

EQA for UNFPA Indonesia Country Programme Evaluation (2011-2015)



Title of Evaluation Report: UNFPA COUNTRY PROGRAMME EVALUATION: INDONESIA 2011 – 2015

Overall Quality Rating: Poor

Overall Assessment: The evaluation report has attempted to reflect the fact that the UNFPA programme in Indonesia is both large and complex and has undergone modification during the period. However, this has adversely affected the structure of the report and its readability. Although some sections such as gender findings do measure progress against the country programme results, the report is organized around “core programmes”, and not results. The evaluation is largely an assessment of the extent to which these “core programs” have been implemented. The report contains a mix of data collection methods and stakeholder consultation is clearly indicated, but there is a lack of data to support findings. The findings are organized by programme area and within them by an excessive number of evaluation questions which make the evaluation dense and there is no clear sense of the main results of the programme. The conclusions are presented clearly. The recommendations are poorly presented, too numerous and could have been made more effective by prioritization.

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<p>I. Structure and Clarity of Reporting <i>To ensure report is user-friendly, comprehensive, logically structured and drafted in accordance with international standards.</i> Checklist of minimum content and sequence required for structure:</p> <ul style="list-style-type: none"> • 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	
			<p>All of the required sections are included; however, the chosen structure undermines the report’s readability. In the findings section, there are ten subsections, some of which cover cross-cutting issues (Youth) whereas some programme components are split into multiple sections. For example, the programme component “Reproductive Health” includes both capacity building for reproductive health policies and implementing the MISP, but these are split into separate sections (p6). Breaking up the findings section into ten different subsections detracts from the presentation of the overall findings. The evaluators stated that they decided to structure the findings by programme area and within that by the questions in response from comments from the stakeholder consultations. In addition, these sections do not</p>	

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
	<p>correspond to the country programme outcomes, which makes it difficult to track progress against expected results.</p> <p>The required list of acronyms is included as an appendix rather than at the front of the document. The only methodological instrument that is included in the annexes is the evaluation matrix (p120). The inclusion of other tools such as interview guides would have strengthened the report.</p> <p>There are a number of poorly formulated sentences and spelling errors in the report that hinder readability, such as “For Gender, while still striving to mainstream gender sensitivity, the gender component <u>payed</u> particular attention to gender based violence” (p30).</p>			
<p>2. Executive Summary <i>To provide an overview of the evaluation, written as a stand-alone section and presenting main results of the evaluation.</i> Structure (paragraph equates to half page max):</p> <ul style="list-style-type: none"> 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. 	<p>Poor While the summary contains all of the required sections and is a stand-alone document, it exceeds the recommended length by almost 2 pages, largely because the recommendations are structured according to the programme categories rather than being synthesized by type, which makes the summary too long and difficult to read and each of the other components is also too long, including the methodology which is six paragraphs rather than one.</p>			
<p>3. Design and Methodology <i>To provide a clear explanation of the following elements/tools</i> Minimum content and sequence:</p> <ul style="list-style-type: none"> 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; 	<p>Good The design is well-explained, in particular the explanation of constraints regarding methodological choice. Some limitations are described in detail and are very thoughtful, particularly the discussion of exclusion of past partners of UNFPA, and frank, such as the discussion of the limitations of the capacities of the evaluation team.</p> <p>Selection of interviews is explained with evaluators noting that two factors influenced the fact that almost all of the interviews and site visits were in Jakarta: the programme emphasis was on national level institutions and the lack of resources to go to more than one other site (Jogjakarta). Within that constraint, the evaluators made an effort</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
<ul style="list-style-type: none"> Details on how cross-cutting issues (vulnerable groups, youth, gender, equality) were addressed in the design and the conduct of the evaluation. 	<p>to triangulate and there were well-described consultations, which were taken into account. However, there were many more evaluation questions (and sub-questions) than recommended and the second question under effectiveness is inappropriately detailed and too unfocused to be considered an evaluation question, and this led to a certain amount of duplication (which was noted by the evaluators). While not undermining the quality of the whole section, this is still a shortcoming of this section.</p> <p>Triangulation is clearly used, and the methodology section includes an example of how they used triangulation between four sources to get a more reliable assessment of one fact. Details of the consultative process are provided. The techniques used for data collection are sufficiently described (selection of interviewees, documents consulted, site visits), but the tools used are not provided, which undermines the quality of the report.</p>			
<p>4. Reliability of Data</p> <p><i>To clarify data collection processes and data quality</i></p> <ul style="list-style-type: none"> 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. 	<p>Poor</p> <p>In many of the findings the source of the data is not discussed. Assertions like “though there has been an increase in knowledge of GOI and stakeholders, usage of the knowledge has as yet not been optimal” (p36) are made without any indication of what data was used to make that finding.</p> <p>There is evidence that triangulation was used to check the validity of data. For instance, the evaluators used information in published reports to triangulate information collected through their own field visits and interviews (ex. “The 2013 publication “<i>What have we learned: Good practices documentation of the UNFPA Humanitarian Programme in Indonesia from 2005-2012</i>” confirms the evaluation team’s impressions...” p51). The evaluators also used information from multiple sources to assess the programme (ex. “...DAC in both</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
	<p>Jayapura and Merauke and NAC reported that the UNFPA system was easy to use for budgeting and reporting.” P62).</p> <p>Quantitative data are rarely presented in the report, and there are many cases where data disaggregated by gender would have been appropriate but is not used. For instance, in the discussion of UNFPA’s work with youth during crises, no data is presented on the number of youth who have been engaged or the gender breakdown of those youth (p51). Similarly, disaggregated data on the Youth Advisory Panel is not presented (p72, p80).</p>			
<p>5. Findings and Analysis <i>To ensure sound analysis and credible findings</i> <u>Findings</u></p> <ul style="list-style-type: none"> • Findings stem from rigorous data analysis; • Findings are substantiated by evidence; • Findings are presented in a clear manner <p><u>Analysis</u></p> <ul style="list-style-type: none"> • 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. 	<p>Poor</p> <p>There are a very large number of findings, partly based on the fact that the programme has seven major components. The report largely assesses the degree to which the programme activities have been completed on-schedule. Because the findings are focused on reporting activities, discussion of context, assumptions, and cause and effect links is often limited. Some of the sections on the effectiveness of advocacy lack discussion of context and cause and effect links. For instance, in evaluating the effectiveness of the MISP in humanitarian settings, the evaluators claim that “significant progress has been made by UNFPA over the course of the country programme in the humanitarian area, which has contributed to the institutionalization of MISP in relevant GOI regulations, guidelines and systems for health disaster preparedness and response” (p51). However, there is no discussion of whether these laws and guidelines would have happened without UNFPA’s intervention, and how UNFPA brought the change about. Some of the findings are unrelated to the evaluation question that they purport to address. For instance, in the discussion of efficiency of Population Dynamics, the evaluators mention that many of the analyses and reports that UNFPA has produced have not been used by the GOI</p>			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
	(p37). In fact, this is less a sign of inefficiency than relevance.			
6. Conclusions <i>To assess the validity of conclusions</i> <ul style="list-style-type: none"> • Conclusions are based on credible findings; • Conclusions are organized in priority order; • Conclusions must convey evaluators' unbiased judgment of the intervention. 	Good The conclusions summarize the findings well and in that sense help compensate for the excessive details in the findings sections. They are not organized in priority order, but are structured by programme area. The conclusions do reflect an unbiased appraisal by the evaluators.			
7. Recommendations <i>To assess the usefulness and clarity of recommendations</i> <ul style="list-style-type: none"> • 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 	Poor Recommendations are not presented in priority order, and are not numbered or presented clearly, instead presented in blocks of narrative containing multiple recommendations. This, in addition to the large number of recommendations, significantly impedes the clarity of the recommendations as a whole: for example, in the section on population dynamics there is a recommendation about UNFPA's recruitment process. While recommendations are linked logically to the conclusions, it is difficult to draw general recommendations for the next country programme. While most recommendations are strategic, targeted, and operationally feasible, not all are specific and actionable. Some are vague in what UNFPA should do (ex. "UNFPA should fully support the conduct of the Inter-Censal Population Survey planned for 2015 in as many ways as it can" on p. 96 and "It is recommended that UNFPA provide assistance to the Ministry of Health..." on p. 94). None of the recommendations have specific time frames.			
8. Meeting Needs To ensure that Evaluation Report responds to requirements (scope & evaluation questions/issues/DAC criteria) stated in the	Poor The ToR requires that the evaluators use fewer than 10 evaluation questions; however, there are 16. The evaluators have not reduced the			

Quality Assessment criteria	Assessment Levels			
	Very good	Good	Poor	Unsatisfactory
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.	number of evaluation questions, and in fact increased the number. The ToR also requests that the evaluation focuses on eight “thematic areas,” which were not ultimately used to organize the report.			

Quality assessment criteria (and Multiplying factor *)	Assessment Levels (*)			
	Very good	Good	Poor	Unsatisfactory
1. Structure and clarity of reporting (2)			2	
2. Executive summary (2)			2	
3. Design and methodology (5)		5		
4. Reliability of data (5)			5	
5. Findings and analysis (50)			50	
6. Conclusions (12)		12		
7. Recommendations (12)			12	
8. Meeting needs (12)			12	
TOTAL		17	81	

(*) 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