural shortcoming that inhibits clarity of the report, despite the structure being clearly outlined in the ToRs.
### 2. Executive Summary

**To provide an overview of the evaluation, written as a stand-alone section and presenting main results of the evaluation.**

Structure (paragraph equates to half page max):
- i) Purpose, including intended audience(s); ii) Objectives and Brief description of intervention (1 para); iii) Methodology (1 para); iv) Main Conclusions (1 para); v) Recommendations (1 para). Maximum length 3-4 page

**Unsatisfactory**

The Executive Summary contains little apart from a list of 25 findings, with a single paragraph outlining the purpose of the evaluation. There is no mention of methodology or design. Overall this chapter fails to provide the reader with a comprehensive overview of the report and does not effectively operate as a stand-alone section.

### 3. Design and Methodology

**To provide a clear explanation of the following elements/tools**

Minimum content and sequence:
- Explanation of methodological choice, including constraints and limitations;
- Techniques and Tools for data collection provided in a detailed manner;
- Triangulation systematically applied throughout the evaluation;
- Details of participatory stakeholders’ consultation process are provided.
- Whenever relevant, specific attention to cross-cutting issues (vulnerable groups, youth, gender equality) in the design of the evaluation

**Poor**

A two page introduction section describes the objectives of the evaluation. The main tasks are outlined with some detail including the document review process, interviews (questions are provided in annex) and data collection and enables triangulation. However methodological choice is not mentioned and only one limitation is listed (documentation language).

### 4. Reliability of Data

**To clarify data collection processes and data quality**

- Sources of qualitative and quantitative data have been identified;
- Credibility of primary (e.g. interviews and focus groups) and secondary (e.g. reports) data established and limitations made explicit;

**Poor**

The majority of data is taken from the CPAP tracking tool, which is relied upon heavily by the evaluators but is not acknowledged as a limited source and the reliability of the data in the Tracking tool is not discussed. Furthermore, several indicators lack baseline data which is not commented upon. Interviewees are also limited to UNFPA and government staff only.
### 5. Findings and Analysis

**To ensure sound analysis and credible findings**

**Findings**
- Findings stem from rigorous data analysis;
- Findings are substantiated by evidence;
- Findings are presented in a clear manner

**Analysis**
- Interpretations are based on carefully described assumptions;
- Contextual factors are identified.
- Cause and effect links between an intervention and its end results (including unintended results) are explained.

**Poor**
This section is arranged by mandate area and by CPAP outputs, with indicators and baselines presented in table form, with a brief comment from the evaluation team. The accompanying narrative is often extensive and arranged in sub-sections including: description of activities; assessment of results; best practices; facilitating factors; constraints/challenges; recommended actions; and cross-cutting themes. The use of this fragmented sub-section approach makes it extremely difficult to focus on evidence-based analysis amongst the extensive activity descriptions.

The evaluation team frequently begins by describing the activities as appropriate (e.g. ‘all three strategies are relevant and appropriate for attainment of the eleven results’) with no additional explanation followed by detailed descriptions of activities without accompanying analysis (e.g. ‘in order to contribute to RH services in the population, a number of training activities have either been implemented or are being implemented’). Overall there is a significant lack of analysis in this section which is mostly descriptive.

Where analysis can be found, for example the described Constraints/Challenges, it does not correspond to the Recommended Actions. There is little reference to data collected independently or interviews.

### 6. Conclusions

**To assess the validity of conclusions**

- Conclusions are based on credible findings;
- Conclusions are organized in priority order;
- Conclusions must convey evaluators’ unbiased judgment of the intervention.

**Unsatisfactory**

The Conclusions section is less than a page, and makes very generic statements (e.g. ‘there have been changes in indicators at both outcome and output levels for all three programme components’). This section performs little useful function in terms of evaluation.
### 7. Recommendations

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

- Recommendations flow logically from conclusions;
- Recommendations must be strategic, targeted and operationally-feasible;
- Recommendations must take into account stakeholders’ consultations whilst remaining impartial;
- Recommendations should be presented in priority order

**Unsatisfactory**

There is no separate recommendations section but Recommended Actions are located at the end of each component output sub-section in the Findings chapter. This scattered structure is extremely unhelpful for the reader. Recommended Actions often lack detail (e.g. ‘build capacity for the collection and analysis of data at district level’), are impractical (e.g. ‘improve quality of care and services in all components of reproductive health’), or are simple statements of problems rather than being action statements. The report also includes recommendations to other bodies, which is not relevant. As well as not being prioritized, the number of recommendations is too great (e.g. 14 for a single output).

### 8. Meeting Needs

**To ensure that Evaluation Report responds to requirements (scope & evaluation questions/issues/DAC criteria) stated in the ToR (ToR must be annexed to the report).**

In the event that the ToR do not conform with commonly agreed quality standards, assess if evaluators have highlighted the deficiencies with the ToR.

**Unsatisfactory**

The ToRs are extensive with detailed information and requirements, including highly specific evaluation questions. However, these have not been wholly fulfilled by the evaluation team and in particular the structure lacks the requested Recommendations section. Furthermore, the MTR was conducted in 2009 and it would have been prudent for the evaluation team to undertake a more focused CPE based on the results rather than ‘using the MTR of 2009 as a base’.
<table>
<thead>
<tr>
<th>Quality assessment criteria (and Multiplying factor *)</th>
<th>Assessment Levels (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unsatisfactory</td>
</tr>
<tr>
<td>5. Findings and analysis (50)</td>
<td>50</td>
</tr>
<tr>
<td>6. Conclusions (12)</td>
<td>12</td>
</tr>
<tr>
<td>7. Recommendations (12)</td>
<td>12</td>
</tr>
<tr>
<td>8. Meeting needs (12)</td>
<td>12</td>
</tr>
<tr>
<td>3. Design and methodology (5)</td>
<td>5</td>
</tr>
<tr>
<td>4. Reliability of data (5)</td>
<td>5</td>
</tr>
<tr>
<td>1. Structure and clarity of reporting (2)</td>
<td>2</td>
</tr>
<tr>
<td>2. Executive summary (2)</td>
<td>2</td>
</tr>
<tr>
<td>TOTAL</td>
<td>38</td>
</tr>
</tbody>
</table>

(*) Insert the multiplying factor associated with the criteria in the corresponding column e.g. - if “Finding and Analysis” has been assessed as “good”, please enter the number 50 into the “Good” column. The Assessment level scoring the higher number of points will determine the overall quality of the Report.

OVERALL QUALITY OF REPORT: Poor