

|   |  |              |
|---|--|--------------|
| <b>Evaluation title</b>   | Evaluation of Humanitarian Response Facilities Network in Pakistan from January 2014 to September 2020 |              |
| <b>Evaluation category and type</b>   | Decentralized  |              |
| <b>Post Hoc Quality Assessment (PHQA) – overall rating</b>  | Satisfactory: 82%  |              |
| <p>The Evaluation of Humanitarian Response Facilities Network in Pakistan from January 2014 to September 2020 is a satisfactory report that evaluation users can rely on the quality and credible evaluation findings provided and can use it with confidence for decision-making. The summary provides relevant information on the evaluation type, context, and overview of the evaluation subject, with key findings for all evaluation questions, conclusions and recommendations clearly summarized. A clear overview of the country context, relevant national policies and strategies, and features of international assistance and other WFP work in Pakistan is found in the report. The report also includes a concise overview of the evaluation subject, with descriptions of thematic area, transfer modality and main partners. The evaluation uses appropriate methods, including a desk review, stakeholder consultations, and observation, drawing from diverse data sources to triangulate findings. There is a focus on gender equality and women’s empowerment (GEWE) dimensions in the methodological approach and questions, specifically related to Relevance and Effectiveness, and in the analysis in the findings section. Key findings are clearly summarized and prioritized from medium to high. Based on the matrix and findings, listed lessons may contribute to wider learning in WFP and guide future action. The report would have benefited from including a description of the affected and targeted population and while the context does provide a limited intersectional analysis, it could have gone further. Conclusions are mapped to related findings and recommendations but could have been mapped to fundamental questions that led to the study objectives.</p> |  |              |
| <b>CRITERION 1: REPORT SUMMARY</b>  | <b>Rating</b>  | Satisfactory |
| <p>The summary is clear and concise, providing a brief overview of the context, purpose, and approach. Key findings for all evaluation questions are clearly summarized. The evaluation conclusions are clearly summarized and prioritized. The recommendations are mapped to findings and presented just as they are found in the main report; however, the summary could have included lessons learned. Further detail on the affected population would also have enhanced the summary.</p>   |  |              |
| <b>CRITERION 2: CONTEXT AND OVERVIEW OF EVALUATION SUBJECT</b>  | <b>Rating</b>  | Satisfactory |
| <p>The report presents a clear overview of the country context, relevant national policies and strategies, and features of international assistance and other WFP work in Pakistan. The evaluation includes an overview of the subject, its activities and thematic areas, including transfer modality and main partners. A reconstructed theory of change (TOC) based on Results Stories (Theories of Change developed by the Country Office in 2016) and available documentation is presented to conceptualize the pathways of change between the outputs of the HRFs and the intended objectives. An intersectional analysis draws the link between humanitarian crises, levels of food security and gender-based violence but could have benefitted by going further.</p>   |  |              |
| <b>CRITERION 3: EVALUATION RATIONALE, OBJECTIVES, AND SCOPE</b>   | <b>Rating</b>  | Satisfactory |
| <p>The evaluation presents dual and mutually reinforcing objectives of accountability and learning, with an emphasis on accountability, as well as two specific objectives on generating evidence and improving effectiveness. Primary intended internal and external users are clearly identified. Evaluation objectives could have been mapped to fundamental evaluation questions. Rationale justification could have benefited from using the sound theoretical framework underlying the evaluation matrix. Narratives on considerations given to human rights and gender equality in setting up the evaluation objectives could have been provided.</p>  |  |              |
| <b>CRITERION 4: METHODOLOGY</b>   | <b>Rating</b>  | Satisfactory |

The evaluation methods included a diverse range of data sources, including KIIs, FGDs, observations and desk reviews, with the sampling frame addressing a diversity of stakeholders. The reconstructed ToC both informed and is reflected in the evaluation matrix. Four evaluation questions with sixteen sub-questions are presented based on the OECD-DAC evaluation criteria of relevance, effectiveness, efficiency, coherence, and sustainability. Ethical considerations and safeguards are mainstreamed throughout. The report could be elaborated further on the evaluability of monitoring data to inform choice of the selected research approach and analysis; and whether, during the implementation period, sufficient information was collected on specific results indicators. The evaluation could have addressed limitations inherent in data analysis for subpopulations.

|                              |               |              |
|------------------------------|---------------|--------------|
| <b>CRITERION 5: FINDINGS</b> | <b>Rating</b> | Satisfactory |
|------------------------------|---------------|--------------|

Findings are presented in neutral language without bias presenting both strengths and weaknesses in a balanced way and mapped to evaluation questions and sub-questions with gaps in evidence clearly flagged. The report makes excellent use of case studies. However, the findings could have been supported by quantitative data analysis based on contents of collected data. The report would also have benefited from adding description of how sample was stratified for data analysis as presentation of findings on indicator nominal variables is suggestive of the assumption that variability is constant among beneficiaries. While findings were not accessed by the International Humanitarian Principles (IHP), post analysis discussions could have informed future investigation to assess performance against the IHP.

