

OVERALL QUALITY RATING: Poor

Summary: The evaluation of the Sierra Leone 4th Country Programme uses a rigorous and clear methodology to collect and triangulate data from a variety of quantitative and qualitative sources. The evaluators meet most of the needs of the ToR and point out deficiencies in its design when appropriate. The evaluators appropriately recognize that they cannot evaluate progress on many indicators laid out by the CPAP due to lack of Baseline data, and use the data that they do have to determine whether UNFPA’s outputs are on track to achieve the unmeasured indicators. However, despite the use of data to establish progress towards outcomes, there is insufficient analysis of the causal relationship between UNFPA’s programs and outcome indicators, which are often measured nationally and may be influenced by many factors. Furthermore, few of the recommendations are actionable and strategic, and many are directed at the Government of Sierra Leone rather than UNFPA, which is not acceptable.

<table>
<thead>
<tr>
<th>Quality Assessment criteria</th>
<th>Assessment Levels</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Structure and Clarity of Reporting</strong></td>
<td>Very good</td>
</tr>
<tr>
<td><em>To ensure report is user-friendly, comprehensive, logically structured and drafted in accordance with international standards.</em></td>
<td><strong>Good</strong></td>
</tr>
</tbody>
</table>
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.

Good

The Executive Summary clearly states the purpose of the evaluation, including the intended audience, objectives and intervention, methodology, main conclusion, and recommendations. The Executive Summary functions as a stand-alone document. The main conclusions and recommendations, however, are too long at 3 and 1.5 pages respectively.

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;
- Details on how cross-cutting issues (vulnerable groups, youth, gender, equality) were addressed in the design and the conduct of the evaluation.

Good

The methodological choices are clearly described and meet UNFPA standards. The evaluators describe limitations to the chosen methodology (i.e. purposive sampling) and justify their choices (i.e. there is no sampling frame for beneficiaries) (p. 5). The data collection process is described in detail and has some notable components, including that the evaluators sent detailed questionnaires in advance to stakeholders and covered 8 of 12 districts (p. 5, see also annexes 5 and 7 for the schedule of interviews and questionnaires). The evaluators used triangulation to compare information from various sources and show good judgment in recognizing variation in the credibility of different sources. Details of the participatory stakeholders’ consultations are provided, but a description of how cross-cutting issues were incorporated into the design and conduct of the evaluation are omitted. The final subsection of the Methodology section (3.4 Structure of the Evaluation Report) is not required and is unnecessary.

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;
- Disaggregated data by gender has been utilized where necessary.

Good

The evaluators use both quantitative and qualitative information in their analysis. They demonstrate a clear understanding of variation in credibility across official and unpublished sources (p. 5). They use data from a broad range of sources (government, UNFPA, implementing
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**
Findings stem from the data analysis. Though many findings are substantiated by some evidence, there is generally not sufficient evidence to prove cause and effect relationships between the intervention and its results. In some cases where adequate evidence does not exist, the evaluators appropriately recognize that they cannot make a finding as to whether the programme has reached its intended goals. For example, on Output 2 of the Reproductive Health component, the evaluators judge that “Given the extent to which activities have been carried out, it is likely that the strategies adopted and activities being carried out will lead to the realization of the output indicators 2 and 3 by the end of the programme cycle; much has been delivered in support of output indicator 1, but there are no baselines and targets to make an informed judgment about the extent of achievement of the output” (p. 17).

Given that the evaluation is not resourced to assess impact, the subsection on impact (4.2.4 Impact, p. 27) should not be included.

The analysis does not provide sufficient evidence of cause and effect links from an intervention to its end results. Contextual factors are identified in some cases (particularly when they have hindered the implementation of the program as on p. 20) but contextual factors that have contributed to the success of programs or attainment of goals are not mentioned. The assumptions that are needed to analyze the cause and effect links should be stated more explicitly.
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.

**Poor**
Conclusions are based on findings. However, the links between conclusions and the relevant findings are not made explicit enough. For example, the authors conclude that UNFPA “helped strengthen at all levels the delivery of maternal and child health care in the project districts in the country” (p. 50) but do not clearly link this conclusion to the supporting findings (p. 27). Conclusions are not organized in priority order. There is no evidence of bias in the conclusions.

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**
Recommendations should not be targeted at the government of Sierra Leone, but several in this evaluation are. For example, “In order to facilitate the activities of the Population Secretariat, Government should make adequate financial provision for it as part of the regular budgeting” (p. 55). Recommendations must be targeted at specific audiences within UNFPA, such as the country office, regional offices, or headquarters, in order to be strategic and operationally feasible.

Not all of the recommendations flow logically from the conclusions, particularly the recommendations and conclusions on the gender component (p. 51, 55). There is no evidence that the recommendations are not impartial. Recommendations are not presented in priority order.

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.*

**Good**
The evaluation meets most of the needs outlined in the ToR. One deficiency is that the evaluation does not clearly select and answer a set of evaluation questions although the ToR provides sample evaluation questions.

The ToR includes evaluation questions on the impact of the
country programme; the evaluators appropriately recognize that they do not have the time or data to measure impact (p. 5).

<table>
<thead>
<tr>
<th>Quality assessment criteria (and Multiplying factor *)</th>
<th>Assessment Levels (*)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Very good</td>
</tr>
<tr>
<td>1. Structure and clarity of reporting (2)</td>
<td></td>
</tr>
<tr>
<td>2. Executive summary (2)</td>
<td></td>
</tr>
<tr>
<td>3. Design and methodology (5)</td>
<td></td>
</tr>
<tr>
<td>4. Reliability of data (5)</td>
<td></td>
</tr>
<tr>
<td>5. Findings and analysis (50)</td>
<td></td>
</tr>
<tr>
<td>6. Conclusions (12)</td>
<td></td>
</tr>
<tr>
<td>7. Recommendations (12)</td>
<td></td>
</tr>
<tr>
<td>8. Meeting needs (12)</td>
<td></td>
</tr>
<tr>
<td><strong>TOTAL</strong></td>
<td>26</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.