

UNFPA Evaluation Quality Assessment Grid

Version: March 2025 annotated

REPORT RATING SUMMARY		
Overall Rating	82%	Satisfactory
● ● ● ● ●	Excellent	5
● ● ● ●	Highly Satisfactory	4
● ● ● ● -	Satisfactory	3
● ● - - -	Fair	2
● - - - -	Unsatisfactory	1

REPORT DETAILS	
Title of the evaluation report	UNFPA Bosnia and Herzegovina 3rd Country Programme Country Programme Evaluation 2021-2025
Region	Eastern Europe & Central Asia
Country	Bosnia and Herzegovina
Year of report	2025
Business Unit/programme country (managing evaluation)	Bosnia and Herzegovina CO
Date of assessment review (dd/mm/yyyy)	5/13/2025
Name of assessment review firm	IOD PARC
CLASSIFICATION OF EVALUATION REPORT	
Primary SDG(s) covered (list provided below)	3,4,5,9,16,17
UNFPA Strategic Plan areas covered (lists provided below)	
Three transformative results	<ul style="list-style-type: none"> • an end to preventable maternal deaths • an end to the unmet need for family planning • an end to gender-based violence and all harmful practices, including female genital mutilation and child, early and forced marriage
Six outputs	(a) policy and accountability, (b) quality of care and services, (c) gender and norms, (d) data and evidence, including on population changes, (e) humanitarian action, and (f) adolescent and youth.
Six accelerators	<ul style="list-style-type: none"> • Human rights-based and gender transformative approaches • Innovation and digitalization • Partnership, South-South and triangular cooperation, and financing; • Data and evidence; • "Leaving no one behind" and "reaching the furthest left behind first"; • Resilience and adaptation
Organizational effectiveness and efficiency	Yes
Humanitarian evaluation	Yes
Evaluation evaluand (e.g. country programme/intervention/policy/thematic)	Country Programme Evaluation
Evaluation type (e.g. formative, summative, developmental)	Formative and summative
Geographic scope (e.g. global, regional, national)	National
EQA Summary: The rater will provide top line issues for this evaluation relevant for feedback to senior management (strengths and weaknesses), summarizing how the evaluation report meets or fails to meet all criteria. As relevant, the rater will highlight good practice/added value elements and the level of complexity of the evaluation. The rater should also highlight how cross-cutting issues were addressed in the report. Considerations of significant constraints (e.g. humanitarian crisis or political turmoil) should also be highlighted here.	
Strengths <ul style="list-style-type: none"> • The purpose, objectives and scope are clearly defined and described. • Evaluation criteria are clearly set out in Table 1 and aligned with evaluation questions. The evaluation matrix also contains all expected information for a UNFPA evaluation, including detailed evidence which is an example of good practice. • The background section clearly outlines the complex political context of BiH, as well social and economic factors, with a focus on UNFPAs mandate. With regard to cross-cutting issues, gender equality is well covered, with mention of disability and other marginalised groups. The key facts table includes detailed information on relevant SDGs. • Findings clearly and systematically address each evaluation question and present a range of evidence to triangulate findings and present evidence from a range of sources. They present both positive and negative evidence with clear reference to sources of data. there is some evidence of intersectional analysis and data sources are well referenced. • Conclusions are well written, relate directly to the findings and provide additional insight and analysis. They are linked to both the EQs to which they relate and the recommendations they inform • Recommendations are clear and directly relate to the conclusions, they also provide useful guidance for implementation. • Gender is well-considered throughout this evaluation with multiple evaluation questions relating to gender and gender equality. The methodology was designed to capture disaggregated data, and the findings do analyse some evidence with an intersectional lens, though as noted this was not always applied consistently. However, the evaluation shows a nuanced understanding of gender and has robustly assessed the CP with a gender lens, which is also reflected in conclusions and recommendations. 	
Weaknesses <ul style="list-style-type: none"> • The ToC/ logic framework is provided in the main body of the text and analysed narratively. However, the report would have greater utility if the ToC had been rcarefully reviewed and revised/improved. • The methodology included only two main types of data collection (Kilis with organisations and document review). While the reasons for this are well explained, it represent a missed opportunity to include the vices of rights holders. The sampling strategy could be improved if site visits were included, and greater clarity on which organisations represented the views of right sholders provided. • The data analysis is well explained, but could provide detail on software packages utilised and clarity on whether AI was used, while limitations could have been more detailed. • Ethical issues are noted in the report, and data collection tools provided in the Annexes. This section could be improved if more contextualised information was provided and consent sought that included the right for KII participants to withdraw at any point in the evaluation process. • Findings could have been improved if there was a stronger and systematic analysis of casual factors for each evaluation question. • Recommendations could be improved if there was a more detailed explanation of how they were developed and if the ERG was involved in a co-creation process (not only validation), as well as greater precision in identifying responsible parties that goes beyond naming the BiH CO as being responsible for all recommendations. • While there is significant analysis of marginalised and vulnerable people, including those with disabilities, migrants and others, and evidence of an intersectional lens being applied to the analysis of evidence in the findings section, this could have been improved were it noted in the methodology that an intersectional lens would be applied and enable it to be applied consistently. • Data is disaggregated where possible, noting the limitations of secondary sources. However a differentiated analysis of results across different types of partner would have been possible if Kilis had been differentiated as government partners, implementing partners or other types of partners to enable a situated and nuanced understanding of different types, and the people they represent. 	
Suggestions for future evaluators: The rater will identify key suggestions to improve the evaluation, and be specific to the sections of the report where shortcomings were found. As relevant, examples will be cited to assist evaluation managers in overseeing future evaluations.	

Areas for improvement are outlined below.

- The stakeholder map from the design report should have been included as an Annex to this CPE. Additionally, stakeholders should have been identified by type of partners, and their interests, power, influence and impact on the evaluand provided.
- The ToC should have been reformulated to ensure key elements of a good quality ToC are included, notably critical assumptions about causal relationships, contextual factors and stakeholder characteristics and a clear impact and problem statement.
- The methodology section should provide clarity on the software packages used to analyse data, and explicitly state if AI was used, and a key limitation noted as being the lack of rights holders' voices in the data collection.
- Ethical issues should be contextualised for this evaluation, and the data collection tools which explain the process of participation for KIs should not give a one week time limit for interview participants to withdraw their consent to participate, but should note they can withdraw their consent at any point in the evaluation process.
- KIs should be disaggregated by type of partner to enable an analysis across different groups of stakeholders, and provide footnotes that illustrate this in the findings section to enable a nuanced and situated understanding of the views of different types of partners, and by extension the voices of different stakeholder groups (acknowledging that no direct beneficiaries/rights holders were interviewed in this evaluation).
- Recommendations should have greater clarity on how they were developed, and greater precision with regard to responsible parties for implementation.
- Given that a gendered analysis occurred, it would have been helpful to state in the methodology that an intersectional lens was applied, to ensure consistency across the evaluation.

SECTION RATINGS			
SECTION A:	EXECUTIVE SUMMARY (weight 5%)	83%	Comments on Rating
Question 1.	Can the executive summary inform decision-making?		
	i Is a clear, standalone document useful for informing decision making, (a minimum of 5 pages, up to a maximum of 7 pages).	Yes	The executive summary is six pages in length and is useful as a standalone document for informing decision-making.
	ii Includes all necessary components of the evaluation report, including: (1) overview of the context and intervention, (2) evaluation purpose, objectives and intended users, 3) scope and evaluation methodology, (4) summary of most significant findings, (5) main conclusions and (6) key recommendations	Partially	Almost all necessary components are included in the executive summary, with an overview, thorough description of the purpose, objectives and scope as well as intended users, methodology, summary of most significant findings (broken down by criteria), as well as a summary of conclusions and recommendations. The background/ context however, would benefit from explicit inclusion to support understanding of the context of UNFPA's work in BiH.
	iii Includes all significant information in a concise yet clear manner to understand the theme, intervention, programme, project and the evaluation.	Yes	The executive summary includes all significant information in a concise and clear manner to understand the CP, and evaluation of it.
SECTION B:	BACKGROUND (weight 5%)	80%	Comments on Rating
Question 2.	Is the evaluand (i.e. intervention/policy/thematic area etc. that is to be evaluated) and context of the evaluation clearly described?		
	i Clear description of the evaluand (e.g. intervention), including: geographic coverage, implementation period, main partners, cost/budget, and implementation status.	Yes	There is a clear description of the BiH 3rd country programme including its implementation status, geographic coverage, implementation period and cost/ budget.
	ii Clear description of the context of the evaluand (e.g. economic, social and political context, relevant aspects of UNFPA's institutional, normative and strategic framework, cross cutting issues such as gender equality and human rights, disability and LNOB dimensions) and how the context relates to the evaluand (e.g. key drivers and challenges that affect the implementation of the intervention/policy/thematic area	Yes	There is a clear description of the complex political and administrative system in BiH, with two main state entities, and the districts/cantons within each. The multiple national, regional and local government entities are set out and the context within which UNFPA operates in BiH is well described, particularly the challenges of working across so many administrative and political boundaries. EU candidacy status is also well noted as an influencing factor in BiH. Social and economic issues are also well-described, with a focus on population demographics, reproductive health, gender equality and young people, along with relevant aspects of UNFPA's normative and strategic framework. With regard to cross-cutting issues gender equality is well covered and explained in the context of BiH and UNFPA, and disability is mentioned in relation to access to reproductive rights, as well as gender equality. The principles of LNOB are noted and taken into account in the BiH context.
	iii Linkages drawn between the evaluand and the ICPD benchmarks and SDGs relevant targets and indicators.	Yes	Information on SDGs is provided in the key facts table in detail, as is expected of a CPE. However, there is little further discussion of ICPD related SDGs in the background section. The report could be more explicit about which SDGs relate to which area of UNFPA's work in BiH. However, there is discussion of key indicators as they relate to UNFPAs overarching mandate.
Question 3.	Are key stakeholders clearly identified and analysed?		
	i Clear identification of key stakeholders which should include implementing partner(s), development partners, rights holders, and duty bearers among others; and of linkages between them (e.g., stakeholder map).	Partially	A Stakeholder Map was provided to the evaluation team by the BiH CO, and while this is included in the design report it is not included in the final report, where it ideally should have been provided as an Annex, along with detail of the type of stakeholders (for example whether they were implementing partners, development partners, rights holders or duty bearers).
	ii Stakeholders are analysed to understand their specific rights, duties, needs, interests, concerns, and potential impact on the evaluand.	Partially	The evaluation does not analyse stakeholder interests, power, or potential impact on the evaluand, and the stakeholder map provided in the design report does not break down stakeholders by type of partner. However, it is acknowledged that the UNFPA Evaluation Handbook provides limited guidance on this matter, and that there has been some attempt to align key stakeholders against CP outputs in the inception/design report.
SECTION C:	EVALUATION PURPOSE, OBJECTIVES AND SCOPE (weight 5%)	100%	Comments on Rating
Question 4.	Is the purpose of the evaluation clearly described?		

	i Purpose of evaluation is clearly defined, including why it was needed at that point in time, its intended use, and key intended users.	Yes	<p>The purpose of this CPE is clearly defined as being to ensure oversight and demonstrate accountability to stakeholders on performance in achieving development results and on invested resources; to support evidence-based decision making to inform development, humanitarian response and peace-responsive programming; to aggregate and facilitate sharing of good practices and credible evaluative evidence to support organisational learning on how to achieve the best results; and to empower community, country and regional stakeholders. It is needed at this specific time for the above accountability and learning reasons, and to inform the development of the 4th CP.</p> <p>The report clearly states that the main audience and primary intended users of the evaluation are: the United Nations Population Fund Country Office Bosnia and Herzegovina (UNFPA CO BiH), the authorities in BiH, implementing partners of the UNFPA CO BiH, rights-holders involved in UNFPA interventions and the organizations that represent them (in particular women, adolescents and youth), the United Nations Country Team (UNCT), UNFPA Eastern Europe and Central Asia Regional Office (EECA RO), and donors, as well as other interested audiences</p>
•	Are the objectives and scope of the evaluation clear and realistic?		
	i Clear and complete description of the objectives of the evaluation, including reference to any changes made to the objectives included in the ToR (if applicable).	Yes	<p>The report outlined the purpose and scope clearly in the introduction followed by setting out the overall objectives, which include providing the UNFPA CO BiH, the country stakeholders and rights-holders, the UNFPA EECARO, UNFPA Headquarters as well as a wider audience with an independent assessment of the UNFPA BiH CP 2021-2025, and broadening the evidence base to inform the design of the next programme cycle.</p> <p>The report also outlines specific objectives of this CPE which are:</p> <ul style="list-style-type: none"> • to provide an independent assessment of the relevance, coherence, effectiveness, efficiency and sustainability of UNFPA support • to provide an assessment of the geographic and demographic coverage of UNFPA humanitarian assistance and the ability of UNFPA to connect immediate, life-saving support with long-term development objectives • to provide an assessment of the role played by the UNFPA CO BiH in the coordination mechanisms of the United Nations Country Team, with a view to enhancing the United Nations collective contribution to country development results, and an assessment of the role of the UNFPA CO BiH in the coordination mechanisms of the Mixed-Migration Group, with a view to improving humanitarian response and ensuring contribution to longer-term recovery; and to draw key conclusions from past and current cooperation and provide a set of clear, forward-looking and actionable recommendations for the next programme cycle.
	ii Clear and relevant description of the scope (e.g. thematic, geographic, and temporal) of the evaluation, covering what will and will not be covered, as well as, if applicable, the reasons for this scope (e.g., specifications by the ToRs, lack of access to particular geographic areas for political, humanitarian or safety reasons at the time of the evaluation, lack of data/evidence on particular elements of the intervention).	Yes	<p>The scope is clearly defined and described. Thematically the CPE covers the following areas:</p> <ol style="list-style-type: none"> 1. Strengthening systemic capacities for provision of sexual and reproductive health services and reproductive rights; 2. Achieving gender equality through women's empowerment and peace building 3. Strengthening individual capacities of adolescents and youth on family planning and gender equality through healthy lifestyles education 4. Strengthening systemic capacities for data collection, analysis and dissemination as well as evidence-based policy development. <p>While geographically the report outlines that the CPE covers all administrative areas across the country – Bosnia and Herzegovina (BiH), Federation of Bosnia and Herzegovina (FBiH), Republika Srpska (RS), and the Brčko District (BD), as well as some cantons where the UNFPA CO BiH worked.</p> <p>Temporally the report outlines that the CPE covers interventions planned and/or implemented in the period from January 2021 through to December 2024.</p>
SECTION D: EVALUATION DESIGN AND METHODOLOGY (weight 20%)		73%	Comments on Rating
Question 6.	<p>Are the selected evaluation questions and evaluation criteria appropriate for the purpose of the evaluation and is there clear justification for their use?</p> <p><i>Note: UNFPA evaluation standards refer to the OECD/DAC criteria such as: relevance, coherence, effectiveness, efficiency and sustainability (not necessarily applicable to all evaluations) and, for country programmes that include circumscribed and limited humanitarian and/or emergency interventions, the criteria of coverage and connectedness</i></p>		
	i Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation. The relevant criteria are specified and are aligned with the questions.	Yes	<p>Evaluation questions and sub-questions are appropriate for meeting the objectives and purpose of the evaluation, and outlined in Table 1, where it is noted they are derived from the ToR. The OECD DAC criteria of relevance, coherence, effectiveness, and sustainability are included, as well as the criteria of coverage and connectedness. These are aligned with questions, and include sub-questions for the criterion of effectiveness which are clearly outlined, and appropriate. A minor point to note is that a short narrative paragraph explaining the selection and use of each criterion would be helpful to include, along with the table setting out criteria aligned against questions and sub-questions. Additionally, in the case of this evaluation it would be helpful to include the assumptions associated with the evaluation questions in Table 1; this is due to the fact that in the evaluation matrix in Annex 1 it is hard to differentiate evidence from assumptions, which makes it more challenging to easily understand what the assumptions were that were tested as part of the evaluation questions.</p>

	ii	Evaluation matrix clearly presents the evaluation criteria used as well as the corresponding evaluation questions, indicators, lines of inquiry, benchmarks, assumptions, source of data, methods for data collection and analysis, and/or other processes from which the analysis can be based, and conclusions drawn.	Yes	The evaluation matrix outlines the thematic scope of the evaluation by specifying the evaluation criterion for each evaluation question, the related assumptions for verification, the indicators and the methods and tools/sources for data collection. Data collection tools are sufficiently detailed and outline the type of data collection (e.g. KII) and with what type of partner, e.g. government officials. An element of good practice in this evaluation matrix, taking into account the minor areas for improvement, is that the matrix is populated with complete evidence, which makes it possible to verify the strength of evidence underlying each assumption/evaluation question and the associated indicators. It enables an easier assessment of the evidence underpinning this report, and a deeper understanding of the data collected.
Question 7.		Is the theory of change, results chain, logical framework, or equivalent framework well-articulated?		
	i	Clear description of the intervention's intended results, or of the parts of the results chain that are applicable to, or are being tested by, the evaluation.	Yes	There is a clear description of the intended results, and a simplified logic framework is provided in the main body of the report (Figure 2), and the full ToC at Annex 6 (as part of the ToR). However, the report would benefit from a more detailed narrative description of the ToC, to compliment the figurative framework.
	ii	Causal relationships between the various elements (e.g. outcomes, including the three or relevant Transformative Results, outputs) of the theory of change, results chain or logical framework are presented in narrative and/or graphic form).	Partially	Some casual relationships between elements of the ToC are well explained narratively, with two main output areas which are aligned with eight outcomes. These are aligned with the UNSDCF outcomes. This information is also presented narratively in Figure 2 where a logic framework for the interventions of the CP in BiH aligned against outputs and outcomes. While risks and assumptions are acknowledged, it would be helpful to outline more clearly what they are in the main body of the report, though it is noted some of this is provided in footnote number 4 (pp. 9). Nevertheless, the information is incomplete and it does not cover all critical assumptions, nor indicate where the assumptions are tested as part of the evaluation questions and matrix. Additionally, the ToC would also benefit from the inclusion of a problem statement and clear impact statement, and in areas where evaluators criticise elements of the ToC, it would be helpful to suggest improvements and provide a reformulated ToC.
	iii	Comprehensive analysis and assessment of the theory of change, results chain or logical framework, and if requested in the ToR, it is retrofitted/reconstructed by the evaluators.	Partially	The ToC is analysed in the narrative text, with areas noted where there are gaps/where improvements could lie for example in being more precise about what risks and assumptions apply to different interventions, or outlining the challenges of measuring outcome targets, due to lack of appropriate indicators used. The ToC has not been validated and refined, and while the ToR could have been clearer about expectations, the UNFPA Evaluation Handbook clearly states that the ToC should not only have been assessed, but refined, and this evaluation could have been more useful if a reconstructed theory of change had been produced by the evaluation team.
Question 8.		Does the report specify adequate methods for data collection, analysis, and sampling?		
	i	Evaluation design and set of methods are clearly described, and are relevant and robust for the evaluation's purpose, objectives and scope, including the use of AI in the evaluation process if applicable.	Partially	A mixed-method and participatory approach was applied which relied on collection and analysis of both quantitative and qualitative data. The design and methods are relevant for the evaluations purpose and the report clearly outlines that it relied on two main methods of data collection - documentary review and individual key informant interviews. The reasons are clearly explained and relevant for the situation. These were supplemented by site visits. However, the lack of rights holders' voices in the evaluation represents a weakness in the design, and is a missed opportunity to access direct feedback from beneficiaries with regard to UNFPA support.
	ii	Data sources are all clearly described and are relevant and robust; these would normally include qualitative and quantitative sources (unless otherwise specified in the ToR).	Yes	A mixed-method approach has been applied with a limited, but appropriate toolkit which primarily included quantitative data collection via document review, with sources provided at Annex 2, and KIIs to gather qualitative evidence with sources provided at Annex 3. This does mean that direct beneficiaries were not included in data collection; however, organisations which represent them, for example youth groups, teachers and other NGOs, as well as local and national government officials were included. However, even although organisations representing rights holders were consulted, beneficiaries are not included as a data source in their own right.
	iii	Sampling strategy is provided - it should include a description of how diverse perspectives are captured (or if not, provide reasons for this).	Partially	The range of documents reviewed is clearly and succinctly explained, from the strategic plan to country programme documents, and national policy documents and datasets. While for KIIs the evaluation team was provided with a stakeholder map by the CO and selected people to interview from this, with a sampling strategy provided in the design report, and summarised in this final report, which states that KIIs, and site visits, were selected on the following basis: <ul style="list-style-type: none"> • type of institution/organization (governmental, non-governmental, UN agency, donor, or other), • jurisdiction (BiH, FBiH, RS, BD, for governmental institutions and NGOs), and • distribution across UNFPA BiH CP outputs. A purposive and non-random selection of stakeholders for KIIs was undertaken to try to achieve a balance by administrative area, type of institution/organization, and intervention area/output (outlined in Table 2). However, this section could be improved if the sampling selection for site visits was provided, along with greater clarity on how LNOB dimensions were taken account of to ensure diversity of voices in the evaluation.
	iv	Methods allow for rigorous testing of the theory of change, results chain or logical framework (e.g. methods help to understand the causal connections, if any, between outputs and expected outcomes (3TRs).	Yes	Methods allow for testing of the ToC, acknowledging the limitations of the ToC.

	v Clear and complete description of the methods of analysis, including explainability and full disclosure of the use of AI in the evaluation process, if applicable.	Yes	There is a description of the methods of analysis, a contribution analysis method was applied, and data extracted from documents against outcomes and impacts, while for evidence for KIIs data was input into a pre-designed Key Informant Interview Matrix, organized by evaluation criteria and Interview Guide questions, while it is explicitly stated that the evaluation matrix was used as a repository of data and the main analytical tool. The report explicitly notes that triangulation of data took place and that KIIs were used to fact check assumptions and evidence gathered from documentary review, with site visits providing an additional level of triangulation. However, it is not clear what software packages were used, and whether AI was used in the analysis process.
	vi Clear and complete description of limitations and constraints faced by the evaluation in its data collection and analysis, including gaps in the evidence that was generated and mitigation of bias, and how these were addressed by the evaluators (as feasible).	Partially	There is a concise limitations section which covers the main issues in relation to the conduct of the CPE: challenges in accessing a wide range of stakeholders, reliability of data, and reducing bias. Appropriate mitigation measures are outlined for each, where feasible. However, this section could have been improved if the ToC had been reconstructed to take into account the two SP time periods that the evaluation covered, additionally the absence of rightsholders voices in the evaluation is a limitation that was partially mitigated by the inclusion of organisations which represent them, but this could have been more explicit. While with regard to document review it may also have been helpful to note the challenges of accessing higher level monitoring data, which meant that the evaluation is more focused on activities and outputs, rather than outcomes or even impacts.
Question 9.	Are ethical issues and considerations described? The evaluation should be guided by the UNEG ethical standards for evaluation. As such, the evaluation report should include:		
	i Explicit and contextualized reference to the UNEG obligations of evaluators (independence, impartiality, credibility, conflicts of interest, accountability) and/or UNEG Ethical Principles.	Partially	There is explicit reference to the UNEG ethical principles and guidelines in the main report. However, this is not contextualised for this particular report, particularly with regard to different stakeholder groups.
	ii Clear description of ethical issues and considerations (e.g. respect for dignity and diversity, fair representation, confidentiality, and avoidance of harm) that may arise in the evaluation, safeguard mechanisms for respondents (e.g. parental consent forms for adolescents, compliance with codes for vulnerable groups; WHO standards of safe data collection on GBV) and ethical considerations in the use of AI as applicable (e.g., transparency of use, explainability, privacy, data protection, accuracy, human rights). If AI is used in the evaluation, there should be transparency and disclosure on the ethical and responsible use of AI in the report.	Partially	There is mention of ethical issues and respect for dignity, and the annexes include data collection tools which include a preamble for KIIs that outlines the evaluation process and how data will be stored with regard to KII participation. However there is not a consideration of ethical issues across all phases of the evaluation. Additionally the interview tool states that participants can remove their right to participate/ use their information, within one week of completing the interview - interviewees should be given the right to withdraw their consent at any stage of the evaluation process.
Question 10.	Does the evaluation incorporate innovative practice that adds value to the evaluation process?		
	i Innovation practice is used to improve the quality of evaluation process. This could include efforts to optimize the evaluation process (e.g., use of AI or new technology for data gathering, content analysis, outcome harvesting among others), or components introduced to enhance inclusion and participation in the evaluation processes (e.g. a youth steering committee), or ways of sharing of evaluation results.	Not Rated	Not rated as not a requirement of a UNFPA evaluation and ToR.
SECTION E:	EVALUATION FINDINGS (weight 25%)	90%	Comments on Rating
Question 11.	Do the findings clearly and adequately address all evaluation questions and sub-questions?		
	i Findings are presented clearly and provide sufficient levels of evidence to systematically address all the evaluation's questions	Yes	Findings are clearly presented, in accordance with UNFPA expectations each finding section corresponds to an evaluation question (with criteria noted) followed by summary of findings, and links to assumptions in the evaluation matrix. This is followed by a numbered findings, under which the evidence to support it is provided, with references to sources in the main text or as a footnote. All of the evaluations questions are answered systematically with evidence to support each finding, or a clear reasons as to why evidence is not available (for example a lack of recent census data) or indicators within the CP which do not have data to be measured against.
	ii Explicit use of the evaluand's theory of change, results chain, logical framework in the formulation of the findings.	Yes	There is explicit use of the evaluands theory of change/ logical framework, which is analysed and tested to inform the formulation of the findings.
Question 12.	Are evaluation findings derived from credible data sources as well as a rigorous data analysis?		
	i Evaluation uses credible forms of qualitative and quantitative data. It presents both output and outcome-level data as relevant to the evaluation framework. Triangulation is evident using multiple data sources.	Yes	The evaluation uses credible forms of qualitative and quantitative data, and does an excellent job at presenting a wide range of quantitative data in particular, which is clearly and consistently referenced in footnote form. Qualitative data and evidence from site visits is also clearly recorded, either within the text or through the use of footnotes. There is strong evidence of triangulation, across different documentary sources, for example data sets held by government and UNFPA documentation, as well the use of KIIs to triangulate evidence from documentary sources. For example Finding 17 outlines that UNFPA has made good use of its human resources – core staff and well-selected consultants and implementing agencies, and this is validated through financial data, KIIs, and CO annual reports. One area which would aid in clarity would be to provide an indication of the type of KII which is supporting (or disputing) the findings, for example not simply to note "KII" as a foot note but instead to differentiate evidence from a KII with government, or KII with NGO, to enable a situated understanding of the different view of partners.

	ii Findings are clearly supported by the evidence presented, both positive and negative. Findings are based on clear performance indicators, standards, benchmarks, or other means of comparison as relevant for each question.	Yes	Findings are clearly supported by the evidence and both positive and negative findings are presented. For example finding 8 notes the success of the health programmes in multiple issues, but includes the deployment of HPV immunisation. The key role that UNFPA played in introducing the first HPV immunisation as a result of partnership between governmental partners and UNFPA CO BiH is clearly outlined, and the fact that governments in BiH took over the responsibility for the provision of vaccines, and vaccines are available in all jurisdictions for recommended age groups is evidenced through KIIs, documentary evidence and CP indicators. Negative evidence is also presented, for example Finding 8 also outlines that UNFPA contributed to the development of clinical guidelines for maternal health, and development of MISP capacities in 10 local communities; however, these guidelines have not yet been implemented and the CPE did not find evidence of formal integration of MISP into local preparedness plans/development strategies. The findings are based on clear performance standards and indicators, with explicit reference to CP indicators throughout the findings section, and with data provided in relevant sections on CPD output indicators, for example Table 10.
	iii Causal factors (contextual, organizational, managerial, etc.) leading to achievement or non-achievement of results are clearly identified. For theory-based evaluations, findings analyse the logical chain (progression -or not- from outputs to high level results).	Partially	Causal factors are analysed as they relate to achievement or non-achievement of results, for example in relation to GBV services the impact of previous conflict and sexual based violence is noted as relating to provision of GBV services today, along with differences in provision in the two governments (RS and FBiH) and district of Brko. Organisational factors are also considered, for some questions, for example it is noted that UNFPA CPD BiH does not have a specific component dedicated to gender equality/women's empowerment and GBV, and this has been included under youth services, meaning that it is challenging to measure performance accurately. The CPE does a good job of analysing the ToC and logical chain progression, though was noted in the limitations section that not all areas have appropriate (or any) outcomes indicators, making it challenging to translate outputs to outcomes. Additionally, as the report mainly analyses achievement or non-achievement of results, there is a lack of systematic analysis of causal factors.
Question 13.	Does the evaluation assess and use the intervention's Results Based Management elements?		
	i Assessment of the adequacy of the intervention's planning, monitoring, and reporting system (including completeness and appropriateness of results/performance framework - including vertical and horizontal logic, M&E tools and their usage) to support decision-making.	Not Rated	Not rated as this was not requested in the ToR.
SECTION F:	EVALUATION CONCLUSIONS (weight 10%)	75%	Comments on Rating
Question 14.	Do the conclusions clearly present an unbiased overall assessment of the evaluand?		
	i Conclusions are clearly formulated and present unbiased summative statements that respond to the evaluation questions.	Yes	Conclusions are clearly articulated statements that present unbiased summative statements. Each conclusion is numbered and presents a statement/judgement, followed by a more detailed description. All conclusions indicate to which evaluation question they relate, as well as to which recommendation they are linked. Conclusions are broken down by programmatic and strategic conclusions. However, it should be noted that evaluation question 7 on connectedness would benefit from more detail, and an explanation of why an operational conclusion was added, and how this differs from programmatic conclusions, would be helpful to include.
	ii Conclusions are well substantiated and derived from findings and add deeper insight and analysis beyond the findings.	Partially	Conclusions are well supported by the evidence presented in the findings and demonstrate a logical connection between findings and conclusions. They are well written and succinctly summarise evidence from the findings, demonstrating a comprehensive, and independent analysis of the findings. However, they are still largely a restatement of the evaluations findings, and as such add limited additional insight and analysis beyond the findings themselves.
Question 15.	Are lessons learned identified? [N/A if lessons are not referenced or requested in ToR]		
	i Lessons learned are derived from the findings and are well substantiated with practical, illustrative examples.	Not Rated	Not rated as not a requirement of a UNFPA evaluation and ToR.
	ii Lessons learned are clearly presented and provide actionable insights on the positive aspects of the evaluand as well as any areas of improvement.	Not Rated	Not rated as not a requirement of a UNFPA evaluation and ToR.
SECTION G:	EVALUATION RECOMMENDATIONS (weight 15%)	75%	Comments on Rating
Question 16.	Are recommendations well-grounded and articulated?		
	i Recommendations are clearly formulated and logically derived from the findings and/or conclusions.	Yes	Recommendations are clearly formulated and logically derived from conclusions. It is clearly noted under each recommendation which conclusions they relate to.
	ii Recommendations are useful and actionable for primary intended users. Specific guidance is provided for its implementation (e.g. actions, deadlines, responsible actors), as appropriate.	Partially	Recommendations are broken down by strategic and programmatic recommendations, and specific guidance is given on implementation (under operational implications). Responsible actors are identified; however, all are directed towards the BiH CO, and other actors, but all at the high level with no named division, units, teams or individuals. As such the lack of precision in identifying responsible parties may reduce the utility and ability to action each recommendation in a timely manner.
	iii Process for developing the recommendations is described, and includes the involvement of key stakeholders (e.g. evaluation reference group members), including those who will be affected by the recommendations.	Partially	Section 1.3 notes that "the preliminary findings, conclusions, and recommendations of the CPE were presented to and validated with the Evaluation Reference Group (ERG)". While this is helpful to provide, the evaluation does not clearly describe how the process of co-creation was implemented in practice, or indeed if co-creation occurred.
	iv Recommendations are clearly articulated and prioritized based on their importance, urgency, and potential impact.	Yes	Recommendations are clearly prioritised as high or medium. It is assumed that none are considered low priority as every recommendation has a priority rating.
SECTION H:	REPORT STRUCTURE AND PRESENTATION (weight 5%)	92%	Comments on Rating
Question 17.	Does the evaluation report include all required information?		

	i Opening pages include: Name of evaluation and/title of evaluation, timeframe of the evaluation, date of report, location of evaluand, names and/or organization(s) of the evaluator(s), name of organization commissioning the evaluation, table of contents (including, as relevant, tables, graphs, figures, annexes); list of acronyms/abbreviations.	Yes	Opening pages contain all expected information, including membership of the ERG and a key facts table of the country context.
	ii Annexes include, if not in body of report: terms of reference, evaluation matrix, list of respondents, results chain/ToC/logical framework, list of site visits, data collection instruments (such as survey or interview questionnaires), list of documentary evidence. Other appropriate annexes could include: additional details on methodology (e.g. inception report), case study reports.	Yes	The Annexes contain almost all expected information, they could be improved if the ToC was presented as a separate Annex, not simply referred to in the ToRs.
Question 18. Is the report logically structured and of reasonable length?			
	i The report has a logical structure that is easy to identify and navigate (for instance, with numbered sections, clear titles, well formatted).	Yes	The report has a logical structure that is easy to identify and navigate with numbered sections and clear titles. Overall it is well formatted, but a minor point to note is that use of bolding and italics to highlight key information is used inconsistently in different sections of the report (for example in section 1.3.1).
	ii Structure and length accords to UNFPA guidelines for evaluation reports; it does not exceed number of pages that may be specified in ToR. <i>Note: Maximum pages for the main report, excluding executive summary and annexes: 60 for institutional evaluations; 70 for CPEs; 80 for thematic evaluations and 50 for other types of evaluations)</i>	Yes	Structure and length (70 pages, excluding the executive summary) meet expectations for a CPE.
Question 19. Is the report well presented?			
	i Report is easy to understand (written in an accessible way for the intended audience) and generally free from grammar, spelling and punctuation errors.	Yes	The report is easy to understand and written in an accessible way for the intended audience. It is generally free from grammar spelling and punctuation errors.
	ii Frequent use of visual aids (such as infographics, maps, tables, figures, photos) to convey key information. These are clearly presented, labelled, and referenced in text.	Partially	There is frequent use of visual aids to convey information, and they are clearly presented and labelled in the text. However, these are primarily in the introduction and background sections and the findings could benefit from the use of figures (in particular) to explain complex information in more easy to understand formats.
SECTION I: CROSS-CUTTING ISSUES (weight 10%)		79%	Comments on Rating
Question 20. Are cross cutting issues - in particular, human rights-based approach, gender equality, disability inclusion, LNOB - integrated in the core elements of the evaluation (e.g. evaluation design, methodology, findings, conclusions and recommendations)?			
	i Evaluation's data collection methods designed to capture the voices/perspectives of a wide range of stakeholders including right holders, marginalized and vulnerable persons, young people, people with disabilities, migrants or refugee populations, indigenous communities, and other persons that are often left behind.	Yes	The evaluations data collection methods are designed to capture a range of data that represent diverse perspectives, and include marginalised and vulnerable people, including people with disabilities, LGBTQI people, migrants, people from the Roma community and young people. Given that no direct rights holders/beneficiaries were interviewed the evaluation does a good job of considering how marginalised groups have been impacted by UNFPA interventions, and where progress is still to be made.
	ii Evaluation questions address cross cutting issues, such as human rights-based approach, gender equality, disability inclusion, LNOB, social and environmental standards as appropriate.	Yes	Evaluation questions address cross-cutting issues such as gender, and inclusion more widely, and an LNOB lens has been applied in the analysis of evidence. There are specific evaluation question to ensure that cross-cutting issues are addressed, including gender equality and how UNFPA meets the needs of marginalised and vulnerable people.
	iii Data is disaggregated by population groups (e.g. persons with disability, age, gender, etc.) where there are implications related to UNFPA's portfolio/interventions for these population groups; differential results are assessed (distribution of results across different groups).	Partially	Data is disaggregated by population group where possible though the evaluation notes that national data sets and census data is not up to date so identifying marginalised communities is challenging. Results are presented across different groups, where feasible. However, an area where this could have been improved is in the use of primary disaggregated data, and breakdown of KIs by type of organisation, at the moment the findings present secondary disaggregated where possible, but do not break down the findings from primary evidence by gender, or indeed by government or by implementing partner. If the evidence from KIs was presented by different types of partner it would enable an analysis of findings across different groups of partners, for example national or local government, NGOs representing marginalised communities or other UN partners.
	iv Intersectional lens is applied in the data analysis, looking at various and multiple forms of exclusion and discrimination (and how they overlap with each other) and how this may impact the performance or results of the evaluand.	Partially	An intersectional lens is applied in parts of the findings. For example, it is noted that marginalised women have more challenges accessing health services, and reproductive health, but there is not an intentional and systematic intersectional analysis applied. There is, however, consideration of multiple forms of exclusion and discrimination, and it is clear the evaluation has a nuanced understanding of exclusion, though this could be more systematically explored in relation to the performance of the CP.
	v Findings, conclusions and recommendations, address cross-cutting issues such as equality and vulnerability, disability inclusion, leave no-one behind, social and environmental as relevant.	Yes	Eevaluation question 1 explicitly asks "To what extent is UNFPA support adapted to the needs of vulnerable and marginalised groups?" As such this and other evaluation questions make explicit reference to marginalisation, vulnerability, gender equality, refugees and migrants as well as youth. This has ensured that findings, conclusions and recommendations address cross-cutting issues, the findings do so repeatedly in addressing each relevant question, while multiple conclusions and recommendations address cross-cutting issues and explicitly mention marginalised and vulnerable groups as well as LNOB. Eight of the nine recommendations relate to vulnerable groups, or issues affecting them, such as (but not limited to) GBV, youth services, migration and targeted healthcare.
	vi Inclusion of young people in the evaluation team and/or Reference Group [N/A if not requested in ToR]	Not Rated	Not rated as not a requirement of a UNFPA evaluation and ToR.

Question 21.	Does the evaluation meet UN SWAP evaluation performance indicators? <i>Note: this question will be rated according to UN SWAP standards with detail provided below</i>	7	Comments on Rating
	i GEEW is integrated in the Evaluation Scope of analysis, and evaluation criteria and questions are designed in a way that ensures GEEW-related data will be collected.	Fully integrated	GEEW is well integrated into the scope of analysis with specific evaluation questions designed to assess to what extent marginalised and vulnerable groups have been supported by UNFPA, while other questions explicitly relate to gender equality, SRH and GBV. Both human rights and gender equality are therefore fully integrated into the scope of the evaluation. Data collection tools also gathered data that was disaggregated, and acknowledging that some secondary data sources (including the national census) did not provide up to data, or disaggregated data, the ability to collect data that did provide information about gender, and gendered experiences was built into the design of this CPE.
	ii A gender-responsive methodology, methods and tools, and data analysis techniques are selected.	Satisfactorily integrated	The methodology and data collection tools allow for the collection of gendered data, and the report is clear about the challenges of disaggregated data with regard to secondary data sources in particular. While rights holders were not included in the methodology, and the primary data collection methods were KIs with stakeholders and document review, the report provides a list of stakeholders interviewed, and also disaggregates this by sex. The sampling framework is provided, but more clarity would be useful to provide on how marginalised voices/voices of their representatives would be included, and how meaningful participation would be enabled in practice, given the evaluation noted it was participatory in nature. The lack of rights holders' voices was a key limitation of this evaluation, as mentioned above.
	iii The evaluation Findings, Conclusions and Recommendations reflect a gender analysis.	Satisfactorily integrated	<p>The evaluation does analyse some specific social groups and notes analyses some issues of intersectionality, for example in accessing reproductive health services it is clearly noted that marginalised women such as migrants or those with disabilities are less likely to be able to access support. The findings also do a good job of disaggregating data wherever feasible, and drawing on a wide range of documentary evidence to supplement the KIs with stakeholders to provide an analysis of gendered issues, and gender equality more widely. Gender and human rights are explicitly addressed in the findings, conclusions and recommendations along with specific mention and attention to vulnerable and marginalised groups. Almost all of the ten conclusions explicitly mention issues relating to gender equality, women's empowerment, or issues affecting women such as GBV, SRH, and women's healthcare issues such as cervical screening and HPV vaccinations. While almost all recommendations contain reference to women, and to issues which affect them, with a strong focus on GBV prevention and support, women's health, women's empowerment and equality more widely.</p> <p>This element could have been improved if there was specific and intentional application of intersectionality from the methodology onwards, and if voices of different groups have been more transparently included in the findings section in particular.</p>

List of SDGs

1. No Poverty
2. Zero Hunger
3. Good Health and Well-being
4. Quality Education
5. Gender Equality
6. Clean Water and Sanitation
7. Affordable and Clean Energy
8. Decent Work and Economic Growth
9. Industry, Innovation and Infrastructure
10. Reduced Inequality
11. Sustainable Cities and Communities
12. Responsible Consumption and Production
13. Climate Action
14. Life Below Water
15. Life on Land
16. Peace, Justice and Strong Institutions
17. Partnerships for the Goals

Three transformative results

1. Ending unmet need for family planning
2. Ending preventable maternal deaths
3. Ending gender-based violence and harmful practices

Six outputs

1. Policy and accountability
2. Quality of care and services
3. Gender and social norms
4. Population change and data
5. Humanitarian action
6. Adolescents and youth

Six accelerators

1. Human rights-based and gender-transformative approaches
2. Innovation and digitalization
3. Partnerships, South-South and triangular cooperation, and financing
4. Data and evidence
5. Leaving no one behind and reaching the furthest behind first
6. Resilience and adaptation, and complementarity among development, humanitarian and peace-responsive efforts