|   |               |                        |
|---|---------------|------------------------|
| <b>CRITERION 6: CONCLUSIONS AND LESSONS</b> | <b>Rating</b> | Partially Satisfactory |
|---|---------------|------------------------|

The evaluation conclusions are mapped to related findings and recommendations. However, the conclusions could have been substantiated by statistical evidence from quantitative analysis results. Lessons could also have been drawn out exclusively to show how they may contribute to wider learning in WFP and guide future action.

|                                     |               |                     |
|-------------------------------------|---------------|---------------------|
| <b>CRITERION 7: RECOMMENDATIONS</b> | <b>Rating</b> | Highly Satisfactory |
|-------------------------------------|---------------|---------------------|

Four recommendations with sub-recommendations are grouped as either strategic or operational and mapped to the findings, targeted, specific and actionable with responsible actors clearly identified. Specific recommendations prioritize GEWE, protection, and accountability to affected populations (AAP) more systematically in interventions related to emergency preparedness and response. The report could have described how recommendations from previous evaluations have or have not been addressed through the present intervention design and implementation.

|   |               |                     |
|---|---------------|---------------------|
| <b>CRITERION 8: ACCESSIBILITY AND CLARITY</b> | <b>Rating</b> | Highly Satisfactory |
|---|---------------|---------------------|

The report follows the WFP template and includes all required lists. It is written in clear and understandable language. Sources for all data and quotes are provided, as are cross-references to refer to relevant information entailed in annexes. References in the report are cited in appropriate formats, with key messages highlighted in bold for ease of reading. Tables and figures are appropriately headed and labelled, and visual aids are used consistently throughout the report to convey key information. Annex could have been listed in the order that they are referenced in the main report.

**Integration of Gender Equality and Women's Empowerment (GEWE) considerations in the evaluation report based on the UN System-Wide Action Plan (UN-SWAP) Evaluation Performance Indicator (EPI) scorecard**

|  |                    |
|--|--------------------|
| <b>UN-SWAP EPI – individual evaluation score</b> | Meets Requirements |
|--|--------------------|

Gender is mainstreamed throughout the evaluation with special attention to the perspectives of women, girls, and other vulnerable population groups to reflect and assess attention given to diversity and gender-responsiveness. Questions related to Relevance and Effectiveness ensured GEWE-related data was collected with the effective use of case studies to highlight the voices of affected populations. The report noted limitations in identify women in Pakistan working in relation to the WFP HRF. Several findings speak to cross-cutting dimensions (or the lack there of) during the intervention period, noting that various phases of HRF construction predated WFP's Gender Policy (2015-2020) and identifying key gaps related to considering the needs of women/girls and vulnerable population groups. However, findings could have been presented to include data specific social groups affected by the issue. The report could also have drawn more explicitly on unanticipated effects related to human rights and gender quality as this was identified as an objective of the evaluation. Short narratives, Recommendation 3 focus on GEWE and inclusion; however, sections of the report could have been

dedicated to approaches towards evaluating GEWE and inclusion. if any, on considerations given to human rights and gender equality in setting up the evaluation objectives could have been included.

| Post Hoc Quality Assessment – Rating scale and definitions at overall report and criteria levels |   |
|--|---|
| <b>Highly Satisfactory</b>   | <i>Definition at overall report level:</i> Evaluation users can rely on the credible and useful evaluation findings provided and can use the evaluation with a high degree of confidence for decision-making. The report is considered an excellent example.                                |
|  | <i>Definition at criterion level:</i> The criterion is addressed without any gaps or limitations.   |
| <b>Satisfactory</b>  | <i>Definition at overall report level:</i> Evaluation users can rely on the quality and credible evaluation findings provided and can use it with confidence for decision-making.   |
|  | <i>Definition at criterion level:</i> There are no significant gaps or limitations in addressing the criterion.   |
| <b>Partly Satisfactory</b>   | <i>Definition at overall report level:</i> Evaluation users can rely on the evaluation findings provided and may use it for decision-making noting that there are some gaps/shortcomings in the information provided.   |
|  | <i>Definition at criterion level:</i> There are some significant gaps or limitations in addressing the criterion.   |
| <b>Unsatisfactory</b>  | <i>Definition at overall report level:</i> Evaluation users can use some of the learning from the evaluation, noting that there are significant gaps/ shortcomings in the evaluation findings provided. The report may still contribute to decision making but should be used with caution. |
|  | <i>Definition at criterion level:</i> There are critical gaps or limitations in addressing the criterion. Most of the required parameters are not met.  |