

Organizational unit: Jordan **Year of report:** 2022

Title of evaluation report: Evaluation of the Government of Jordan/United Nations Population Fund 9th Country Programme (2018-2022)

Overall quality of report: Very good **Date of assessment:** 25 June 2022

Overall comments: This is a very solid evaluation report that can be used by decision-makers with high confidence. Within the constraints of having to conduct data collection remotely due to pandemic restrictions, evaluators were able to gather input from a diverse range of stakeholders, including rights holders. The findings are presented by criteria and follow the logic of evaluation questions, providing an appropriate structure for the intended audience. They provide a well-evidenced and balanced picture of the programme's strengths and weaknesses, and data gaps are frequently identified. The perspectives of different stakeholders are apparent, particularly through the inclusion of participant quotes. The evaluation provides concise conclusions that are linked to each evaluation question. Recommendations include cross-cutting issues that address the most-at-risk populations and vulnerable communities, and appear useful for preparing the next programme cycle. There are few shortcomings. One is that a more comprehensive stakeholder map could have been provided. Evaluators could also have been more clear about the make-up of the rightsholders that participated in the process, and provided a gender disaggregated list of all evaluation participants.

The evaluation is disability inclusive. The background section notes that approximately 11 - 15% of the country's population have a disability. There is then a thorough analysis of the extent to which the CP addressed persons with disabilities (PWD) with this issue being discussed under EQ3 (effectiveness) and EQ9 (coverage), as well as in conclusions and recommendations.

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:	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	Overall, the report is logically structured, well written and generally free from grammatical or spelling errors. There is also a clear distinction between each section. However, the report does not include page numbers which makes it more challenging to navigate.
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, including the Executive Summary, is 58 pages which is within the expected 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 report contains all required annexes - the ToRs; documents consulted; a list of interviewees; the evaluation matrix; final stakeholder map and methodological and data collection tools used (interview guides, focus group notes, outline of surveys).

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 reads as a standalone document and includes all of the required sub-sections.
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Partial	The executive summary is a few lines over 5 pages in length, but would have been longer if there was consistent spacing between paragraphs and if same sized font was used throughout. Finding and conclusions could each be more briefly presented or combined to reduce overlap and increase readability.

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	The context section clearly describes the political and socio-economic situation in Jordan. In addition, it provides a good overview of the role of external assistance in the country.
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Yes	The report clearly outlines the outcomes and outputs of the UNFPA Programme in Jordan (2018-2022). In addition, Figure 2 shows the interventions logic which is then assessed in chapter 4 in subsection 2.5. The evaluators identified a need to focus on the stand-alone pillar on Young People, which would reflect a more logical chain of how the programme inputs and outputs achieve the results. According to the report, this would have also presented a structured alignment with the UNFPA strategic plan with its four interlinked outcomes.

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 matrix is described in the text and also included as an annex. It includes evaluation questions by criterion, assumptions, indicators and appropriate data sources, and methods for data collection aligned with the indicators. The key findings are detailed for each criterion. The evaluators also noted the changes made to the evaluation questions from what was proposed in the ToR.
--	-----	---

4. Are the tools for data collection described and their choice justified?	Yes	The methods/tools for data collection are briefly described and the rationale for their choice explained. These included key informant interviews (KIIs), focus group discussions (FGDs), and document review.
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)?	Partial	A basic stakeholder map is provided in Annex 5, which just lists institutional stakeholders and beneficiaries by CP output. The map would be more complete if details such as the roles and geographic areas were provided. In respect to stakeholder consultation, it is noted that an ERG was formed and provided input throughout the process. The introduction to the recommendations states that these were discussed with the CO and ERG. FGDs were held with rights holders although the evaluation team noted that "it was challenging to identify the direct beneficiaries of the interventions".
6. Are the methods for analysis clearly described for all types of data?	Yes	The evaluation report notes that narrative analysis and contribution analysis were the primary methods of qualitative data analysis and that descriptive statistical methods were used for quantitative data. Each is briefly explained.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	The methodological limitations are noted and what was done to minimize the limitations briefly described.
8. Is the sampling strategy described?	Yes	The sampling strategy is explained as being based on a 'participatory approach', that it followed the UNFPA Handbook and was illustrative, not statistically representative. The resultant sample is provided in a table that lists the interviewees by outcome area and institution. The number of beneficiaries participating in each FGD are also listed by outcome area and implementing partner.
9. Does the methodology enable the collection and analysis of disaggregated data?	Partial	Data collection tools did not ask for disaggregated data, either sex, and/or disability status in the tools for key informant interviews and focus groups. The stakeholder/beneficiary list is not disaggregated by sex or disability status. The findings, however, suggest that sex-disaggregated data was collected since quotes from both males and females stakeholders/beneficiaries are mentioned.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)?	Yes	Evaluation questions 1, 4 and 9 assess the integration of cross-cutting issues (the needs of diverse populations including the needs of marginalized and vulnerable groups, i.e. women and adolescents and youth with disabilities; those of racial, ethnic, religious and national minorities; and LGBTQI populations, among others. Covid restrictions made it difficult to engage the range of beneficiaries in the evaluation process and this is explained.
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 evaluation findings triangulate data across sources/ methods. Multiple data sources are identified for each evaluation sub-question and the evaluation matrix shows clear intent to collect triangulated data for each evaluation question.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Partial	Most of the data sources are identified (interviewees are specified by stakeholder type and institution) and the reliability of sources is briefly discussed in the methodology section. Based on Table 3, 51 KIIs took place and 70 beneficiaries were either interviewed or participated in FGDs. However, evaluators could have provided more details on the beneficiaries as there is no distinction between direct and indirect, gender, age, etc. It is mentioned that "The target beneficiaries for the FGD sessions included women, adolescents, youth, men, most-at-risk populations (MARPs) and refugees.", but it is not clear whether the eventual sample did include these groups.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Yes	There is a subsection on Ethical Considerations which discusses the UNFPA and UNEG guidance that was followed, including the UNEG Ethical Guidelines and Code of Conduct. There is also a sentence that mentions additional considerations were made in conducting interviews where data on GBV and disability issues was collected, and one that notes parental permission was obtained when adolescents below the age of 18 were involved. The data collection tools are included in Annex 5, and the introductory paragraphs make note of the ethical considerations followed.
4. Analysis and Findings		
	Yes No Partial	Assessment Level: Very good
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	The findings present a coherent and complete narrative that respond to each evaluation question with sufficient depth. Findings also provide a balanced picture of the programme's strengths and weaknesses. In addition, data gaps are frequently identified.
2. Is the basis for interpretations carefully described?	Yes	A strength of this evaluation is the thorough explanation of the findings. □
3. Is the analysis presented against the evaluation questions?	Yes	The findings are structured along the evaluation questions, making it easy for the reader to understand how these have been answered.

4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	The methodology notes the limits to drawing cause-and-effect links between the intervention and end results due to the evaluation design, however evaluators do effectively use the information available to show what was expected to be done, what was done in practice, and extent of achievement of planned results (both quantitative and qualitative). The methodology section indicates that contribution analysis was used to assess both intended and unintended outcomes, and to some extent the latter was evident in the organization's response to the pandemic.
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	The findings clearly discuss the extent to which the programme has reached women, adolescents and girls, as well as refugees and vulnerable Jordanians, and to an extent, people with disabilities.
6. Is the analysis presented against contextual factors?	Yes	The findings are shown against contextual factors, frequently through the use of text boxes to present the summary of challenges to implementation and achievement of indicators.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	The findings present a good analysis of cross-cutting issues, including gender equality and human rights (in particular sexual and reproductive rights). There is a solid analysis of the extent to which the CP addressed persons with disabilities (PWD).
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	Conclusions clearly emerge from findings. Each conclusion is directly linked to particular evaluation questions and evaluation criteria.
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?	Yes	The conclusions are analytical, go beyond the findings and demonstrate a strong understanding of the key issues underlying the country programme in Jordan. In addition, they adequately integrate cross-cutting issues, including those 'those furthest behind', such as the elderly, refugees of other nationalities, LGBTQ communities, migrant workers, survivors of human trafficking and commercial sex workers."
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	The conclusions convey the evaluators' unbiased judgement by presenting balanced and independent arguments.
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	Each recommendation is linked explicitly to its corresponding conclusion.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Yes	Each recommendation clearly identifies the users responsible for implementation. The recommendations also propose clear and useful actions and also discuss implications for their implementation.
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 are balanced and impartial and include cross-cutting issues. In particular, recommendations 9 and 10 mention focus on "Leaving no one behind" and "Addressing the root causes of GBV", respectively.
4. Are the recommendations prioritized?	Yes	The level of priority is identified for each recommendation (high or medium).
7. Gender	0 1 2 3 (**)	Assessment Level: Very 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?		<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) This was not evident in the objectives but the thematic scope directly addresses human rights and gender. (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) Gender and human rights were mainstreamed into multiple criteria. (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) There are dedicated evaluation questions and assumptions which address GEEW, specifically EQ 1, 4 and 9. (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) The evaluation assesses the results indicators. Out of 6 gender output indicators, 2 were overachieved and 3 were fully achieved. (3)</p>

<p>2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?</p>	<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 matrix includes indicators that are designed to collect disaggregated data. There is also a discussion on how the evaluation was gender-responsive. (3)</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 evaluation methodology used mixed methods appropriate for evaluating GEEW considerations, however the total sample was not gender disaggregated (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) There was sufficient diversity of stakeholders and methods involved to enable triangulation. (3)</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) The sample shows that a range of stakeholders were engaged. (3)</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 are described in general terms. Regarding vulnerable individuals, the report mentions that "the special needs around GBV, and disability-related work were considered, while ensuring confidentiality with adequate and informed consent," however it is not clear how these special needs were considered. (2)</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) The background section includes a solid gender and intersectional analysis. (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)</p> <p>The findings include the voices of different social groups. (3)</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) Unanticipated effects of the interventions on gender equality are addressed in respect to UNFPA being able to adapt emerging needs and priorities around the pandemic. (3)</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) The evaluation includes a recommendation (9) on addressing the root causes of GBV with actionable points for future implementation. (3)</p>
--	---

(*) 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	82	18		

Overall assessment level of evaluation report **Very good**

The evaluation integrates adequately cross cutting issues of gender equality, human rights and disability inclusion, even though is not included as part of the evaluation objective.

Consideration of significant constraints (e.g. COVID-19 or civil unrest)

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

If yes, please explain: