

META EVALUATION OF IOM EVALUATIONS 2020–2024

This evaluation brief presents a summary of the key findings, conclusions, and recommendations, as identified by the evaluator(s) for use by key stakeholders, including internally by IOM staff and externally by project partners. More details can be found in the full evaluation report.

Evaluation type: Centralized evaluation

Evaluators: Leticia Bendelac (Lead eval.), Inmaculada Román, Carlos Madridejos, Lucas Rey and Salvador Bustamante.

Final report date: July 2025

Commissioned and

Managed by: IOM Central Evaluation Division (CED)

SUMMARY

This meta-evaluation examined IOM evaluations published between 2020 and 2024 to develop actionable recommendations for enhancing the quality and utilization of IOM evaluations.

In particular, the evaluation assessed the quality of 90 evaluation reports (from a pool of 265 reports) along with 19 terms of reference. The evaluation also analyzed the evolution of evaluation quality in comparison with the previous meta-evaluation of evaluations 2017–2019, the relevance and application of IOM quality assessment tools, and the internal capacities to ensure the quality of evaluations.

The evaluation was commissioned by the IOM Central Evaluation Division to Artival Research Ltd., in accordance with the IOM Evaluation Strategy 2024–2028 and the Biennial Evaluation Plan 2023–2024.

KEY FINDINGS & CONCLUSIONS

The 2020–2024 meta-evaluation reveals moderate improvement in IOM evaluation quality, with 39% of reports meeting or exceeding standards, 44% approaching requirements and 17% falling below standards, representing an overall score increase from 2.6 (2017–2019) to 2.8 out of 5.

High-performing components included "findings" and "evaluation criteria," while "lessons learned" and "good practices" remained consistently weak across evaluations.

Evaluation purpose: The main purpose was to assess the quality of internal and external evaluations (centralized and decentralized) conducted between 2020 and 2024 and provide recommendations to enhance the quality and utilization of evaluations. The evaluation also assessed the review process, institutional capacities, and the use and relevance of evaluation tools.

Evaluation criteria: Quality, comprehensiveness and clarity, as well as adherence to institutional standards.

Evaluation methodology:

The meta-evaluation employed a utilization-focused approach using mixed-methods: document review, scoring tools, key informant interviews, and surveys administered to internal and external evaluators, and evaluation managers.

Project information:

Geographical coverage: Global

Budget: 37.741 USD

The quality of ToRs and inception reports positively influenced outcomes, though significant gaps persist in disability inclusion, environmental sustainability, and documenting emerging practices. Many ToRs continue to exhibit inconsistencies, with blanket inclusion of all OECD-DAC criteria without assessing contextual applicability, leading to broad scopes and diluted analysis.

For decentralized evaluations, the approach led to significant variability in scope, methodological rigor, and consistency, with ad hoc planning and limited systematic oversight, especially for smaller or short-term projects. The absence of standardized review processes for decentralized evaluations and irregular involvement of RPMEOs contributed to uneven quality. Quality assurance tools exist but remain inconsistently applied across offices due to competing priorities, limited awareness, and voluntary compliance approaches.

Structural challenges include fragmented responsibilities, absence of mandatory procedures for reviewing evaluation deliverables, and insufficient involvement of RPMEOs in evaluation processes.

EVALUATION BRIEF

RPMEOs are underutilized in evaluation processes and disproportionately tasked with strategic planning and monitoring. Meanwhile, recent initiatives—like broader dissemination of evaluation briefs and enhanced donor engagement—show promise for strengthening learning, strategic alignment and evaluation use.

The proportion of budget allocated to evaluation significantly influences quality of evaluations. Inadequate time allocation has led to tight timelines, superficial analysis, and methodological compromises, with many reports prioritizing donor compliance over meaningful learning.

In terms of capacities and skills, the number of participants to internal evaluation training dropped from 256 in 2021 to just 1 by 2024. High staff turnover, sporadic training, and inconsistent consultation of critical guidance documents contribute to persistent technical capacity gaps. While core resources like the IOM Project Handbook and M&E Guidelines remain widely used, inconsistent application of specialized guidance (such as gender evaluation guidance) contributes to cross-cutting issue gaps.

Regional offices play a crucial role in enhancing quality through practical technical guidance and capacity-building, with evaluation quality highest where RPMEOs actively review ToRs and deliverables. However, support remains uneven and often dependent on individual project manager initiative.

Despite improvements since 2020, evaluations are still largely viewed as contractual obligations rather than integral learning tools. Weak sampling strategies, insufficient triangulation, and inadequate data collection transparency often lead to superficial findings and broad, non-evidence-based recommendations. The absence of structured stakeholder engagement further weakens recommendation relevance.

A structured, mandatory quality assurance system embedded in approval workflows, combined with more practical and flexible guidance, sustained training pathways, and strengthened institutional mechanisms are essential to improve evaluation quality, foster a culture of learning and accountability, and enhance organizational decision-making effectiveness.

CROSS-CUTTING ISSUES

The assessment examined how gender, human rights, disability inclusion, and environmental issues were integrated into the design, findings, conclusions, and recommendations of evaluations.

GOOD PRACTICES AND LESSONS LEARNED

In offices where RPMEO are actively involved in the review of evaluation deliverables, evaluations rigour and coherence are improved. Standardized templates enhance reporting consistency, though not necessarily the depth or quality of content.

Weak integration between monitoring and evaluation, technical expertise gaps, and inconsistent evaluation requirements have undermined the depth, data quality and prioritization of evaluations. This has resulted in methodological gaps and a perception of evaluations as an optional formality rather than an essential management and learning tool.

KEY RECOMMENDATIONS

1. Keep promoting the use of the evaluation report templates and ensure that their use is mandatory to ensure that all the required sections of such reports are included.
2. Continue advocating for the inclusion of gender equality and RBA in programming as cross-cutting issues as well as promoting environmental sustainability and disability inclusion during evaluations. This integration should begin with early needs assessments, stakeholder identification, and the development of the evaluation methodology, continuing through the findings and recommendations sections.
3. Continue promoting that a fixed percentage of the overall project budget be allocated for evaluation to ensure the quality of evaluations, as recommended in the IOM Monitoring and Evaluation Guidelines. This approach appears to ensure evaluations have the necessary resources to be of high quality.
4. Promote a culture of adaptation that enables methodological adjustments tailored to the specific context of each evaluation, to improve the utility and impact of the framework.
5. Enhance the quality of ToRs by strengthening the clarity, focus and methodological rigour, and by establishing review procedures, ultimately leading to more impactful and actionable insights.
6. Make the quality assurance tools more accessible and practical to promote their consistent use across different teams.
7. Build long-term capacity in evaluation through structured and continuous training efforts. Relying on one-off webinars or ad hoc learning opportunities has proven insufficient, particularly considering high staff turnover.