

Organizational unit: UNFPA Evaluation Office **Year of report:** 2020

Title of evaluation report: Evaluation of the UNFPA support to the HIV response (2016-2019)

Overall quality of report: Good **Date of assessment:** 2020 AUG

Overall comments: This report provides a comprehensive evaluation of UNFPA global support to the HIV response in five country offices (Turkey, Zambia, Georgia, Namibia, and Indonesia) and two regions in Eastern Europe and Central Asia (EECA) and East and Southern Africa (ESA) covering the period from 2016 (when the current UNAIDS strategy was rolled out) to 2019. The strengths of the evaluation include a well-organized report structure and clarity of reporting, as well as good analysis and integration of gender. The evaluation uses a mixed-methods and case study approach linked to an extraordinarily elaborate and long evaluation matrix and a well-constructed theory of change. Findings are consistently presented in the context of these models and key intervention activities, assumptions, and indicators. The key areas for improvement relate primarily to various aspects of the evaluation design and methodology. Missing elements include information on intended audience, stakeholder mapping, and a description of the consultation process. The report provides findings that reflect a reasonable analysis of the intervention's strengths and weaknesses contrasted against contextual factors and performance indicators, and a clear gender and vulnerability analysis. The conclusions provide a balanced perspective and are grounded on the evaluation findings with very good integration of contextual factors. The recommendations appear useful for informing the development of the next phase of the 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:	Very good
--	----------------------	-------------------	-----------

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	This report is accessible and easy to understand.
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	This report is within the page limits of 80 pages for a thematic evaluation.
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 structure is coherent and the sections are clearly delineated. The use of pictures, textboxes, maps, glossary of terms, and diagrams contributed to the clarity and thoughtful organization of the report.
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?	Yes	The annexes contains all the required information.

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	The executive summary is written as a standalone section.
6. Is there a clear structure of the executive summary, (i.e. i) Purpose; ii) Objectives and brief description of intervention; iii) Methodology; iv) Main conclusions; v) Recommendations)?	Yes	The executive summary is clearly structured and it includes all the required sections. The main sections are capitalized and the key points in the conclusion and recommendation are numbered.
7. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	At three pages, the executive summary is reasonably concise and well within the required five-page limit.
2. Design and Methodology		
	Yes No Partial	Assessment Level: Fair
<i>To ensure that the evaluation is put within its context</i>		
1. Does the evaluation describe the target audience for the evaluation?	Partial	The evaluation purpose is clearly stated in the introduction and the executive summary, but the report does not describe the target audience for this evaluation.
2. Is the development and institutional context of the evaluation clearly described and constraints explained?	Yes	The global context for UNFPA support to the HIV response, the global initiatives and targets, and constraints are described in sections 2.1 and 2.2.
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 report provides a brief assessment of the reconstructed intervention theory of change for UNFPA support with an elaborate graphic illustration presented in the methodology section.
<i>To ensure a rigorous design and methodology</i>		
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	A detailed 175 page long evaluation matrix with all necessary components and data points is presented in the annex. There is a clear description of the data collection and analysis methods in the main report.
5. Are the tools for data collection described and their choice justified?	Yes	The use of case studies with supplemental data collection methods is described and justified and the rationale for the selection of the two regional and five country case studies is provided in Table 6. The report provides the interview and online survey protocols separately in the annex. The case studies are provided as separate documents.
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	Although reference is made in the acknowledgement section about contributions from the evaluation reference group and validation of the case study findings by the local stakeholders at the country level, the report does not provide a comprehensive stakeholder map or a description of the consultation process.
7. Are the methods for analysis clearly described for all types of data?	Yes	The report explicitly describes the data analysis strategy.
8. Are methodological limitations acknowledged and their effect on the evaluation described? (Does the report discuss how any bias has been overcome?)	Yes	Three key methodological limitations and their mitigation strategies are described and summarized in Table 7 of the main report. The report acknowledges that the effect of the limitations on the evaluation is minimal due to the mitigation strategies.
9. Is the sampling strategy described?	Partial	The report does not describe the sampling strategy although the rationale for selecting countries is given.
10. Does the methodology enable the collection and analysis of disaggregated data?	Partial	The use of mixed methods approach enabled the collection and analysis although the list of persons consulted is not disaggregated by gender, and the interview and survey protocols do not capture gender of respondent.
11. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, gender equality and human rights)?	Yes	The theory-based evaluation design and mixed methods approach were appropriate in assessing cross-cutting issues. This is evidenced by the findings discussed in section 4.2 on the needs and rights of key populations and the most vulnerable.

3. Reliability of Data	Yes No Partial	Assessment Level:	Good
<i>To ensure quality of data and robust data collection processes</i>			
1. Did the evaluation triangulate data collected as appropriate?	Yes	The triangulation of data collected is demonstrated by the detailed sources of evidence in both the evaluation matrix and footnotes throughout the report.	
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	The evaluators noted that documents were the primary source of quantitative data and were supplemented with the data from the online survey, while qualitative data was drawn from interviews.	
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 evaluators discuss the diverse programme contexts, the lack of an explicit theory of change, and the challenges of attribution as limitations to the generalization of findings and conclusions. The mitigation strategies for each, including the reconstruction of the ToC, are explained.	
4. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	The methodology section does not specifically discuss how the evaluators' approach to data collection was sensitive to ethical considerations (i.e., how confidentiality and informed consent were ensured, whether UNEG standards were adhered to), nor are these mentioned in the data collection protocols. However, the evaluation approach does allow for the participation of a full range of stakeholders at the field level and preliminary evaluation findings were also presented to local stakeholders for discussion, comment and validation at the end of the field mission in each country - suggesting a consultative process that was open and transparent.	
4. Analysis and Findings	Yes No Partial	Assessment Level:	Good
<i>To ensure sound analysis and credible findings</i>			
1. Are the findings substantiated by evidence?	Yes	The report provides sources of evidence in the text, textboxes, 128 footnotes, and summaries in the evaluation matrix referencing documents, key informant interviews, and observations to substantiate findings. This report is highly evidence-based.	
2. Is the basis for interpretations carefully described?	Yes	The evaluators describe the basis for interpretations of qualitative and quantitative data as indicators, assumptions, and relevant evaluation criteria presented in the evaluation matrix.	
3. Is the analysis presented against the evaluation questions?	Yes	The report provides a detailed analytical framework in the evaluation matrix presenting analysis by evaluation questions, assumptions, evaluation criteria, and context.	
4. Is the analysis transparent about the sources and quality of data?	Partial	The evaluators are transparent about the sources of data in text, footnotes, and summaries in the evaluation matrix. They are not explicit about the quality of data.	
5. Are cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The evaluators used the theory of change model to demonstrate the cause-to-effect links between UNFPA contribution to the results of the intervention and the unintended outcomes noted in different activity focus areas in the findings section. The evaluators noted key contextual and strategic challenges to contribution and attribution analysis, such as a lack of well-defined programme of UNFPA support to the HIV response and diverse national and regional contexts.	
6. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The evaluation findings are presented at the key activity level showing relevant outcomes for different target groups.	

7. Is the analysis presented against contextual factors?	Yes	The findings are presented against relevant contextual factors at the national, regional, and global levels showing both the policy and programmatic factors affecting the intervention.
8. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, gender equality and human rights?	Yes	The report provides specific sub-sections to elaborate on cross-cutting issues such as equity, vulnerability, and gender and human rights issues in the relevant findings sections.
5. Conclusions	Yes No Partial	Assessment Level: Very good
<i>To assess the validity of conclusions</i>		
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions flow clearly from the findings and are organized around the same activities on which the findings were presented with clear visual presentations in Figure 9. The report references the findings, evaluation question, and relevant sections for each conclusion.
2. Do the conclusions go beyond the findings and provide a thorough understanding of the underlying issues of the programme/initiative/system being evaluated?	Yes	The evaluation conclusions are well-rounded on both the strengths and weaknesses of UNFPA interventions demonstrating the evaluators' understanding of the underlying contextual and programmatic issues.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of bias. The conclusions are clearly based on evidence from the findings showing both positive and negative elements and relevant sections of the findings.
6. Recommendations	Yes No Partial	Assessment Level: Very good
<i>To ensure the usefulness and clarity of recommendations</i>		
1. Do recommendations flow logically from conclusions?	Yes	The recommendations flow logically from the conclusions, and each recommendation notes the conclusions on which it is based.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	The recommendations clearly specify to whom they are directed, provide operational requirements, and the conclusions based on.
3. Do recommendations appear balanced and impartial?	Yes	As noted above, the recommendations flow clearly from the findings and do not show any partiality or bias.
4. Is a timeframe for implementation proposed?	Yes	Where possible, the recommendations do try to address the time frame in the operational requirements; e.g. rec 3 "This includes defining the UNFPA role at country level in support of programming, in particular throughout the new three-year Global Fund cycle". Reference is also made to ensuring synchronicity with the UNFPA strategic plan for 2018-2021.
5. Are the recommendations prioritized?	Yes	Although not explicitly prioritized, the recommendations appear to be in priority order for implementation with a level of urgency noted in recommendation 1.

7. Gender	0 1 2 3 (**)	Assessment Level:	Good
<i>To assess the integration of Gender Equality and Empowerment of Women (GEEW) (*)</i>			
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	<p>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) GEEW is considered as a cross-cutting theme both within the evaluation scope and objectives. (Score = 3).</p> <p>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 no standalone criterion on GEEW, but as mentioned, it is included as a cross-cutting theme to be addressed, and the assumptions and indicators for EQ1 and EQ3 assesses how GEEW is implemented in line with the priorities set by the international and national policy frameworks and aligned with the UNFPA policies and strategies. (Score = 3).</p> <p>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 was integrated into questions under the Relevance, Effectiveness, and Efficiency criterion. (Score = 3).</p> <p>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) Assessment of the sufficiency of information collected during the implementation period on specific results on GEEW is not specifically addressed, though there is some analysis provided within the relevance section. (Score = 1).</p>	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?	2	<p>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) The evaluation methodology in the main report and the case studies documents do not specify how gender issues are addressed although the findings section is gender responsive. (Score = 1).</p> <p>b. Does the evaluation methodology employ a mixed-methods approach, appropriate to evaluating GEEW considerations (collecting and analysing both quantitative and qualitative data, and ensuring the appropriate sample size)? (Score: 0-3) The methods and data collection tools include key informant interviews and observations but the stakeholders consulted are not disaggregated by sex. There is no information on how the sample size was drawn. (Score = 2).</p> <p>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 were diverse but it is not clearly stated how the triangulation and validation processes ensured inclusion and gender equality. (Score = 2)</p> <p>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) It is not clearly stated what sampling frame was used to select interview participants although the criteria for selecting countries and regions were stated. (Score = 2)</p> <p>e. Were ethical standards considered throughout the evaluation and were all stakeholder groups treated with integrity and respect for confidentiality? (Score: 0-3) Ethical considerations were not addressed. The FGD and interview protocols did not include directions about informed consent. (Score = 1)</p>	

<p>3. Do the evaluation findings, conclusions and recommendations reflect a gender analysis?</p>		<p>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) Gender analysis is evident in the background section, clearly showing an intersectional analysis of how different social groups are affected by policies, practices, and social norms under the global initiatives and targets such as the MDGs, SDGs, and the UNFPA Family Planning Strategy 2012-2020. (Score = 3).</p> <p>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) While the findings include data analysis that explicitly triangulates the sources of data, there is minimal use of quotes or stories to highlight the voices of different groups or clear gender-disaggregated data. Most quotes are from document sources (Score = 2).</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) There is anecdotal evidence of unanticipated effects of the intervention on human rights and gender equality. This could be clearly reported to improve the report clarity on unintended consequences of the intervention (Score = 1).</p> <p>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) Recommendation #2 addresses gender equality in the broader categorization of vulnerability in terms of meeting the needs of those left behind. (Score = 3).</p>
--	--	--

(*) This assessment criteria is fully based on the UN-SWAP Scoring Tool. Each sub-criteria shall be equally weighted (in correlation with the calculation in the tool and totaling the scores 11-12 = very good, 8-10 = good, 4-7 = Fair, 0-3=unsatisfactory).

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

0 = Not at all integrated. Applies when none of the elements under a criterion are met.

1 = Partially integrated. Applies when some minimal elements are met but further progress is needed and remedial action to meet the standard is required.

2 = Satisfactorily integrated. Applies when a satisfactory level has been reached and many of the elements are met but still improvement could be done.

3 = Fully integrated. Applies when all of the elements under a criterion are met, used and fully integrated in the evaluation and no remedial action is required.

Overall Evaluation Quality Assessment

	Assessment Levels (*)			
Quality assessment criteria (scoring points*)	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	29	58	13	
Overall assessment level of evaluation report		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?

- 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: