Organizational unit:	Pacific Sub-regional office	**Year of report:**	**2022**
Title of evaluation report:	The Pacific Sub-regional Programme Evaluation		
Overall quality of report:	**Good**	**Date of assessment:**	**7 January 2023**

Overall comments: This was a complex evaluation to undertake given the complexities of the multi-faceted programme and implementation environment. Nonetheless, the evaluators presented a well evidenced set of findings that showed the accomplishments and challenges faced by the SRP. Particularly notable aspects include a solid analysis of the covid response, the attention to climate change, and the extent to which the evaluation was disability inclusive. DI was evident in the context/background information, methodology (given that respondents included at least one organization representing persons with disabilities), in the analysis (the topic is specifically considered under three criteria with the perspectives of DI representatives being brought out), and in the conclusions. However, there were also several shortcomings in the report. One is the completeness of the methodology. Although the evaluators highlighted the participatory and inclusive approach to the evaluation and sampling, there is insufficient detail on the stakeholder groups, the sampling criteria, the final sample and the extent to which it was representative of stakeholder group, gender, location or other factors. It appears more of this information was provided in the inception report but it should also be included in the evaluation report. There is also no information on data analysis processes or on one of the methods of data collection - the case assessments and how these were used. The conclusions section was quite detailed and could have presented a higher-level analysis that went beyond the findings providing a clearer basis for the recommendations. Recommendations could also have been more clearly formulated in order to more easily facilitate a management response.

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 <u>assessment level</u> followed by main <u>comments</u>. (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	The report is well structured and the writing is clear. However, the report would have benefitted from close final proofreading - for example, there are several typos including in the recommendations section and the annex on results includes notes of people who are still to provide final data. In addition, more attention could be given to the presentation of visual aids as Figures 1 and 2 are not at a resolution that can be read.	
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 just under 60 pages.	
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	There is an extensive set of annexes that include that required elements and, among others, a more detailed explanation of the limitations of the evaluation, further explanation/justification of the evaluation approach, a glossary of relevant Conventions, a listing of ERG members, and further details on the three programme areas.	
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?	Partial	All elements are present. However, although the methodological approach is described and its inclusiveness is emphasized, the number of stakeholders that participated is not provided; this limits the readers' understanding of the scope and robustness of the evaluation.	
5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)?	Yes	This section includes considerable detail but is within the 5 page limit.	
2. Design and Methodology	Yes No Partial	Assessment Level:	**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	There is a solid overview of the development context which extends to GEWE and persons with disabilities. The challenges of climate change and covid are also addressed. The section on the UNFPA response includes a description of the outcome areas and indicates challenges and emergent issues.	
2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?	Partial	The is a discussion on how the ToC was used as part of a theory-based approach. Evaluators noted that the ToC formed the basis of the assumptions in the evaluation matrix, and indicated that it was 'extremely limited in content'. However, no further explanation is given on the shortcomings. Figure 1 provides a graphic depiction of the results framework, however as noted above it cannot easily be read.	
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 annexes include a complete evaluation matrix. The main findings for each criteria are succinctly presented in bullet point form along with an overall progress rating based on a 4-point scale. The matrix is also referenced in the main part of the report.	

4. Are the tools for data collection described and their choice justified?	Partial	Evaluators provide detailed justification for the chosen evaluation design. Tools included a literature review process, individual and group interview instruments, a FGD instrument, a case assessment instrument, and a tool to track progress against intended results. These are briefly described and justified. However, it would be useful to have more information on the case assessments and how these contributed to the methodological approach.
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	The challenges of conducting a comprehensive mapping process (and following the UNFPA recommended steps) for a multi-country programme are discussed and the approach eventually used is explained. Although the map was only included in the design report, the annexed list of stakeholders consulted appears to serve as a proxy. However, there could still be more clarity on the different groups of duty bearers and rights-holders and the extent to which each participated in the evaluation (for example, through a table of participants disaggregated by stakeholder group and other identifying factors). The description of the participatory approach to the evaluation indicates that the evaluation recommendations were discussed with stakeholders.
6. Are the methods for analysis clearly described for all types of data?	No	There is no specific discussion of data analysis processes.
7. Are methodological limitations acknowledged and their effect on the evaluation described? Does the report discuss what was done to minimize such issues?	Yes	Five limitations are discussed and presented in Annex O, with mitigation actions given where feasible. Limitations include the complexity of the multi-faceted programme and implementation environment.
8. Is the sampling strategy described?	No	The sampling process was purposive and is described in general terms in section 1.3.3. However, there is no mention of the sampling criteria or the universe from which the sample was drawn. The final sample can only be found in the annexed list of respondents but the final sample is not provided in the methodology section. As such the adequacy and representativeness of the sample is not readily apparent.
9. Does the methodology enable the collection and analysis of disaggregated data?	Yes	Annex P includes a brief discussion on data being disaggregated by gender, age, disability and vulnerability. Annex Q, which provides an overview of each programme, provides statistics related to persons with disabilities, youth, and percent of urban population, etc.
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 design included a participatory approach that was explicit about 'hearing unheard voices'. Indicators for evaluation questions cover cross-cutting issues including disability inclusion.
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	Triangulation is discussed and is reflected in findings
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	Sources appear appropriate and reliable, and include an extensive list of documents. There was heavy reliance on qualitative data in terms of primary sources with quantitative data being drawn from secondary sources.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations?	Partial	Adherence to UNEG Ethical Guidelines is noted in the methodology section and ethical considerations - including informed consent, confidentiality anonymity, following of covid protocols - are evident in the data collection protocols. However, there is not a specific discussion on ethical practices or of how the process accommodated the range of rightsholders in particular.
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 well supported by document sources, project monitoring data, references to interviews, and direct quotes from respondents.
2. Is the basis for interpretations carefully described?	Partial	Evaluators are careful to interpret and explain the data. However, they frequently present information from other reports and evaluations including those that have been done at the global level for UNFPA and the UN Pacific Strategy; and pull out quotes are often drawn from these secondary sources. While it is good practice to reference and place the object of the evaluation in the context of these other studies, this additional information sometimes obscures the findings related directly to this SRP evaluation.
3. Is the analysis presented against the evaluation questions?	Yes	There is one main evaluation question clearly shown for each criteria. The subsequent analysis addresses the question.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted?	Yes	There is adequate attention to linkages particularly in the Effectiveness discussion where evaluators also note where there is still insufficient data on outputs to show outcome-level achievements. The Effectiveness evaluation question includes a component on unexpected results. Although the subsequent analysis does not explicitly address this, there is a thorough discussion on the covid response and some of the particular achievements and challenges during that time, as well as of innovative work done.

5. Does the analysis show different outcomes for different target groups, as relevant?	Partial	The analysis has some reference to performance levels in different countries, however beyond this and the discussions on disability there is minimal attention to other target groups. Given that one of the conclusions of the evaluation is that there have been gains in the collection of increasingly disaggregated data by national statistical systems, it seems reasonable to expect there to be greater reflection of disaggregated data in this report.
6. Is the analysis presented against contextual factors?	Yes	Context is consistently provided. Examples in the discussion on Effectiveness include where the evaluators explain the significance of the CP exceeding the targeted number of health facilities making referrals to multisectoral services, and the explanation of how different approaches to rating systems give different pictures of performance.
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	Section 4.2.4 Vulnerability and Relevance looks specifically at disability inclusion and diverse sexual identities, as well as at climate change/environment. DI is well mainstreamed into the analysis as there is a also a specific subsection under Effectiveness on disability inclusion and it is discussed under Coordination.

5. Conclusions	Yes No Partial	Assessment Level:	**Good**

To assess the validity of conclusions

1. Do the conclusions flow clearly from the findings?	Yes	After a page of overall comments, conclusions are presented by criteria and clearly reflect the respective 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 and reflect as appropriate cross-cutting issues such as equality and vulnerability, disability inclusion, gender equality and human rights?	Partial	The conclusions are detailed and span 4.5 pages. Although they do present the accomplishments and challenges of the CP, some of the content is at a findings- and explanatory-level (an example of the latter is that the definition of each criteria is provided). At the same time, attention to cross-cutting themes is quite minimal.
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no indication of bias.

6. Recommendations	Yes No Partial	Assessment Level:	**Fair**

To ensure the usefulness and clarity of recommendations

1. Do recommendations flow logically from conclusions?	Partial	The links are not explicitly made to conclusions, and at times the relationship is not apparent. For example, there is a main recommendation on climate change adaptation - there is a relevant discussion about this in Findings but not in Conclusions. There is also a sub recommendation on dignity kits without a corresponding conclusion.
2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?	Partial	Most recommendation are directed to UNFPA in general, and a few to UNFPA PRSO. This distinction seems somewhat arbitrary and more clarity about the organizational level would be helpful. Operational guidance is provided in some cases but not all.
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	These appear to be impartial and reflect the strengths and weakness found through the evaluation process. There is some attention to cross-cutting themes, mainly gender, but they do not address LNOB or disability inclusion.
4. Are the recommendations prioritized?	Partial	They are prioritized to the extent that there are two groupings - 'main recommendations' and 'other recommendations'. However, the rational for the distinction could be more clear particularly as the latter includes high-level recommendations such as the need for a systematic review of the structure of the operations in the Pacific.

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	**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)** Although not covered in the objectives, the thematic scope includes GEWE as a cross-cutting theme. = 2 **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)** HRGE issues are mainstreamed into the criteria = 3 **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 was not a dedicated question or subquestion on GEEW. = 0 **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)** Data availability on gender and gaps in this regard are considered = 3

2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?		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) Evaluators note using the UNEG guidance on HRGE in evaluations but are not explicit about how the approach was gender-responsive. = 1
		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 mixed methods and participatory approach is appropriate for assessing GEEW, however the sampling process is not clear about the extent women's voices are incorporated. = 1
	2	c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3) Although sampling is not well defined, evaluators state that the participatory approach ensured inclusion. Both triangulation and validation were evident. = 2
		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) Although the annexed list of evaluation respondents shows the organizational affiliation of each and location, there is no cumulative presentation of the total number of evaluation participants or disaggregation by stakeholder group, gender or location. The list does indicate that there was representation from organizations focused on women, human rights, and persons with disabilities. = 2
		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 apparent but evaluators could have been more specific about their application to rightsholders = 2
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) The context section provides an appropriate analysis, including on GBV. = 3
	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) The findings include a reasonably solid analysis of GE, however although the voices of those representing disability inclusion actors are apparent, the voices of women are not. = 2
		c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) Unanticipated effects are only implicitly addressed in respect to the covid response and the delivery of services, including the recognition of the likely increase in GBV during this time. = 1
		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 recommendations related to GEEW. = 3

(*) 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	51	38	11	
Overall assessment level of evaluation report	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: