

**Organizational unit:** UNFPA Evaluation Office **Year of report:** 2023

**Title of evaluation report:** Evaluation of UNFPA Support to Population dynamics and data

**Overall quality of report:** **Very good** **Date of assessment:** 15 May 2023

**Overall comments:** This is the first thematic evaluation of UNFPA support to population dynamics and data with a specific focus on demographic dividend, population ageing and low fertility. The evaluators undertook a comprehensive examination of the topics including through the use of six in-depth case studies. The case studies were supplemented by data from 500+ documents, further KIs, an online survey with 122 respondents, and a scan of 8735 social media posts. As per good practice, the evaluation process also entailed re-formulation and validation of the theory of change, which was then also used to structure the evaluation matrix and analysis. The report further stands out for having well-evidenced and triangulated findings, for frequent use of clear visual aids, and for clarity in its presentation. In respect to cross-cutting themes, there is a solid assessment of gender and leave no one behind; disability issues are explored in the findings but are not explicitly discussed in the context, conclusions or recommendations. Evaluation participants included representatives of organizations serving persons with disabilities. The minor shortcomings of the evaluation include the executive summary and main report both being somewhat longer than the preferred length, the need for an explanation of how contribution analysis was applied, and insufficient clarity on the involvement of CSOs and academic/research institutions in the implementation of the programs.

**Assessment Levels**

<b>Very Good</b> (blue)	strong, above average, best practice	<b>Good</b> (green)	satisfactory, respectable	<b>Fair</b> (yellow)	with some weaknesses, still acceptable	<b>Unsatisfactory</b> (red)	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)*

<b>1. Structure and Clarity of Reporting</b>	Yes No Partial	Assessment Level:	<b>Good</b>
--	----------------------	-------------------	-------------

*To ensure the report is comprehensive and user-friendly*

<b>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)?</b>	Yes	The report is logically organized and professionally designed with well-defined sections and sub-sections. The evaluators have maximized the accessibility of the document through clear writing, good use of visual aids, and by providing succinct contextual information (such as Box 1: key terms and definitions).
---	-----	---

<b>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) Note: YES - the report is within the indicated maximum page length. PARTIAL - the report exceeds the maximum page limit by 1- 5 pages. NO - the report exceeds the maximum page limit</b>	Partial	The main part of the report covers 93 pages with 5 of these being full page photos. There is also fairly extensive use of white space and large visual aids. Taking these into account, the text still extends beyond the 80 page limit requested for thematic evaluations.
--	---------	---

<b>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)? Note: YES - the report contains all the annexes indicated. PARTIAL - if the report is missing the ToRs or the bibliography in the annexes. NO - if the report is missing any of the following annexes: a list of interviewees; the evaluation matrix; methodological and data collection tools used (e.g. interview guides; focus group notes, outline of surveys).</b>	Partial	There is an extensive set of annexes, however a missing piece is the interview protocols. Although these protocols were included in the Inception Report, the expectation is for them to be in the final evaluation report as well.
--	---------	---

*Executive summary*

<b>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?</b>	Yes	The Executive Summary presents all of the expected key components, including intended audience. Although the number of evaluation participants involved in the interviews and survey would have provided additional insight on the extent of stakeholder involvement, the more detailed description of case studies helps to establish the level of rigour of the methodological approach.
---	-----	--

<b>5. Is the executive summary reasonably concise (e.g. with a maximum length of 5 pages)? Note: YES - the executive summary is within the indicated maximum page limit. PARTIAL - the executive summary exceeds the maximum page limit by 1 to 2 pages. NO - the executive summary exceeds the maximum page limit beyond 2 pages.</b>	Partial	Although this section is written in a concise style, it does extend 1.5 pages beyond the preferred length of 5 pages. This is in part due to the recommendations including more detail than is normally presented in an Executive Summary.
--	---------	--

<b>2. Design and Methodology</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
----------------------------------	----------------------	-------------------	------------------

*To ensure that the evaluation is put within its context*

<b>1. Is the development and institutional context of the evaluation clearly described and constraints</b>	Yes	There is a useful and appropriate level of discussion on the background context and on UNFPA's response.
--	-----	--

<b>2. Does the evaluation report discuss and assess the intervention logic and/or theory of change?</b>	Yes	Evaluators reconstructed the ToC and note that it was initially used as a foundation tool for the development of the evaluation matrix, but later was validated and expanded upon via consultations with the Evaluation Reference Group which will give it broader utility to UNFPA programming. An explanation of how it was developed and the resultant ToC are provided in the annexes, with the ToC also being inserted in the main report. The presentation goes beyond the different levels of results to also include barriers.
---	-----	--

*To ensure a rigorous design and methodology*

<b>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?</b>	Yes	The framework is clearly set out in the evaluation matrix. The matrix includes questions, assumptions, linkages to the ToC, indicators, and methods and sources of information. Figure 2 provides a useful graphic of the overall evaluation design, evidence base, and stakeholder involvement. Although key findings are not provided in the matrix, they are shown in evidence tables for each of the case studies in Annex 3.
---	-----	---

<b>4. Are the tools for data collection described and their choice justified? Remember: Please address both aspects of this sub-criteria in the comment: 1) are data collection tools described (i.e. documentary review, interviews, focus group discussions etc.) and 2) is the rationale for their selection detailed</b>	Yes	The methodology relied heavily on six regional and country-level case studies. It is noted that this was supplemented by global and regional-level 'datasets' (500+ documents, 210 KIs, online survey with 122 respondents, and a scan of 8735 social media posts). The justification is sufficient.
--	-----	--

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)? <i>Remember: Please address all three aspects of this sub-criteria in the comment: 1) is a comprehensive stakeholder map included (in either the report itself or the annexes) 2) Is the overall stakeholder consultation process described and 3) within the consultation process were key stakeholders consulted on the recommendations specifically? 4) does the evaluation stakeholder mapping and data collection methods involve vulnerable and marginalized groups, including persons with disabilities and their representative organizations?</i>	Partial	The different stakeholder groups consulted are listed in the explanation of KIIs, and then are disaggregated by location and sex. In terms of stakeholder participation, there was a range of groups that provided input including representatives of organizations working with persons with disabilities. It is also apparent that the ERG was actively engaged at various stages of the process with their consultation on the recommendations being inferred in figure 2. However, there is not a comprehensive map provided that would, for example, clarify the different types of CSOs and other partners including research institutions that are involved in this work and the types of roles they play. Such an explanation would provide further clarity on the extent to which the sample was representative of vulnerable groups. (It appears that more mapping information was provided in the Inception Report, however the key points should also be included in the evaluation)
6. Are the methods for analysis clearly described for all types of data? <i>Remember: This sub-criteria is asking about the data analysis methods used and whether they are clearly described - was contribution analysis used, or qualitative comparative analysis, for example, or descriptive statistics? Triangulation is not a method of analysis; it is a validation technique.</i>	Partial	There was a participatory element to the analysis. Evaluators explain using excel databases for collating, coding and analysing each of the primary types of data. This was followed by a data analysis workshop that brought together the analysis of all datasets and set the stage for preliminary findings. The methodological approach would have been more robust, however, if the application of contribution analysis had been 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 main report and annexes include a chart with the anticipated limitations and mitigations measures, as well as the actual limitations and mitigations.
8. Is the sampling strategy described? <i>This criteria is asking whether the methodological approach to determining the sample of stakeholders consulted and the sample of site visits is described. Reviewers should examine whether the evaluation report includes information on how the universe was determined; the sampling approach used (i.e. purposive); the indicators used to develop the sample to be consulted (or visited); the resulting sample; and importantly limitations to the approach (including</i>	Yes	The eight sampling criteria for the selection of case studies are listed, and the targeting of individual respondents is sufficiently justified. It is noted that a more detailed explanation is provided in the Inception Report.
9. Does the methodology enable the collection and analysis of disaggregated data? <i>Remember: The default should be to disaggregate by sex. Whenever possible, this sub-criterion is also asking systematic disaggregation of data related to population groups (e.g. persons with disability) where there are implications related to UNFPA's portfolio/interventions for these</i>	Yes	Data is disaggregated by gender.
10. Is the design and methodology appropriate for assessing the cross-cutting issues (equity and vulnerability, disability inclusion, gender equality and human rights)? <i>Remember: This sub-criteria is asking about the evaluation methodology itself – specifically does the evaluation's design / methodology ensure that the evaluation is able to assess the extent to</i>	Yes	EQ 2 asks, "to what extent has UNFPA integrated cross-cutting issues: particularly gender, human rights, disability, leaving no one behind (LNOB) and emerging megatrends, within population dynamics and data programming and support?", and the assumptions for this question further probe each of these issues. There is also a dedicated sub-section of the methodology that explains how cross-cutting issues were
<b>3. Reliability of Data</b>		
	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To ensure quality of data and robust data collection processes</i>		
1. Did the evaluation triangulate data collected as appropriate?	Yes	The evaluation matrix shows multiple data sources for each question, and the data analysis process also clearly enabled triangulation. Triangulation is furthermore evident in the findings.
2. Did the evaluation clearly identify and make use of reliable qualitative and quantitative data sources?	Yes	All sources appear to be relevant and reliable. There is reference to reliability in the limitations and mitigation section.
3. Is there evidence that data has been collected with a sensitivity to issues of discrimination and other ethical considerations? <i>Remember: Ensure that reviewers are assessing the sub-criteria is concerned with whether there is</i>	Yes	Section 2.4 addresses ethical issues, including evaluators compliance with UNEG Ethical Guidelines and Code of Conduct. There is a particularly good treatment of data security processes for each source of data shown in Table 1.
<b>4. Analysis and Findings</b>		
	Yes No Partial	Assessment Level: <b>Very good</b>
<i>To ensure sound analysis and credible findings</i>		
1. Are the findings substantiated by evidence?	Yes	The evidence base for each main finding is effectively shown in footnotes (by stakeholder group, document source and/or method). Evaluators drew on both qualitative and quantitative sources.
2. Is the basis for interpretations carefully described?	Yes	Evaluators are careful to explain the basis for each key finding. An example is the discussion on aging and the different ways that the issue of low fertility rates is viewed and taken up, including by increasingly conservative pre-natalist policies (p 29).
3. Is the analysis presented against the evaluation questions?	Yes	This section is well organized with the main questions highlighted and the respective sub-questions provided in footnotes; key findings for each question are bolded.
4. Are the cause and effect links between an intervention and its end results explained and any unintended outcomes highlighted? <i>Remember: Please address both parts of the sub-criteria in the comment, namely: are the cause/effect links (between UNFPA contribution and outputs/contribution to results/outcomes) explained as well as the results and 2) are unintended outcomes discussed. On the latter, please note in the comment whether evaluators considered/looked for unintended outcomes and noted whether there were (or were not) any; or whether the report does not mention unintended outcomes.</i>	Yes	The use of contribution analysis helps to ensure causal linkages are assessed. Although there is not overt use of the ToC, examples of key achievements in each program area are highlighted and their contribution to end results are discussed. The evaluators also look to other mechanisms to establish causality for the different types of programming. For example, for assessing effectiveness in strengthening data systems, they propose use of data value chains (although the application of this is limited by the lack of a programme specific ToC). A gender marker tag is used to analyze financial expenditures in order to consider effectiveness in integrating gender priorities into population and development work. The extent of contribution of the different modes of engagement within each program area at country, regional and global levels is then graphically depicted in Table 3. Although there is not a specific question on unexpected outcomes, one of the findings addresses the effects of there being inconsistent understandings regarding
5. Does the analysis show different outcomes for different target groups, as relevant?	Yes	This is done in some cases, for example, on the limited amount of census data on LGBTQI+ populations (p 35) and migrant populations. As is appropriate for a global evaluation, the different outcomes of UNFPA's work in different countries is also brought out.
6. Is the analysis presented against contextual factors?	Yes	Context is consistently provided, including through footnotes. An example is in the discussion on migration where there is a footnote that discusses UNFPA pledges under the Global Compact on Refugees (p 36).
7. Does the analysis elaborate on cross-cutting issues such as equity and vulnerability, disability inclusion, gender equality and human rights?	Yes	There is a solid discussion of these themes, including of how HRBA, gender and issues of inclusion (particularly LNOB) are often conflated. Disability and ableism are also addressed with several examples provided of how disability has been reflected in different types of programming.

<b>5. Conclusions</b>	Yes No Partial	Assessment Level:	<b>Very good</b>
<i>To assess the validity of conclusions</i>			
1. Do the conclusions flow clearly from the findings?	Yes	The conclusions are based on the findings. To emphasize this, the respective evaluation question is identified for each main conclusion statement.	
<b>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</b>	Yes	This section is clearly articulated and provides a useful higher level of analysis than findings. The cross-cutting themes are evident although disability inclusion is not explicitly mentioned but appears to be covered under LNOB.	
3. Do the conclusions appear to convey the evaluators' unbiased judgement?	Yes	There is no evidence of bias, particularly as the basis for the evidence was strong.	
<b>6. Recommendations</b>			
Yes No Partial			
Assessment Level:			
<b>Very good</b>			
<i>To ensure the usefulness and clarity of recommendations</i>			
1. Do recommendations flow logically from conclusions?	Yes	The conclusions on which each recommendation is based are identified.	
<b>2. Are the recommendations targeted at the intended users and action-oriented (with information on their human, financial and technical implications)?</b>	Yes	The recommendations are very well formulated. The main recommendation is succinctly stated, and then is further explained and action-oriented sub-recommendations are provided. The target audience is also shown. There is one mention of financial implications; this is in respect to developing a resourcing and recruitment plan to address skillset gaps.	
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?	Partial	They appear balanced and impartial. Cross-cutting issues are reflected at the sub-recommendation level but are not featured prominently. Disability inclusion is not specifically addressed.	
4. Are the recommendations prioritized?	Yes	A priority level is assigned to each recommendation.	
<b>7. Gender</b>			
0 1 2 3 (**)			
Assessment Level:			
<b>Very good</b>			
<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?  <i>Remember: Ensure that the reviewer does not confuse the content of the country programme with the approach to the evaluation. This question is asking whether the evaluation criteria and questions are gender responsive and inclusion of GEWE dimensions in its scope of analysis (i.e. in the objectives for example) or the indicators the evaluation selects against which data will be collected so that the evaluation is able to assess whether the country programme is gender responsive.</i>  <i>A general note on UNFPA programming: While there may be evidence of gender being referred to as a cornerstone of UNFPA programming - in the sense that most UNFPA programmes target</i>	3	<p><b>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)</b> The objectives include reference to integrating a gender-responsive approach across all aspects of the evaluation. = 3</p> <p><b>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)</b> Gender is mainstreamed = 3</p> <p><b>c. Is there a dedicated evaluation question or sub-question regarding how GEEW was integrated into the subject of the evaluation? (Score: 0-3)</b> There is a dedicated question on gender and other x-cutting themes, with a gender lens applied to other questions as well. =3</p> <p><b>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)</b> As the program includes a major focus on developing accurate</p>	
2. Is a gender-responsive methodology used, including gender-responsive methods and tools, and data analysis techniques?  <i>Remember: Ensure that the reviewer does not confuse the content of the country programme with the approach to the evaluation. This sub-criteria is asking whether the evaluation criteria and evaluation questions (i.e. the evaluation itself) are gender responsive; in other words, are the criteria interpreted/operationalized and evaluations questions developed in a way that is able to capture whether (or not) gender equality/human rights/the empowerment of women has been integrated into UNFPA's country programmelsupport (in the design/planning, implementation and results)?</i>	3	<p><b>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)</b> There is a sub-section on integrating x-cutting issues. It notes the use of an intersectional approach, the application of a gender lens and gender-balance of respondents. = 3</p> <p><b>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)</b> A mixed-methods approach was used with both types of data being collected. The sampled was gender balanced. = 3</p> <p><b>c. Are a diverse range of data sources and processes employed (i.e. triangulation, validation) to guarantee inclusion, accuracy and credibility? (Score: 0-3)</b> Data sources were diverse; both triangulation and validation were evident. =3</p> <p><b>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)</b> The list of stakeholder consulted shows that people who worked with vulnerable groups were included. However, as noted above, an explanation of the CSOs involved in program implementation and in the evaluation process would have made it more clear whether there was adequate representation from the diversity of stakeholders. The absence of KII protocols also hampers the assessment of this criteria. = 1</p> <p><b>e. Were ethical standards considered throughout the evaluation and were all stakeholder</b></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 context section addresses issues of gender and population data, and references the ICPD framework (noted as a milestone in highlighting the rights of women and girls being central to development) amongst others. = 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>Although the findings are well evidenced, the perspectives of female vs male respondents are not articulated. = 0</p> <p>c. Are unanticipated effects of the intervention on human rights and gender equality described? (Score: 0-3) This was not apparent. = 0</p> <p>d. Does the evaluation report provide specific recommendations addressing GEEW issues,</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			
<b>Total scoring points</b>	<b>93</b>	<b>7</b>		
<b>Overall assessment level of evaluation report</b>	<b>Very good</b>			

**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.**  
 There is good treatment of gender and LNOB. Disability inclusion was covered in the evaluation design; it was included as a question and representatives of organizations working with persons with disabilities were included as evaluation participants. As noted above, disability and abism are addressed in the analysis with several examples provided of how disability has been reflected in different types of programming. However, this theme was not specifically reflected in conclusions or recommendations.

**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: