Organizational unit:	Syria CO		Year of report:	2020

Title of evaluation report: UNFPA CPE: SYRIA 8th COUNTRY PROGRAMME 2016-2018

Overall quality of report:	Very good		Date of assessment:	26 July 2021

Overall comments: This evaluation provides a good example of how to overcome potential methodological limitations inherent of complex and conflict-affected environments to deliver a comprehensive evaluation with robust findings. The evaluation consults an extensive group of stakeholders, both in terms of diversity and size, whose voices and perspectives are analyzed, triangulated and presented in the report. The analysis clearly elaborates on cross-cutting issues, presenting disaggregated data by stakeholders and various demographic factors to support a more in-depth gender and social inclusion analysis and an understanding of the extent to which outcomes are experienced differently across vulnerable groups in the Syrian context. The conclusions and recommendations flow from the findings, though appear to have been rushed in their construction, limiting their usability/action-orientation. Persons with disabilities were considered as a vulnerable population within the evaluation design and analysis, ensuring the evaluation assessed, through specific evaluation questions under relevance and coherence, whether PwDs diverse needs were considered in the humanitarian response and presenting disaggregated data on the CPE's effectiveness in reaching PwDs, where information is available. While disability was adequately mainstreamed into the findings analysis and conclusions, there was not a specific recommendation tied to disability inclusion (though PwD can be implied to be included as a 'vulnerable population' under recommendation three requiring more extensive data collection and disaggregated of data to better understand population trends and needs).

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:	**Very good**

To ensure the report is comprehensive and user-friendly

1. Is the report structured in a logical way? 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? Is there a clear distinction made between analysis/findings, conclusions, recommendations and lessons learned (where applicable)?	Partial	The report is largely structured in a logical way, though the formatting of multiple columns is sometimes inconsistent, breaking in different places, and therefore difficult for the reader to follow (for example, see differences in column breaks on page 11 and 12). There are also occasional grammatical errors and typos throughout the report, as well as poorly constructed and run-on sentences which make it sometimes difficult to read and understand. While the report is largely well-written, the mistakes are noticeable enough and suggest copy-editing was not done to the extent needed. Finally, some graphs are not clear (page 41, for example).
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	At 68 pages, the report is within the indicated maximum length for a CPE.
3. Do the annexes contain – at a minimum – the ToRs; a bibliography; a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys)?	Yes	The annexes contain all of the required documents in addition to a more extensive overview of programme activities.

Executive summary

4. Is an executive summary written as a stand-alone section, presenting the i) Purpose; ii) Objectives, scope and brief description of interventions; iii) intended audience; iv) Methodology; v) Main results; Vi) Conclusions and Recommendations?	Yes	The executive summary is written as a stand-alone section, including details about the Syria Country Programme Evaluation purpose, the intervention areas and context, as well as the target audience. Main results and conclusions are presented together and organized by evaluation criteria with a separate section on recommendations.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	The executive summary is four pages and therefore meets requirements.

2. Design and Methodology	Yes No Partial	Assessment Level:	**Very good**

To ensure that the evaluation is put within its context

1. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	Chapter two covers the country context, and constraints faced as a result of the context are described in the section on limitations.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The evaluation report describes the process and basis for developing a reconstructed theory of change and also includes a visual representation in Figure 1 on page 3.

To ensure a rigorous design and methodology

3. 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 established and the evaluation matrix clear and complete.
4. Are the tools for data collection described and their choice justified?	Yes	A description of data collection tools is provided, which includes ongoing desk review, key informant and group interviews, focus group discussions, a survey and site visits. The rationale for their selection and target stakeholders for response to the different tools is also detailed, with interviews focusing primarily on duty bearers and focus group discussions and the survey focused on rightsholders.

5. 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)?	Yes	A comprehensive stakeholder map is not provided, however the consultation process is clearly and extensively described in the descriptions of data collection methods and the data collection process. The consultation process included time for feedback from the CO and ERG across phases of the evaluation, including the development of recommendations.
6. Are the methods for analysis clearly described for all types of data?	Yes	Data analysis methods are described and include contribution and trend analysis for qualitative data and the generation of descriptive statistics with SPSS for quantitative data.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Partial	The methodological limitations are acknowledged, the effect on the evaluation described and clear and appropriate mitigation mechanisms applied. The effect of COVID-19 on the evaluation process is not mentioned within the limitations and could be better highlighted for the purpose of learning (especially considering the extra deliverable on providing guidance for evaluations in conflict/humanitarian settings).
8. Is the sampling strategy described?	Yes	The sampling strategy is described and included both purposive and random approaches.
9. Does the methodology enable the collection and analysis of disaggregated data?	Yes	The methodology enables the collection and analysis of sex disaggregated data.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	The data collection methods are comprehensive and appropriate for collecting both quantitative and qualitative data on cross-cutting issues of equity and vulnerability, disability inclusion, gender equality and human rights. It is specifically noted that focus group discussions were designed in a way to capture and present a balanced perspective from persons of different socioeconomic backgrounds, including women of reproductive age, adolescents and youth, refugees and internally displaced persons and persons with disabilities. Included in the evaluation matrix are assumptions to be tested which reflect cross-cutting issues.
3. Reliability of Data	Yes No Partial	Assessment Level: **Very good**
To ensure quality of data and robust data collection processes		
1. Did the evaluation triangulate data collected as appropriate?	Yes	There is clear evidence that triangulation has been applied as a validation technique.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Considering the quite significant constraints faced due to COVID-19 and Syria's humanitarian context, the evaluation did well to consult an extensive and diverse group of stakeholders through various means.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	The evaluators noted adherence to the standards and principles of evaluation at UNFPA, particularly utility, credibility, independence, impartiality, ethics, transparency, and human rights and gender equality. The application of these principles is captured in the detailed methodology section, including through a thorough description of the sampling strategy and process for obtaining consent for the survey, which ultimately saw few refusals to participate (115) and high response rates (above 80%).
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	Findings are clearly substantiated by multiple lines of evidence.
2. Is the basis for interpretations carefully described?	Partial	When provided, the basis for interpretations are carefully described. In some cases, however, no interpretation is provided though it would have been useful; for example, the survey results revealed varied participation levels in trainings amongst youth across geographic areas though no analysis on why this might be the case (except that it is only the view of the sample, p.36).
3. Is the analysis presented against the evaluation questions?	Yes	The analysis is presented against the evaluation questions and indicators.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	Cause and effect links between UNFPA interventions, outputs and expected outcomes (as defined in the UNFPA Strategic Plan and reconstructed Theory of Change) are well explained and justified using triangulated data from documents, interviews and surveys. Unintended outcomes are highlighted throughout the section on effectiveness.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The analysis includes comparisons of different outcomes for different target groups, for example, using the survey data to analyze the extent to which returnees and IDPs perceive their access to SRH services relative to permanent residents. This type of analysis is most evident and extensive in the section assessing 'Coverage'.
6. Is the analysis presented against contextual factors?	Yes	The analysis presents contextual factors facilitating or hindering results, such as COVID-19, budget constraints, or varying priorities between government, implementing partners and UNFPA, for example.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	The analysis clearly elaborates on cross-cutting issues, presenting disaggregated data by stakeholders and various demographic factors to support a more in-depth gender and social inclusion analysis.
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 directly linked to evaluation questions in the text.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?	Partial	The conclusions reflect a thorough understanding of the underlying issues of the programme and context, while still remaining sensitive to cross-cutting issues particularly related to gender and human rights. However, the findings are thorough (if not occasionally repetitive) and the conclusions do not extensively move the analysis beyond what is reflected in the findings.

3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of bias in the conclusions.
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.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	Recommendations are not clearly targeted towards intended users. While operational implications are provided, they are sometimes presented as an unstructured list of tasks without clear explanation of why that task falls under the specific recommendation, nor detail on who should be charged with taking the task on. For example, recommendation 5 is focused on strengthening governance and state building, though there is reference to disaggregated monitoring data which might be better suited under recommendation 7 or 12 (or at least requires further explanation to understand its relevance to recommendation 5).
3. Do recommendations appear balanced and impartial and address, as relevant, key cross cutting issues such as equity and vulnerability, disability-inclusion, gender equality and human rights?	Yes	Recommendations appear impartial and address relevant cross-cutting issues.
4. Are the recommendations prioritized?	Yes	Recommendations are prioritized though most of the 12 recommendations are considered "high priority" (10).
7. Gender	0 1 2 3 (**)	Assessment Level: **Very 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?	3	**a. Does the evaluation include an objective specific to assessment of human rights and gender equality considerations or was it mainstreamed in other objectives? (Score: 0-3)** The evaluation mainstreams human rights and gender equality into the evaluation objectives through its aim to provide an assessment of geographic and demographic coverage. (3) **b. Was a standalone criterion on gender and/or human rights included in the evaluation framework or mainstreamed into other evaluation criteria? (Score: 0-3)** There is not a standalone criterion on gender or human rights, though it is mainstreamed into the evaluation framework through assumptions. (2) **c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3)** GEEW is integrated into evaluation sub-questions (3). **d. Does the evaluation assess whether sufficient information was collected during the implementation period on specific result indicators to measure progress on human rights and gender equality results ?(Score: 0-3)** Yes, this is covered in multiple sections including on context, CP background and the findings. (3)
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?	3	**a. 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 clearly integrate gender considerations and ensure disaggregated data is collected, which is further evidenced in the findings. (3) **b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analyzing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3)** The evaluation employs a mixed-methods approach and a large sample represented by diverse stakeholders. (3) **c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3)** Data sources are consistently triangulated to demonstrate consideration of inclusion, accuracy and credibility. (3) **d. Do the evaluation methods and sampling frame address the diversity of stakeholders affected by the intervention, particularly the most vulnerable, where appropriate? (Score: 0-3)** Yes, the sample includes a diverse range of stakeholders, including the most vulnerable who were primarily consulted through surveys and focus group discussions. (3) **e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3)** Yes, ethical procedures, including obtainment of informed consents, were sufficiently described in the methodology section and further reflected in the annexed data collection tools. (3)

3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?		**a. Does the evaluation have a background section that includes an intersectional analysis of the specific social groups affected by the issue or spell out the relevant normative instruments or policies related to human rights and gender equality? (Score: 0-3)** Chapter two provides an intersectional analysis as best possible based on available data. (2) **b. Do the findings include data analysis that explicitly and transparently triangulates the voices of different social role groups, and/or disaggregates quantitative data, where applicable? (Score: 0-3)** Findings clearly triangulate the voices of different social groups, especially within the section on coverage. (3) **c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3)** Unanticipated effects are described, including barriers to access based on social norms and other constraints. A distinct section on unintended results could have drawn these out more clearly. (2) **d. Does the evaluation report provide specific recommendations addressing GEEW issues, and priorities for action to improve GEEW or the intervention or future initiatives in this area? (Score: 0-3)** There are several recommendations which specifically address GEEW issues, though recommendations could be made more clear and actionable. (2)

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted.

(**) Scoring uses a four point scale (0-3).

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			
5. Conclusions (11)		11		
6. Recommendations (11)		11		
7. Integration of gender (7)	7			
Total scoring points	78	22		
Overall assessment level of evaluation report	Very good			