Organizational unit:	UNFPA		Year of report:	2018
Title of evaluation report:	GOVERNMENT OF MALAWI/UNFPA SEVENTH COUNTRY PROGRAMME 2012-2018			
Overall quality of report:	**Very Good**		Date of assessment:	8 June 2018

Overall comments: The report has been rated very good particularly because it has provided a clear analysis of the causal connections between the UNFPA-funded programs and projects and the observed results. The evaluation carefully reviewed the theory of change used by the programme and chose methods of data collection and analysis that would permit showing which outputs led to the expected outcomes and why. These are reflected in the findings which, on the whole, are backed up by the data and analysis. Conclusions are clearly written and provide a solid understanding of the findings. The recommendations address the conclusions and offer a solid basis for improvements in the next country programme.

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		*Insert assessment level followed by main comments. (use 'shading' function to give cells corresponding colour)*	
1. Structure and Clarity of Reporting	Yes No Partial	Assessment Level:	**Fair**
To ensure the report is comprehensive and user-friendly			
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	The report is easy to read and understand, with the consultants using tables and graphs to present the data. While there are some minimal grammatical and design errors, these did not diminish the readability of the text.	
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	The report is 75 pages long, excluding the annexes.	
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 is structured in a logical way: analysis and findings, conclusions, and recommendations are each presented as separate chapters. While this is not a distinct section in the report, there is a section 4.5. "Lessons Learnt" in the Executive Summary.	
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?	Partial	Information on the stakeholder consultation process is briefly presented. There is a description of the consultation process in the section "stakeholder selection" (pp. 5 - 6), but the evaluation does not go into detail with regard to how a participatory process was designed and implemented.	
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?	Yes	An executive summary is written as a stand-alone section and presents the main results of the evaluation.	

6. Is there a clear structure of the executive summary, (i.e. i) Purpose, including intended audience(s); ii) Objectives and brief description of intervention; iii) Methodology; iv) Main conclusions; v) Recommendations)?	Partial	There is a clear structure of the executive summary, but the evaluators do not specify the purpose and intended audiences explictly, though it can be iimplied from the key objectives "…to demonstrate accountability of the UNFPA 7th Country Programme for the relevance of its programme to a wide range of stakeholders as it relates to the Malawian context and to generate evaluative evidence from the 7th Country Programme (2012-2016) and its extension (2017 to 2018) and draw lessons that will guide the design of next Country Programme" (xiii). The executive summary does not have a section on the "main conclusions" as such, but conclusion statements are included in the section titled "Findings."
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	No	The executive summary is 8 pages long.

2. Design and Methodology	Yes No Partial	Assessment Level: **Good**
To ensure that the evaluation is put within its context		
1. Does the evaluation describe the target audience for the evaluation?	Partial	While the Design and Methodology section does not describe the target audience for the evaluation, but it is explained in the Terms of Reference for the evaluation (which is an annex to the evaluation report as well) "…UNFPA at Country, Regional and headquarter Offices, the Organization's executive border, national internal and external key stakeholders" (p. ii).
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The development and institutional context of the evaluation is clearly described in chapter 2 with regards to programmatic areas, including the role of external assistance. The evaluation effectively utilizes statistics from international sources, national government, and surveys. Constraints and challenges are highlighted, including, for example, poverty and income inequalities as well as the devaluation of Malawian Kwacha.
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 intervention logic is presented in the chapter 4 as a summary achievements table for program components. The consultants reviewed and discussed the theory of change in the section 3.2.3. "Theory of change process." The section has a very thorough and easy-to-read theory of change that is shown as a figure, which has been referenced and used throughout the evaluation.
To ensure a rigorous design and methodology		
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 clearly described in Chapter 1. The evaluation matrix establishes the evaluation questions, Criteria/Focus Areas, Assumptions to be assessed, Indicators, Sources of Information, Methods and Tools for data collection.
5. Are the tools for data collection described and their choice justified?	Yes	The tools for data collection are described and explained in the Introduction (p. 4) and the Evaluation Matrix. Most of the data come from document reviews, but the interviews and field visits are clearly described and 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)?	Partial	There is no comprehensive stakeholder map, but the stakeholder consultation process is covered in the Introduction section: Inception Phase, Field phase, Methods for data analysis and validation, Stakeholder selection (pp. 3-6). The consultants note that "Relevant stakeholders were involved at the different stages of the country programme evaluation including design, data collection, data analysis, and reporting especially at the recommendation formulation process, debriefing, and dissemination stages, as were appropriate. A list of stakeholders selected and interviewed is included in Annex 5" (p. 6).
7. Are the methods for analysis clearly described for all types of data?	Yes	The methods for analysis are clearly described for all types of data: content analysis, contribution analysis, and triangulated analysis (pp. 5-6).
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	The consultants describe and explain the methodological limitations and bias: the duration of the assignment, inability to collect quantitative data from the field, the availability and quality of relevant documents and reports. They note, for example, "Accessibility and availability of data from hard to reach areas: Access to most of the targeted sites is in remote parts of the locations in the country where the roads are not tar marked" and "Some of the methodological challenges encountered by the evaluation team include: missing quarterly progress reports for some years leading to incomplete information; and unavailability of most stakeholders and beneficiaries for some interviews. These limitations were, however, mitigated by extensive document reviews and other information sources."
9. Is the sampling strategy described?	Yes	While the sampling approach is not explicitly described in detail in the methodology (p. 4) and the annexes, in fact the evaluation used a purposive sample, based on visiting all of the districts in which the UNFPA programme worked directly. There were five districts for SRE work and 8 additional for gender and youth.
10. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The methodology (p. 4) explains the collection and analysis of disaggregated data by type of stakeholders (policy makers, programme heads in government, the UN agencies, and civil society organisations) and by location (interviews were held at national level and in each of the nine districts). It is also evident from the annex 4 that the evaluators interviewed both men and women (annex 4) and disaggregated data by cadre, sex, level of operation (p. 33), and by other criteria when necessary.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	Cross-cutting issues are addressed in the analysis of the findings as program components mainstream gender, advocacy and human rights (pp. 21-24). Interview Guide for Beneficiaries (SRH/P&D/GE/HIV/AIDS) incorporates questions with regards to all program components, including gender equality (Annex 3).

3. Reliability of Data	Yes No Partial		Assessment Level:	Good
To ensure quality of data and robust data collection processes				
1. Did the evaluation triangulate data collected as appropriate?	Yes		The evaluators triangulated data "based on systematic cross-comparison of findings by data sources and by data collection methods" (pp. 4-5). There are examples in the text such as "Analysis of documents, annual reports from the implementation partners and CO Programme Officers showed…" (p. 31).	

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 (p. 4, the annex 2, the annex 4, the annex 6).
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?		Yes	The consultants discuss the limitations in primary and secondary data sources: "The scope of this exercise did not allow the team to collect quantitative data from the field, thus our analysis and conclusions are based on quantitative data collected from the Country Office and by secondary sources. This is already a source of bias. However, our use of triangulated methodology mitigated the bias that would have been introduced into the evaluation" (p. 7).
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?		Partial	While the UNEG documents on discrimination are referenced in the text (p. 6), the evaluators do not explain how the data being collected has been with a sensitivity to issues of discrimination and other ethical considerations.

4. Analysis and Findings		Yes No Partial	Assessment Level:	**Very good**
To ensure sound analysis and credible findings				
1. Are the findings substantiated by evidence?		Yes	The findings are substantiated by evidence. For instance, the evaluators provide references to the analysis of documents, annual reports, and surveys.	
2. Is the basis for interpretations carefully described?		Yes	The consultants provide examples, references, and explanations for their interpretations. For instance, the evaluators discuss that "From document reviews, performance indicators show that 66% of health facilities in the 5 focus districts provide EmONC functions, above the target of 40%... For example, in 2012, BEmOC facilities under UNFPA in Dedza were five (5)…On the other hand, other facilities … were either not integrating services or partially integrating the services. Reasons for partial or no integration were space availability with fears of integration causing congestion... However, even for those who had fully integrated the services like Mtakataka health facility, some rooms could not allow full integration of services due to limited space and if provider is deficient in one area of service provision. There is need for UNFPA to consider renovation of structures to accommodate full integration of services" (p. 36).	
3. Is the analysis presented against the evaluation questions?		Yes	Yes, the findings and analysis were structured in terms of the evaluation questions, and then by the UNFPA programme area under evaluated (of which there were three).	
4. Is the analysis transparent about the sources and quality of data?		Yes	The analysis is transparent about the sources and quality of data. While most of the data has been extracted from document review, field observations and interviews are prioritized in the data analysis. For instance, the evaluators discuss the quality of data obtained from different sources: "From annual reports and interviews with IPs, UNFPA supported the development of national SRH/HIV integration strategy, handbook for community health workers and National SRH/HIV Reference manual and guidelines. The IPs claimed these were made use of by them as the needs arose, although the team did not see any evidence" (p. 37).	

5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The evaluators were clear about the causal connections, where UNFPA activities led to outputs (such as "increased capacity") and then to outcomes. Additionally, it is noted that "This evaluation did not cover the second level of outcomes and the impact level as the scope and focus of the assessment is at the level of output and outcomes which are short and medium-term changes." Outcomes and outputs are presented in detail in the report, cause and effect links between an intervention and its end results are discussed in Chapter 4 "Analysis and Findings of Country Programme" and Annex 2 "Evaluation Matrix." However, from the narrative description of the outcomes and outputs the causal connection is not always clear. In the following example there is a clear outcome, but how the outputs led to it is not discussed in detail: "…there is a tendency of increased enrolment of girls in schools due to the back to school campaign which is championed with the UNFPA supported interventions under the JPGE and JPAG projects" (pp. xxxiii).
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis shows different outcomes for different target groups: women and men, midwifes, MDSR committees, and health workers (pp. 31-34).
7. Is the analysis presented against contextual factors?	Yes	The analysis is presented against contextual factors: global, political (national needs and national standards – p. 27), social (for instance, fears of increase in workload of service providers and provider's attitude towards service integration – p. 36), and others.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The analysis elaborates effectively on cross-cutting issues such as equity and vulnerability (e.g., meeting human rights standards for quality of care and equity), gender equality (e.g., strengthening legislative framework for gender equality and equity) and human rights (e.g., increased national capacity to conduct evidence based advocacy for incorporating adolescents and youth and their human rights/needs in national laws, policies, programmes including humanitarian settings).

5. Conclusions	Yes No Partial	Assessment Level:	**Good**
To assess the validity of conclusions			
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions flow clearly from the findings and are organized in three clusters: strategic, programmatic and programme design and management levels. Programmatic conclusions relate to the evaluation criteria. A minor issue, however, is the design/presentation of the section: conclusions are written as narrative description and are not numbered. There are no direct references to the Findings section. This makes it time-consuming to see how the conclusions flow clearly from the findings.	

2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Partial	From the conclusions it is possible to understand the underlying issues of the programme: causal relationship between different factors (program implementation, human, social, political, etc.) that affect the achievement of results. However, some of the links are not always presented clearly. For instance, the evaluators explain that "expected population changes or effects of the various interventions remain unclear as surveys at measuring the indicators have not been conducted," but do not comment on why "…there is low motivation of health workers to attend to fistula victims, and the role of men in gender-based violence prevention is not clearly implemented" (p. 64).
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	In describing the process, it is clear that the evaluators considered comments on conclusions, but maintained independence/their own point of view while considering their uptake.

6. Recommendations	Yes No Partial	Assessment Level:	**Very good**
To ensure the usefulness and clarity of recommendations			
1. Do recommendations flow logically from conclusions?	Yes	Recommendations flow logically from conclusions. In each case, the conclusions on which the recommendation is based are shown.	
2. Are the recommendations clearly written, targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	In general, the recommendations are clearly written and easy to understand. For instance: "A clear-cut exit strategy should be integrated in all the activities" (p. 71). The recommendations are targeted at the intended users, for instance: "Target Level: UNFPA CO, Coordination Ministry." There is, in each case, a section indicating what would have to be done to implement the recommendation. Some recommendations include information on their human, financial and technical implications like "staff training on integration of population issues" (p. 71), "…CO invest in providing technical capacity-building for its IPs in M&E mechanism" (p. 72). But, in some cases the recommendations are not detailed enough like "investing in capacity-building, strengthening structures to be used in CP interventions. Joint resource mobilisation should also be embarked upon."	
3. Do recommendations appear balanced and impartial?	Yes	Recommendations appear balanced and impartial as they address relevant conclusions and supported by evidence.	
4. Is a timeframe for implementation proposed?	Yes	Most of the recommendations are directed to the next country programme both in terms of design and implementation. For example, "The 8th CP should also encourage cooperation among the IPs that work with different populations to achieve greater impact" (p. 70), but there are no more details about the timeframe.	
5. Are the recommendations prioritised and clearly presented to facilitate appropriate management response and follow up on each specific recommendation?	Yes	While the recommendations are not formally prioritized, in that they all have "High" priority, they were sufficiently specific to make a management response possible.	

7. Gender	0 1 2 3	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?	2	While indicators are designed in a way that ensures GEEW-related data to be collected, for instance: "Significant changes in marginalised population's i.e. poor women in both rural and urban settings…(p. xxx)", there is no GEEW-specific objective among the evaluation objectives/scope of analysis (p. 1). That is, the evaluation does not state that it will assess the extent to which UNFPA programme incorporated GEEW objectives and / or GEEW mainstreaming prinicples in its design or implementation, as well as how/to what extent GEEW results have been achieved.
2. Do evaluation criteria and evaluation questions specifically address how GEEW has been integrated into design, planning, implementation of the intervention and the results achieved?	2	Evaluation criteria specifically address GEEW "effectiveness (the extent the interventions supported by UNFPA in all programmatic areas contributed or are likely to contribute to the achievement of planned results; extent the programme integrated gender and rights-based approaches)" (p. 3). There is an EQ1 which includes GEEW dimension (i) To what extent is the 7th Country Programme responded (addressed) the country's needs, national priorities, internationally agreed commitments on … gender equality including GBV" (p. xxv), but other formulations do not contain GEEW.
3. Have gender-responsive evaluation methodology, methods and tools, and data analysis techniques been selected?	2	Even though the evaluation methodology does not specifcally include a description of any specific gender-responsive methods and tools, and data analysis techniques, gender-responsive data was collected and presented in the findings section and in the list of people interviewed.
4. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?	3	The evaluation findings reflect GEEW as a cross-cutting issue and a Gender and Youth program component. Both strategic and programmatic evaluation conclusions reflect a gender analysis, for instance: "It is a leading UN agency in ...gender equality and women's empowerment" (p. 63); "It responds to existing needs in the country in terms of gender equality and women's empowerment" (p. 64), "This component contributed to improving the policy and legislative framework as well as strengthened capacity for gender-based violence prevention and response" (p. 64). While most of the recommendations do not specifically mention gender, the first and most important one states:"Although it is important to ensure that the next CP is in line with the ICPD PoA, SDGs and Agenda 2063, it is even more important to ensure that the support provided to communities and beneficiaries on the ground addresses the specific SRH and gender equality needs of those communities."

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool, see Annex 7. 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).

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			

	Very good	Good	Fair	Unsatisfactory
5. Conclusions (11)		11		
6. Recommendations (11)	11			
7. Integration of gender (7)		7		
Total scoring points	51	42	7	
Overall assessment level of evaluation report	Very Good			
	Very good very confident to use	**Good** confident to use	**Fair** use with caution	**Unsatisfactory** not confident to use

(*) **(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?
The fair rating is based on unevenness in the evaluation design and in its applications in terms of findings. While the weaknesses of the methods and findings need to be considered, the fact that the conclusions are strong and clearly expressed, suggests that the evaluation can be used to help design the next country program.

• What aspects to be cautious about?

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

Consideration of significant constraints

The quality of this evaluation report has been hampered by exceptionally difficult circumstances: ☐ Yes   ☑ No

If yes, please explain: