FINAL REPORT

Assessment of ICRC's Evaluation Function

BY CARA J. WINTERS, GLENN O'NEIL
PATRICIA GOLDSCHMID & HUGUES JUILLERAT

20 September 2019

Owl RE
Wise research & evaluation

TRAASS INTERNATIONAL

Consortium of TRAASS International & Owl RE

ABOUT THE AUTHORS

The Assessment was conducted by a four-person team in a consortium of two organizations, TRAASS International SA and OWL RE. The Consortium specializes in supporting organizations in designing and establishing M&E systems, as well as leading individual evaluations, research initiatives, and training. The four consultants in the Consortium include:

- *Cara Winters,* Team Leader - Cara has over 15 years of humanitarian and international development experience as a researcher, technical specialist, senior management leader, and consultant. She specializes in organizational monitoring and evaluation systems, humanitarian accountability, and operational and program strategy. Cara has a background in emergency response programming, development economics, and public health and infectious diseases.
- *Glenn O'Neil,* Senior Evaluator - Founder and a Managing Director of OwlRE, Glenn is an experienced evaluator and humanitarian practitioner with over 25 years of experience as a consultant, researcher and field worker. He has led some 100 evaluations and reviews for over 40 organizations. Evaluations, reviews and research projects have been carried out for the ICRC in the fields of communications, fundraising, cooperation, prevention, operations and archives.
- *Patricia Goldschmid,* Evaluator - Multilingual and multicultural evaluation and communications consultant for over 15 years, Patricia has carried out evaluations spanning over 40 countries in Europe, Latin America, Africa and Asia. Projects encompassed evaluation of communications and advocacy programs and involved running field visits, managing interviews, developing surveys, as well as reviewing methodologies, tools, case studies, reports and other innovative delivery methods such as multimedia presentations.
- *Hugues Juillerat,* Evaluator - Founder and a Managing Director of TRAASS International, Hugues has more than 9-year experience in humanitarian work, predominantly at Médecins Sans Frontières. During the last 6 years he has been directly involved in conducting and managing more than 30 evaluations and evaluation-like exercises as either an evaluation consultant or an evaluation manager.

ACKNOWLEDGEMENTS

Disclaimer: The authors wish to emphasize that the views expressed in this report have been developed based on their assessment and are theirs alone. The views expressed in this publication do not necessarily reflect the views of the ICRC.

TABLE OF CONTENTS

ACRONYMS AND ABBREVIATIONS

ALNAP	Active Learning Network for Accountability and Performance
DAC	Development Assistance Committee
DirGen	Office of the Director General
ECM	Evaluation Coverage Mapping
ECS	Evaluation Case Study
FGD	Focus Group Discussions
HoS	Heads of Sectors
HoU	Heads of Units
HQ	Headquarters
ICRC	International Committee of Red Cross
IFRC	International Federation of the Red Cross and Red Crescent Societies
IR	Inception Report
IT	Information Technology
KII	Key Informant Interviews
MOPAN	Multilateral Organization Performance Assessment Network
MSF	Médecins Sans Frontières (Doctors without Borders)
OECD	Organization for Economic Co-operation and Development
PfR	Planning for Results
PME	Planning Monitoring and Evaluation
RFP	Request for Proposal
ToR	Terms of Reference
UNHCR	United Nations High Commissioner for Refugees

TABLES

EXECUTIVE SUMMARY

This report is an independent assessment ('the Assessment') of current evaluation practice and a future evaluation function at the International Committee of the Red Cross (ICRC). It serves to inform the design of an evaluation function that responds to the ICRC context and the ambitions set forth in the ICRC Strategy 2019-2022. The Assessment was carried out from May to September 2019 through a mixed-methods, participatory approach following an institutional analysis framework and forward-looking gaps analysis.

FINDINGS

Organizational Performance: ICRC is not currently producing or utilizing evaluations in a way that meets organizational expectations as implied in the 2019-2022 strategy or as expressed by staff during the course of Assessment or in key organizational documents. Additionally, ICRC is not effective in producing or utilizing evaluations as benchmarked against key industry standards for evaluation practice in humanitarian organizations, performing well below average at 45% alignment to standards. The primary area of weakness is 'consistent, independent evaluation of results.' This is linked to the absence of a guiding evaluation strategy and clearly defined evaluation function, as well as weak practice around evaluation identification, coverage assurance, and planning.

Organizational Capacity: ICRC does not have sufficient capacity to produce or utilize evaluations in a way that meets organizational expectations or as benchmarked against industry standards. Areas of organizational capacity that need to be strengthened or introduced to meet ICRC's evaluation performance goals include: sustained evaluation leadership and clarified responsibilities for evaluation initiation and participation, improved efficiencies in use of staff time and resources when producing and using evaluations, sufficiently staffing evaluation in terms of the number and expertise of staff, increased predictability and assurance of evaluation funding, application of IT platforms to leverage evaluation management, governance, and evaluation use, application of existing incentive and compliance systems of leverage how evaluations are used for decision making, learning, and accountability.

Organizational Culture: Organizational culture at ICRC influences how staff understand the scope and structure of an evaluation function, attitudes about the purpose of evaluation, levels of enthusiasm or interest in evaluation as compared to worry or scepticism, understandings of 'independence' in evaluation, adaptability to change in the way evaluations are managed or structured, and belief in the utility and potential of evaluations to respond to various needs of organizational decision-making. These areas of organizational culture impact the way staff may react to or support a shift in how evaluation is formalized and resourced and the extent to which ICRC invests in information harmonization. Attitudes about evaluation and divergent interests in the scope of the evaluation function can serve as opportunities or obstacles, depending on how decisions are communicated and capacitated.

External Environment: The political and economic environment of donor and partner expectations influence ICRC decisions about what to evaluate, how or whether to share evaluations externally, and staff perspectives about the role of an evaluation function. The social or outside normative environment primarily influences alignment within the Movement on evaluation definitions and practice. No influences were found in the legal or administrative environment or technology. No direct threats were noted from the external environment. Opportunities include flexibility in the relationship between donor funding and evaluations and ICRC's desire to maintain institutional control over its evaluation agenda. The social and normative environment provide ways for ICRC to improve how it defines, manages, and uses evaluations.

CONCLUSIONS

Although ICRC's current approach to evaluation encourages self-directed learning and evaluation ownership in commissioning teams, it also reinforces an overall weak evaluation culture that incentivizes 'non evaluation' among staff who are unfamiliar with or skeptical of humanitarian evaluation. ICRC struggles to deploy evaluation in support of institutional governance, accountability to affected populations, and consistent approaches to organizational learning. The end result is an evaluation portfolio that neither reflects ICRC's strategic direction nor provides an evidence base commensurate to the size and function of the organization.

Key gaps in evaluation practice are found in geographic coverage, evaluation types (e.g. transversal or global initiatives), and evaluation activity in Protection, Movement, and Prevention métiers. Issues within the evaluation system that need to be addressed include assurance of evaluation independence, how different types of evaluation exercises are defined, the way evaluations are identified budgeted, and capacitated in terms of quality assurance and technical expertise, and the way in which ICRC uses evaluation within its strategic development, planning processes, accountability mechanisms, and in harmonization with other learning functions. In addressing these gaps, ICRC should leverage opportunities in its existing capacities, culture, culture, and from its external environment. Existing good practice should be utilized as an example within the organization.

RECOMMENDATIONS

RECOMMENDATION 1: **Proposed application of industry standards.** ICRC should align with industry standards that promote clarity and consistency in the production and use of quality evaluations. Industry standards can be applied in phases, first prioritizing actions that serve as foundational layers for the improved performance of the entire system. They should respond to the key issues identified in current practice, including a lack of documented definitions, requirements, or expectations for evaluation performance and fragmented evaluation practice.

RECOMMENDATION 2: **Purpose and structure of the evaluation function.** ICRC should adopt a 'hybrid' approach to the evaluation function and structure, wherein the function combines elements of a service-oriented 'centre of excellence' model with more accountability-oriented 'command and control' models. Key areas that the evaluation function would have direct authority over include: quality assurance over centralized and decentralized evaluations, definition and assurance of evaluation planning at centralized and decentralized levels (including enforcing ICRC's evaluation coverage commitments), and leadership on how ICRC embeds evaluation into organizational systems that support evaluation use.

ICRC should consider a 'split responsibility vs. location' approach to the placement of the evaluation function, wherein the function is located in the Director General's (DG) office and responsibilities for aspects of the evaluation function are split between the Assembly and the DG to maximize the learning-oriented benefits of an 'executive office placement' model with the more accountability-oriented 'governing body placement' model supported by donor evaluation standards (reflected in MOPAN criteria) and UNEG.

RECOMMENDATION 3: **Resourcing the evaluation function.** ICRC should staff the evaluation function according to the 'hybrid' model for its purpose and structure. This includes a Head of Sector / Unit with experience in humanitarian 'center of excellence' models and experience in humanitarian evaluation systems, at least three staff with experience in organizational systems to build the function and manage its operations, and hiring a dedicated 2-year post with expertise in contribution analysis, outcome harvesting, and outcome mapping to work specifically with the Protection, Prevention, and Movement metiers. ICRC should consider developing a staff roster for conducting reviews and mixed team evaluations and a consultancy roster once the function is established.

ICRC should move towards a 'split-control' model for evaluation budget ownership that creates three types of dedicated evaluation 'pots' held by the evaluation function, the five HQ Departments, and a delegation pot managed by the Director of Operations. Funds are ringfenced and cannot be reallocated. ICRC should consider sourcing the budget from a combination of re-allocation of core funds and line-item requests to donors, while recognizing the need to be strategic in how it approaches donors for specific evaluation funds. The budget amount should be determined according to an allocation benchmark, where ICRC dedicates a set proportion of its the total operating budget towards evaluation. The required amount should cover both the costs of staffing and individual evaluations. The allocation levels should be phased over the initial development of the function to ensure absorption capacity. Years 1 and 2 would range from .07 to 0.1% (1.4 to 2 million CHF) of ICRC's operating budget, depending on the need to hire new staff (versus restructuring existing positions). It would cover the costs of staffing and evaluations at the HQ level. Year 3 would range from 0.1 to 0.13% (2 to 2.6 million CHF), expanding funds to include the delegations 'pot' of evaluation funding and cover the potential for additional staff capacity at the field level. By Year 5, ICRC should have a stable and set allocation that it continues to use in the years going forward, allocating no less than 0.1% and no more than 0.3% of total organizational expenditure towards evaluation.

RECOMMENDATION 4: **Information requirements for the Evaluation Function.** The evaluation function requires information that enables tracking and quality assurance for evaluations, supports its ability to manage the implementation of the evaluation strategy, and equips the function to promote evaluation use. ICRC should ensure that the function has access to: evaluation planning / evaluation section of the Planning for Results (PfR) (decentralized evaluations), evaluation ToRs and budgets, evaluator bids and contracts, evaluation reports and other key products, evaluation management responses, and documented action on the commitments agreed to in the management response. The function will also need to develop a complete evaluation report library, create an annual evaluation coverage and frequency mapping, and track evaluation spending and compliance against evaluation requirements. A semi-regular health check exercise is also recommended. This can be done internally as a self-assessment review.

RECOMMENDATION 5: **IT systems and tools to leverage evaluation practice.** ICRC should use technology to leverage the management of individual evaluations (including centralized and decentralized evaluations), evaluation governance and the routine operations of the evaluation function (including the management and maintenance of ICRC's evaluation strategy, guidelines, and provision of evaluation quality assurance), and evaluation use (including the use of data on and from evaluations at multiple levels of ICRC decision making, learning, and accountability procedures). Configuration of existing SharePoint and Tableau platforms will be required for aspects of these areas. Additional database and data management capacity will also be required.

RECOMMENDATION 6: **Evaluation Strategy.** ICRC's evaluation strategy should include an evaluation framework that defines different levels and types of evaluation activity, including centralized and decentralized evaluations and reviews and their corresponding parameters for lines of inquiry. The strategy should also address standards and requirements for evaluation practice, the considerations for the purpose, scope, and placement of the evaluation function proposed under Recommendation 1, and how the function is resourced addressed under Recommendation 2. When developing the implementation plan for the strategy, ICRC should review the annexes included in the strategy on phasing the developing of the evaluation function and ways to leverage existing ICRC incentive and compliance systems for improved evaluation use.

RECOMMENDATION 7: **Additional areas.** ICRC should consider ways to better connect evaluation protocol and shared analysis with existing accountability and learning functions, including the AAP framework, linkages with the CORE team and Internal Audit, and complementing MfR reports with evaluation learning. Promotion of content

ASSESSMENT OF ICRC's EVALUATION FUNCTION

SEPTEMBER 2019

KEY STATS AND EVIDENCE

BY CARA J. WINTERS, GLENN O'NEIL, PATRICIA GOLDSCHMID & HUGUES JUILLERAT

Consortium of TRAASS International and Owl RE
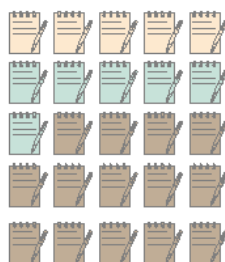

FACTS AND FIGURES
2016 - 2019

25 evaluations and evaluation-like exercises

COMMISSIONING UNIT
- Field Delegation
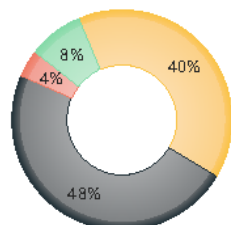- Department/Agency
- Métier

36% of evaluations were 'process' with no defined time period
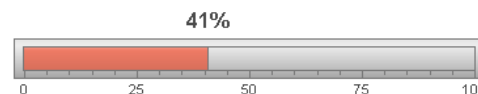
64% of evaluations were ex-post

TYPE OF EXERCISE
- Lessons learnt — 4%
- Assessment — 8%
- Evaluation — 40%
- Review — 48%

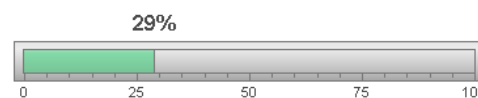**19** countries where evaluation activity took place

---

CURRENT STATE OF EVALUATION

Existence of an independent organizational evaluation function — 41%
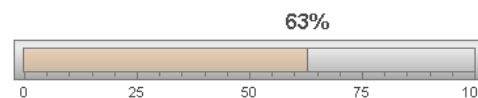
Gaps: No independent evaluation function in operation.

Consistent, independent evaluation of results (coverage) — 29%

Gaps: Evaluation Strategy is no longer in active use and outdated. No organizational evaluation plan.

Systems are applied to ensure the quality of evaluations — 63%

Gaps: Existing organizational-level guidelines not in active use and outdated. Practice is ad-hoc, especially for evaluation management and independence of evaluators.

Uptake of lessons learned and best practice from evaluations — 42%
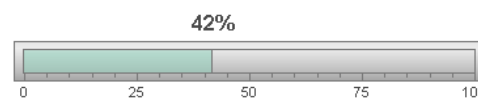
Gaps: No system to track the uptake of lessons learned. No repository of evaluations. Weak dissemination practices.

CASE STUDIES PERFORMANCE AGAINST CRITERIA

Legend:
- Independence
- Quality
- Utility
- Overall meeting objectives

AAP — TTL 71 (19, 18, 17, 17)
Protection Philippines — TTL 40 (8, 11, 8, 13)
HCiD — TTL 71 (14, 19, 19, 19)
EcoSec SSD — TTL 55 (14, 17, 12, 12)

Note: A maximum of 25 points were allocated for each criterion.
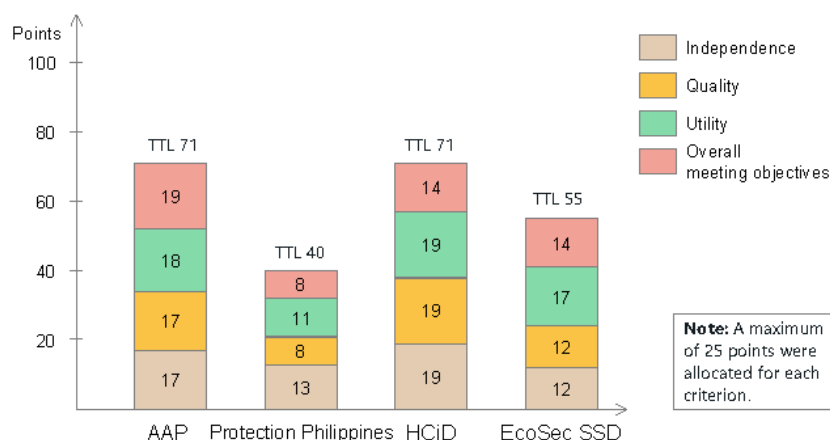
WHAT RESPONDENTS ARE SAYING

AS IS

"Challenge for an evaluation is that **there are many things going on but no connection point.** What is missing is organised transversality."

"There are many reports but **not shared and not accessible to the field.** Often said that they are on "team space" but not found by field. **Evaluation should go beyond what we currently have** which at best are reviews or monitoring."

"Part of the struggle will be that it is not in our **culture to do evaluation**".

TO BE

"The ICRC would need **a standard frame** for developing ToRs and **guidelines on how to do it** as there is currently little knowledge on this. Could have an in-house expert with a network of evaluation experts and a budget to provide **advice on how to manage an evaluation.**"

"How we learn as an organization needs to be strengthened... when evaluations are done – they are parked... We should look into how we can **integrate learning into the strategy process** and how this informs what happens next."

1. INTRODUCTION

1.1 Background

This report is an independent assessment ('the Assessment') of the evaluation function at the International Committee of the Red Cross (ICRC). The Assessment was conducted by a four-person team from the TRAASS International and Owl RE consortium. It was commissioned by the office of the Director General ('DirGen') and managed by the Planning, Monitoring, and Evaluation (PME) Sector and a cross-departmental steering panel. The Assessment was carried out from May to September 2019 through a mixed-methods, participatory approach. The time period covered by this Assessment primarily focuses on current practice, with an analysis of evaluation reports from 2016 until mid-2019. It also reviewed the history of the evaluation function at ICRC from the publication of its evaluation guidelines (2002) until the present day.

1.2 Objectives

In 2006, the ICRC launched the "Institutional Strategy for Independent Evaluation" with an aim to strengthen the use of independent evaluations at all levels of the organisation. Since the approval of the strategy, structures supporting evaluation quality and management shifted from centralised units (Geneva Evaluation Unit 2006-2008, Institutional Performance Monitoring Unit 2009-2013) to the current decentralised approach where evaluations and evaluation-like exercises[1] are largely initiated, funded, and managed from within operational or functional units.

The Assessment of ICRC's Evaluation Function serves to inform the design of an evaluation function that responds to the ICRC context and the ambitions set forth in the ICRC Strategy 2019-2022 (see: Strategic Orientation 2, Objective 2.7). Its **purpose** is to review the *existing evaluation practice* at ICRC, establish a vision for *the desired status of evaluation* that draws from industry evaluation standards to provide a forward-looking *gaps and action analysis* that serve as a basis for recommendations on the scope and structure of the evaluation function, its required level of resourcing and capacitation, and a revised evaluation strategy.

The assessment addresses the following **objectives**:
- I. Define the business model of the evaluation function
- II. Define organisational structure, staffing levels, roles and skills requirements necessary to support the future business operations of the evaluation function
- III. Advise on the IT systems and tools that respond to the future business operations of the Evaluation Function
- IV. Define the information requirements for effective management of the Evaluation Function
- **V.** Develop an Evaluation Strategy

---

[1] "Evaluation" refers to the systematic and objective assessment of the design, implementation, and results of an on-going or completed project, programme, initiative, policy, or otherwise defined area of work (adapted from OECD DAC definition). This team considers "evaluation-like" exercises to include reviews and lessons-learned exercises that likewise assess the performance of an intervention or defined area of organisational work on a periodic or ad hoc basis. Personnel performance assessment, policy research, need assessments, monitoring exercises and situation analyses are excluded from the definition. See also: https://www.oecd.org/dac/evaluation/2754804.pdf

1.3 Methodology

The Assessment followed an institutional analysis framework[2] formulated to examine ICRC's evaluation context, existing practices, and future ambitions as shaped by:
- organizational performance
- organizational capacity
- organizational culture
- external influencing environment

The table below details the final set of 'guiding components' used as entry points for understanding existing and desired practice within each element of analysis.

Table 1: Elements of Analysis and Guiding Components

Elements of Analysis	Guiding Components
Organizational Performance	• Evaluation governance (existence / location of evaluation function, strategy, guidelines) • Evaluation purpose and scope (definition and understanding of) • Evaluation identification and commissioning (protocol for and practice of) • Evaluation selection criteria (existence and use of) • Evaluation frequency and coverage (standards or requirements for, existing rates) • Management of individual evaluations (standards for and practice of) • Evaluation content (standards for and practice of; report structure, lines of inquiry, evidence quality, etc.) • Evaluation utilization (standards for and practice of; visibility, access and knowledge management, usage for learning and accountability)
Organizational Capacity	• Evaluation leadership (vision and championing, management engagement) • Evaluation staffing and overall staff capacity for evaluation • Evaluation funding (budgeting, resourcing) • IT systems and initiatives • Incentive or compliance systems to promote evaluation practice and use (performance management, place of evaluation compliance in internal audit procedures, etc.)
Organizational Culture	• Evaluation function history and motivation for re-establishing the function • Attitudes about evaluation • Adaptability • Information use and evidence-based decision making
External Influencing Environment	• Administrative or legal environment (contractual obligations for evaluation) • Political and economic environment (role of external funding or the politics affecting donor institutions in deciding what, when, or how to evaluate) • Social / normative environment (role of social/normative environment in shaping evaluation purpose and design) • Technological environment (ownership of information technology, reliance on externally managed or developed platforms)

---

[2] The approach is influenced by Ostram's Institutional Analysis and Development Framework, the Institutional and Organisational Assessment Model developed by Universalia and the International Development Research Centre, and approaches focused on knowledge production that are compatible with the OECD DAC criteria. The above guiding components were streamlined from those found in the Inception Report to better match the Assessment findings and reduce redundancy between sections.

The following questions frame the assessment methodology and findings. They form the structure of the Assessment conclusions, in Section 3.

Table 2: Framing Questions

Elements of Analysis	Framing Question
Organizational Performance	1.1 Is ICRC currently effective in producing[3] and utilizing evaluations that meet or exceed organizational expectations? *(As-Is)* 1.2 Is ICRC currently effective in producing and utilizing evaluations as benchmarked against accepted industry standards? *(As-Is)* 1.3 What standards should be in effect at ICRC to promote improved operational impact, organizational learning and accountability, and otherwise reach the objectives of the 2019-2022 ICRC strategy? How are these prioritized for action? *(To-Be)* 1.4 What organizational performance gaps exist between current evaluation practice and the desired end-state? *(Gaps analysis)*
Organizational Capacity	2.1 Does ICRC currently have sufficient capacity to produce and utilize evaluations? *(As-Is)* 2.2 What areas of organizational capacity need to be strengthened, changed, or introduced in order for ICRC to meet its evaluation performance goals and strategic objectives? *(To-Be + Gaps analysis)*
Organizational Culture	3.1 How does organizational culture influence the current practice in evaluation production and the utilization[4] of evaluations? *(As-Is)* 3.2 Are there areas of organizational culture that can be leveraged for improved evaluation practice? Are there areas of organizational culture that could be a hinderance or an obstacle when seeking to improve evaluation practice? *(To-Be + Gaps analysis)*
External Influencing Environment	4.1 How does the external environment (e.g. donors, sector) influence the current production and utilization of evaluations? *(As-Is)* 4.2 What opportunities or threats do these influences present to current and future evaluation performance, capacity, and culture at ICRC? *(To-Be + Gaps analysis)*

Evidence Base and Methodology

The Assessment applied a mixed methods approach, drawing on qualitative and quantitative evidence and analytical techniques. Methodological triangulation is used to ensure the validity and depth of findings. The methodology sequenced evidence generation, with each layer of evidence informing the next (as feasible within the assessment timeline).

Six data gathering and analysis methods are applied in addressing the Assessment questions and objectives: evaluation coverage mapping (ECM), evaluation case study (ECS), key informant interviews (KII), focus group discussions (FGD), a cross-sectional as-is/to-be workshop (Workshop), and an evaluation performance and system

---

[3] 'Producing' and 'production' used here refer to all stages of the evaluation process, including targeting and decisions on when and what to evaluate, evaluation design, resourcing, and process management (e.g. ToR creation, quality control, evaluation team recruitment and oversight, management response formulation, and dissemination).

[4] 'Use' and 'Utilization' here refer to who engages with the evaluation findings and recommendations (including position, level of the organization, and place of evaluation information in different levels of organizational business analytics and institutional or programmatic learning or strategy formulation), follow up and tracking of the management response, and policies on publication and transparency.

health check (Health Check) that benchmarks ICRC performance against industry standards (including donors via MOPAN criteria, Movement actors (IFRC), and fellow humanitarian agencies MSF and UNHCR).

Table 3 highlights the areas of analysis to which the methods primarily responds. Summary reports are provided on each method in Annexes I – V, describing the approach, persons involved or consulted, analysis and findings, and summary conclusions. These annexes are referred to throughout the report as source documents for the presented evidence.

Table 3: Analysis Evidence Base

Element of Analysis	Document Review	Evaluation Mapping	Evaluation Case Study	Key Informant Interviews	Focus Group Discussion	Workshop	System Health Check
Organizational Performance	✓	✓	✓	✓	✓	✓	✓
Organizational Capacity	✓		✓	✓	✓	✓	✓
Organizational Culture	✓		✓	✓	✓	✓	
External Environment	✓			✓	✓		

1.4 Limitations

The Inception Report (IR) identified two potential limitations and notes on how these would be mitigated. In both cases, the limitations did not prove a major obstacle to the Assessment. They were, however, present and should inform how the data is interpreted and used:

- **Relevant ICRC staff are unavailable or inaccessible, limiting the review's overall evidence base.** It proved challenging to identify relevant respondents for the case study survey due to insufficient information in the original evaluation reports and difficulty tracking down persons in the organization who would know or remember the people involved in producing the evaluation. Consequently, the response rate to the survey was lower than desired and it is possible that key persons were not included in the request to participate. The case study limited the reviewed cases to reports with an adequate number of respondents to complete the analysis as required. This reduced the number of reviewed cases to four out of an anticipated ceiling of six.
Although no field staff were available to join the Workshop as was originally planned, the perspective of field staff was captured during the KIIs and FGDs. Participating field staff represented a range of métiers and position levels, as well as years of experience with the ICRC.

The Assessment team was unable to complete an interview with the World Bank as initially intended and described in the IR. This was the result of the Bank not responding to requests for participation.

- **Absent or missing documentation on evaluation practice and processes.** It proved challenging to collate a complete record of evaluation reports commissioned over the past three years due to the absence of an evaluation report library and limited knowledge on where reports are stored or if reports had actually occurred in several métiers. Consequently, the evaluation mapping (Annex I) reflects the total *known and visible* set of evaluation and like exercises conducted between 2016 and mid-2019. It is possible that it is missing information, especially on less formal exercises commissioned or otherwise initiated directly by field delegations (e.g. internal reviews that include evaluation-like questions).

2. FINDINGS - *Existing Practice, Desired Status, Gaps Analysis*

The findings of the Assessment are based on the four analytical elements described in Part 1.3 (Table 1) and cover the existing practice ('As Is') and the desired status ('To Be') for each of the guiding components used as entry points to understand organizational performance, organizational capacity, organizational culture, and external influencing environment. Findings from across all categories inform the recommendations provided against each Assessment objective.

To Be sections express objectives for evaluation performance. They are drawn from two sources of evidence: the ICRC 2019-2022 strategy and the expressed or otherwise documented objectives of ICRC staff. 'Expressed objectives' were gathered and analysed during the Assessment. Key issues identified under the As Is analysis are matched against their most relevant To Be objective in a Gaps Analysis summary for each analytical area. This analysis presents where action on the gap or issue will be addressed in the recommendations.

The following distinctions between each section should be noted:

Section I, Organizational Performance. Benchmarks ICRC systems and practice of evaluation against industry standards. A detailed presentation of the benchmarking is available in the Health Check report, Annex V. This includes the referenced standards, the performance score, and an explanation for the scoring. Relevant findings on performance that are not contained in the Health Check are based on other methods used during the Assessment and cited for their corresponding annexes.

Section II, Organizational Capacity. Reviews challenges and potential opportunities in key areas of organizational capacity, framed as drivers of organizational performance.

Section III, Organizational Culture. Reviews challenges and potential opportunities in key areas of organizational culture, framed as drivers of organizational performance and capacity.

Section IV, External Influencing Environment. Reviews ways the external environment influences organizational performance, capacity, and culture.

✓ I. Organizational Performance

This section reviews the production and utilization of evaluations and evaluation-like ('like') exercises at ICRC *as compared to organizational expectations and as benchmarked against accepted industry standards.* Drivers of performance (including capacity, culture, and external influence) are discussed in more detail in Sections II-IV.

A. Existing Practice (As Is)

FINDING 1: **Evaluation governance.** ICRC is not aligned with industry standards for an independent evaluation function, including the defined existence of a function, separation of evaluation governance from management functions, ownership of an independent budget, and the reporting line of the head of evaluation to the governing body (in this case, the Assembly) (see Health Check criteria 1.1, 1.2, Annex V). This also places ICRC in conflict with its own principle of independence in the way it conducts evaluations (Health Check criteria 3.12). Although the PME Sector has a loosely defined responsibility for evaluation, in practice evaluation governance is fully decentralized. The evaluation strategy from 2006 is no longer active. Organizational-level evaluation guidelines were published in 2002, but are not widely known or referenced. Two metiers have metier-specific strategies for evaluation (EcoSec and Protection-Detention). One metier has a set of metier-specific evaluation guidelines (EcoSec). These guidelines are in active use by EcoSec (not to be confused with the monitoring guidelines that all metiers have to support results-based management system). *Key issues include:*

1.1. *Weak performance on standards for institutional structures that ensure independent management of evaluations.* When fully benchmarked against standards for the 'existence of an independent organizational evaluation function', ICRC scores a failing grade of 41% (out of a possibility of 100%) (Health Check, Standard 1). Additionally, when reviewed for performance against Movement Principles of neutrality, impartiality, and independence as applied for evaluation, ICRC scores a 2 out of possible 4 points. There is currently no defined approach to 'independence' in the conduct of different types of evaluations and no defined approach to mitigating or eliminating potential conflict of interest during different types of evaluations, including those conducted by internal staff and/or managed by the file-holder of the action under review.

1.2. *Independent source of evaluation funds is not determined or guaranteed.* The source of evaluation funds (e.g. allocated from core funding, provided for from earmarked funding, etc.) is not determined in any guiding document. This can potentially compromise the budgetary independence of evaluations, if spending over-relies on earmarked funding (see Health Check criteria 1.5).

1.3. *Expectations for evaluation practice are unclear in the absence of an active ICRC evaluation strategy or organizational-level guidelines.* ICRC's original evaluation strategy and guidelines are no longer active. EcoSec and Protection-Detention have articulated metier-specific goals for evaluation, and EcoSec has a metier-specific set of evaluation guidelines. Beyond these two documents, expectations for evaluation practice are unclear and cited as a point of confusion that drives 'non evaluation' (see FGD Summary Report Annex III and Workshop Summary Report Annex IV).

FINDING 2: **Evaluation purpose and scope.** At the organizational level, definitions for evaluation purpose and scope are not documented according to industry standards for evaluation policies and frameworks (see Health Check criteria 2.1, 2.2., and 2.3, Annex V). The boundaries for what an evaluation will or will not cover are not consistently understood or applied by staff. At the report level, there is mixed performance on establishing the purpose, objectives, and scope of individual evaluations and like activities. *Key issues include*:

2.1 *Lack of organizational-level definitions on the distinct purpose and scope of evaluation versus other learning or accountability-oriented functions (Health Check Criteria 2.1).* While the 2006 Institutional Strategy for Independent Evaluations ('Evaluation Strategy') defines the purpose[5] and types of scope for 'independent evaluations' compared to other learning or accountability functions (e.g. Internal Audit), it is not in active use and the content is outdated. Although several of ICRC's accountability and learning functions have documented and defined parameters for the scope of their work that is relevant for creating distinctions with evaluation, no organizational document directly addresses the topic[6]. Staff sitting outside the accountability and learning functions highlighted this as point of confusion.

2.2 *Inconsistent understanding and application on the purpose and scope of evaluation vs. like activities.* There is no discernible distinction between the purpose and scope of 'reviews' and 'evaluations' as defined in the organization or applied in practice, despite widespread use of 'reviews' as a type of evaluation exercise (see Workshop Summary Report Annex IV and Evaluation Mapping Annex I). The Evaluation Strategy and the 2002 Evaluation Guidelines only address 'independent evaluations' and do not account for 'reviews' as a type of evaluation exercise. While the EcoSec Handbook for Planning, Monitoring, and Evaluation ('EcoSec Handbook') provides a degree of distinction between the intended purpose and scope of 'reviews' versus 'evaluations', the definitions are limited and not consistently applied in practice (e.g. in degrees of externality, in terms of timing, etc.). This is also true for reviews conducted by other metiers.

2.3 *Variable performance on the clarity and quality of the evaluation objectives and scope as defined in a report Terms of Reference (ToR) or Inception Report (IR).* Performance varies on whether report ToRs (1) include a statement on the evaluation purpose, the particular evaluation objectives, and the evaluation scope, and (2) whether the content included on these topics is well defined and positioned to support overall evaluation quality (see Case Study Summary Report Annex IIa). Use of an 'inception phase' to clarify evaluation scope and objectives with the evaluation team varied.

FINDING 3: **Evaluation identification and commissioning.** ICRC is not aligned with industry standards for the independent consideration and identification of an organizational evaluation plan (see Health Check criteria 1.3 and 2.4, Annex V). Evaluation identification occurs primarily during the annual planning cycle and according to a demand-led process. 'Prompts' to identify evaluation as a tool for results-based management (RBM), learning, or accountability are not consistently applied or included in the annual planning cycle. The process of identification and commissioning is largely driven by the metier units based in Geneva. *Key issues include*:

3.1 *Limited engagement on evaluation identification or commissioning from Delegations, or at the Departmental / Organizational level.* In the past 3 years, 56% of all evaluations and like exercises were commissioned by Geneva-based metiers (see Annex I). Only 20% were commissioned by the field delegations and the remaining 24% were commissioned at the departmental or organization level.

---

[5] The 2006 Evaluation Strategy defines the purpose of evaluation at ICRC as 'to inform strategic and operational decision-making, improve the effectiveness of humanitarian action through organizational learning, and to support accountability and transparency through the provision of performance information to partners, donors, and beneficiaries', page 3, Section 5.

[6] For 'learning' functions - this includes the 'Working Model and Definitions' paper for the Centre for Operational Research and Experience (the CORE), the Terms of Reference for the Division of Policy and Humanitarian Diplomacy (DP_POL), and the 'Project Management Framework (v3)' for Gateways 2 and 4 on 'lessons learned' exercises for all headquarters-based projects. Additionally, the Communications Analytics Unit is currently developing background documents to define the scope of research work covered by their team and the Security unit (SCMS) has a flow chart explaining the process for Security Lessons Learned exercises. For 'accountability' functions this includes the 'Internal Control Framework' and 'Internal Control Monitoring Guidelines (DRAFT)' of the Global Compliance Office and the guiding documents for Internal Audit and the annual AAP self-assessments for field delegations.

3.2 Reduced opportunities for teams to identify 'transversal evaluations' (e.g. 'transversal' files at HQ, integrated program responses in field delegations, cross-cutting thematic issues, collective impact measurement, etc.) (see Annex IV). The metier-oriented practice of evaluation identification and commissioning reduces opportunities for teams to identify evaluations on topics that cut across metier-specific categories, such as collective impact on a specific target group or in a certain location. This was a driving point for the recommendation that emerged from the OCP Project Consultation to 'recreate GENEVAL'.[7]

3.3 Reinforced evaluation practice according to staff familiarity and favourability. The demand-led and unprompted process for evaluation identification reinforces evaluation practice in areas of the organization with an established approach to considering evaluations as a tool for learning, accountability, and decision-making. It does not promote evaluation in areas of ICRC where staff are unfamiliar with or sceptical of evaluation. This is evident in the spread of evaluation commissioners. Of the metier-commissioned evaluations since 2016, 86% were identified and led by just one metier – EcoSec – while evaluation was largely absent from two out of the four programs – Protection and Movement. The detrimental impact of an unsupported process for evaluation identification on evaluation coverage was raised in 2010 in a report on evaluations conducted between 2006 and 2009[8]. During this Assessment, staff identified the lack of a prompted or guided process to consider evaluations during annual planning as a driver for 'non evaluation'.

FINDING 4: **Evaluation selection criteria.** Evaluation selection criteria[9] are not defined in the 2006 Evaluation Strategy or the 2002 Evaluation Guidelines. They are mentioned in the EcoSec Handbook (p78), without the label 'selection criteria'. In practice, staff use a range of industry-accepted selection criteria such as 'project closing', 'changing context', or 'problems faced / best practice identified'. In delegation-wide, departmental, or organization-level evaluations, selection criteria are largely 'strategic relevance', 'donor interest', or 'utility'. 'Risk' or 'Size of Operation' was not mentioned or observed as a selection criterion. *Key issues include:*

4.1 Selection according to 'timeliness' is a point of confusion for staff. During interviews and focus group discussions, staff in Protection, Movement, and Prevention noted that applying 'timing/timeliness' as a criterion for evaluation selection is confusing for metiers that work according to long time-horizons for results (e.g. 5 to 20 years). The lack of a clear start or anticipated end point for these métiers makes it difficult to identify when to time an evaluation to maximize its strategic utility. Additionally, the 'program' nature of field operations means that while they plan and budget in one-year cycles, actual activities are frequently designed and implemented over multiple years. This is notable in the absence of any 'mid-term' evaluations in the past 3 years and, instead, the use of 'process' evaluations (36% vs. 64% ex-post) that often occur when the action does not have an established end date. Aligning the different implementation calendars between Assistance vs. Protection, Movement, and Prevention was raised by staff as a challenge for multi-disciplinary evaluations.

4.2 Concerns about 'evaluability' are drivers of 'non evaluation'. Staff consistently raised the concern that key 'impact' questions are not evaluable (often expressed as 'it's impossible to answer this question') (see the KII/FGD Summary Report Annex III). This was most often noted on the question of whether ICRC 'prevented or minimized violations of international humanitarian law' or evaluating the degree to which ICRC 'prevented suffering'. Secondly, staff noted that context complexity can limit evaluability. In both cases, staff said 'evaluability' concerns are a primary driver for 'non evaluation'.

---

[7]See 'The 'OCP – outcomes of workshop – Jan / Feb 2019'

[8] 'Overview of Evaluations 2006-2009 as per evaluation strategy'. IPM Note to the Directorate on Independent evaluations, DIR1478 Annexe, 2010.

[9] Common evaluation selection criteria include: strategic significance, utility, evaluability, timeliness, commitments or requirements, risk or other triggers to evaluate (e.g. size and cost of the operation, signals of poor performance or needed course correction, closure or exit protocol).

FINDING 5: **Evaluation frequency and coverage.** ICRC performance is weakest for 'consistent, independent evaluation of results (coverage)' as defined by industry standards, scoring 29% in the Health Check (Annex V, Standard 2). Evaluation frequency is low and coverage is geographically and thematically imbalanced. This also applies to frequency and coverage by evaluation type (see Evaluation Mapping Annex I). The issue of low evaluation frequency and coverage of strategic-level evaluations was raised by DFID in the 2016 Multilateral Development Review (MDR), to which ICRC agreed an improvement was needed[10]. *Key issues include:*

5.1 *Lack of an active organizational policy and systems to ensure consistent evaluation coverage.* While the 2006 Evaluation Strategy describes requirements for coverage and frequency, it is no longer in active use. The EcoSec Handbook does not address coverage, frequency, or assurance procedures for these areas. There are no organization-level systems in place to track or ensure coverage, including the absence of a prioritized organizational evaluation plan and a lack of routine mapping of evaluation activity.

5.2 *Low comparative frequency for the size and scope of the organization.* On average, ICRC produces around six evaluations (including evaluation and like exercises) per year (2016 - 2019 to date), the majority of which (60%) are reviews or other evaluation-like exercises. By comparison, MSF (an organization of similar size and scope) conducted an average of 17 evaluations per year for the same period (2016-19)[11].

5.3 *Imbalanced geographic coverage between field and HQ/Global and across field delegations.* The majority of evaluations focus on field delegations (60%), while 32% of reports in the past three years looked specifically at actions occurring in headquarters (4 reports) or adopted a global theme (4 reports). Two reports (8%) had a regional focus (EcoSec Lake Chad Evaluation and Ebola Lessons Learned). A total of 19 countries where ICRC operates or operated experienced evaluation activity since 2016, almost half of which were in Africa (47.4%). Using the ICRC figure of presence in 69 field delegations[12], 19 countries indicates a 27.5% coverage rate for the entire 2016 through 2019 timeframe (not per year). The following 'key operations'[13] have not had an evaluation in the past three and a half years (if not longer, beyond the review period of this Assessment): CAR, Libya, Sudan, Ukraine, Jordan, Israel/Occupied Territories, Bangladesh, Syria, and Mexico. The following 'key operations' will be entering into a period where it has been over three years since evaluation activity by the end 2019: Somalia, Philippines, and Myanmar.

5.4 *Imbalanced coverage between 'metier' versus other operational dimensions (e.g. target population, emergency response operation, institutional policies or projects, etc.) and imbalanced coverage between Assistance and other metiers.* The majority of evaluations since 2016 focused on 'metier' themes (72%). Of the seven 'transversal' exercises (28%) included in this time period, only one report focused on a thematic topic spanning several metiers at the delegation level (Colombia, 2016). Assistance represents 78% of all 'metier' themed evaluations, dominated by EcoSec. Only one protection-focused evaluation, commissioned by a field delegation over three years ago that primarily examined HR issues, was identified in the mapping exercise. There have not been any Prevention or Communications evaluations that focused on work in field delegations or field programming. Organizational metiers with no evaluations in the past three years include Movement / Cooperation and Water and Habitation.

---

[10] See the 'Evaluation Function DIRGEN Action Plan' 2017

[11] https://evaluation.msf.org/reports/evaluation

[12] According to the 2020 PfR submissions, there are 69 field delegations where ICRC is active. Based on this, ICRC has a roughly 27.5% coverage rate, using the 19 countries with evaluation activity in the review period (including delegations that have multiple countries). It reflects the coverage rate for the existing number of delegations, not accounting for the possibility that there could have been more or fewer than 69 delegations over the years under review.

[13] Ibid.

FINDING 6: **Management of individual evaluations**[14]. ICRC performs slightly better on industry standards for the process of managing individual evaluations (e.g. 'ensuring evaluation quality'), than in other categories, scoring 63% in the Health Check. (Annex V, Standard 3). ICRC is not aligned with industry best practice for ringfencing a set allocation of operational budget or total organization expenditure towards evaluation (e.g. 0.07% UNHCR, 3-5% IFRC, UN JIU suggestion of 0.5-3%). Findings from the Case Study (Annex IIa) indicate that factors leading to a 'successful' evaluation include assurance of a 'participatory approach' and 'involvement of stakeholders' during the evaluation. An additional success factor includes ensuring the 'qualification and experience of evaluators' during the recruitment of the evaluation team. The lack of a defined or enforced protocol for evaluation management, however, makes it difficult for ICRC to ensure that these success factors are consistently addressed. *Key issues include*:

6.1 *Significant under-resourcing of evaluations for the size and scope of the organization*. Based on the Evaluation Mapping and the ICRC Budget Instructions[15], it is estimated that ICRC spends an average of 480,000CHF a year on evaluations. This is a generous estimate, as the cost per evaluation provided in the ICRC Budget Instructions cover rigorous evaluation conducted by external consultants, not the more frequently done and cheaper 'internal evaluations' or reviews. If ICRC were to follow the lowest funding benchmark in a similarly structured organization and aim for the standard used by UNHCR (0.07% of total expenditure), the amount would be 1.4 million CHF[16]. This indicates that, at minimum and being generous, ICRC is 192% below the lowest industry standard for evaluation resourcing.

6.2 *Inconsistently applied and inappropriate Terms of Reference (ToR) parameters by evaluation type*. ToRs are not consistently developed as part of the process for evaluations or like activity. When developed, ToR content – including evaluation scope, lines of inquiry, degree of intended externality vs. internality – is often not adjusted or appropriately matched for the methodological design or degree of desired independence for the type of evaluation exercise being conducted (e.g. external, mixed team, and internal / evaluations vs. reviews vs. lessons learned exercises). For example, ToRs for 'reviews' often include a scope and set of criteria that industry definitions would set at the 'evaluation' level.

6.3 *No defined or enforced protocol for required steps per evaluation type*. ICRC does not have organizational-level guidelines on procedures to follow when managing an evaluation or like activity. In practice, the following steps are inconsistently applied: creation of a committee to oversee the evaluation and approve deliverables, composition and storage of ToRs / Requests for Proposals (RFP), recruitment process and composition of evaluation teams (e.g. process for assessing required skills for the evaluation scope and type), inclusion of an IR in the evaluation set of deliverables, and requirements for documented evaluation management responses.

6.4 *No clearly defined or consistently practiced procedure to evaluation quality control*. Quality assurance parameters, including responsibilities for it and how it is provided, are not defined in current evaluation practice at ICRC. According to interviews, actual quality control practices employed by evaluation managers vary by the commissioning team. Evaluations supported by the PME Sector in 2017/2018 and those managed by teams with a longer experience in evaluation (eg EcoSec) receive better oversight compared to evaluations managed by teams with less experience or embedded expertise.

---

[14] Management of individual evaluations consists of the following parameters: budgeting, evaluation design and ToR creation, quality control, evaluation team recruitment and ongoing contract management, establishment of a steering or review committee, quality control procedures, establishment of required deliverables such as inception report and 'pre-findings' validation, management response formulation, and dissemination.

[15] ICRC conducts an average of 6 evaluations a year (based on evaluations mapped for 2016, 2017, 2018, and 2019). The budget instructions estimate 80,000CHF per evaluation. 6 evaluations a year x 80,000CHF an evaluation = an average of 480,000CHF /year on evaluations.

[16] Based on the 2018 ICRC budget of around 2 billion CHF

FINDING 7: **Evaluation content.** The content of evaluations and like activities are aligned with industry standards for the usage of OECD-DAC criteria. The application of criteria, however, strongly favour questions of relevance and effectiveness, with limited inclusion of criteria related to efficiency, impact, or sustainability / connectedness. Reports largely – though not completely - present methodological concerns, evidence, findings, conclusions, and recommendations in a balanced way (see Health Check criteria 3.3 and 3.4, Annex V). Three out of the four reports included in the Case Study received a 'strong' reading on 'quality' according to industry benchmarks on evaluation report content and as assessed by evaluation commissioners and users (see Annex IIa). The primary issues noted in the Case Study include 'missing objectives' in some reports (as compared to the ToRs / RFPs), limited visualization of findings, and overlong final report documents. *Key issues include:*

7.1 *Inconsistent use of appropriate methodologies for a given scope and evaluation type.* Parameters for expected methodological rigor are not defined or consistently applied according to evaluation type and scope. This is linked to issues noted in findings 1.2, 1.3, and 5.2 above. For example, the scope of the exercise's ToR may be suitable for an evaluation, but type of evaluation selected is an 'internal review' that applies methodologies appropriate for a review but not matched to the scope of the ToR.

7.2 *Lack of a clearly defined or applied approach to ensuring the inclusion of protection questions or concerns in evaluations.* ICRC does not meet the standard for 'ensuring accountability for actions taken to address protection concerns' across all evaluations (see Health Check criteria 3.9). It lacks a defined and consistently applied set of parameters for the inclusion of protection questions in every ICRC evaluation, including distinctions between protection results and protection mainstreaming, as defined for the evaluated action.

7.3 *Misunderstood or inconsistently interpreted definition of the DAC 'impact' criteria.* In staff interviews and focus group discussions, the issue of measuring impact was consistently raised as a concern. There was not, however, an established definition or interpretation of what is implied by 'impact' as judged by industry definitions. This includes particular confusion on whether 'impact' should be understood as the DAC definition of 'positive and negative changes produced by an intervention, directly or indirectly, intended or unintended...on the local social, economic, or environmental indicators' or as the primarily academic research definition for impact evaluation of 'the extent to which changes in outcomes of interest can be attributed to a particular intervention'[17]. The distinction is important, as most humanitarian organizations following the DAC definition focus on issues of longer-term, societal change and/or assessing unintended consequences without applying the type of methodology suited for measuring direct attribution to change. Frequently applied methodologies are more qualitative in nature, including contribution tracing or analysis, outcome mapping, and outcome harvesting. Organizations focused on the attribution definition, however, focus on quantitatively measuring a single or limited set of results on a direct beneficiary population (vs. the wider social context), which may occur from short to long term time horizons. This confusion is linked to finding 3.2 on concerns about the 'evaluability' of the impact question. Staff also struggled to distinguish between 'impact' and 'outcomes' for the purpose of deciding what is monitored versus what is evaluated.

7.4 *Lack of contribution-oriented approaches to measuring results achievement in protection programs.* There has never been the practice or the intent to explore contribution analysis in evaluating protection programs (see Health Check criteria 3.10). This approach is linked to finding 6.3 above. Methodologies framed to assess 'contribution to change' (vs. attribution) are well aligned with the DAC criteria definition for impact.

---

[17] Impact evaluations and development – Nonie Guidance on Impact Evaluation. http://siteresources.worldbank.org/EXTOED/Resources/chap4.pdf

FINDING 8: **Evaluation Utilization.** ICRC performance is low on 'ensuring response and follow up to and use of evaluation recommendations' and 'uptake of lessons learned and best practice' from evaluations as defined by industry standards, scoring 38% and 42% in the Health Check (see Annex V, overall scores on Standard 4 and 5). Case Study findings noted similar gaps in follow-up on recommendations and all survey respondents whose evaluation experience was 'unsuccessful' stated that it was due to 'insufficient follow-up'. *Key issues include:*

8.1 *Limited practice of including documented management responses and no systematic tracking of implementation of evaluation recommendations.* Documented management responses are not required in the Evaluation Strategy, Evaluation Guidelines, or in the EcoSec Handbook. In practice, 80% of evaluations mapped between 2016 and 2019 did not include a documented management response. Related to this, there is no systematic tracking of implementation of evaluation recommendations at any level of ICRC.

8.2 *Limited evaluation visibility and knowledge management.* ICRC does not have a complete and current evaluation report library or repository. With the addition of basic database and warehousing tools, existing IT infrastructure (namely, SharePoint and Tableau) supports a system where, in theory, evaluation reports and their annexes are searchable by keywords and could enable a 'top layer' dashboard for all filed evaluations. In practice, however, there is not a clear and consistent filing structure that is followed across the organization for evaluations (e.g. glossary of keywords, report location in the platform, owners of filing responsibility, etc). Even staff sitting within the same team / metier had different understandings of where to find evaluation reports in the SharePoint team space. This difference was most acute between HQ-based metier staff and metier staff in the field.

8.3 *Inconsistent practice on evaluation dissemination (internal and external).* ICRC does not have a consistent protocol for the way evaluations are shared internally or externally. This includes confusion around how – or whether - to share the report beyond the commissioning team, its wider metier, or immediate users of the evaluation (internally), as well as whether to publish reports on a public website, share evaluations that included a National Society partner with the IFRC, or share the report with donors (directly or via the donor extranet). According to the new Access to Information Policy (scheduled for launch in 2020), the Executive Summary for all independent evaluations will be published publicly. Implementing this policy would require definitions for the purpose and scope of 'independent evaluations' as compared to 'like' activities.

8.4 *Limited systematic application of evaluations for organizational learning.* Evidence on the application of lessons learned from evaluations is mixed, where available, and largely based on staff perception. There is no documented, systematic protocol for the application of evaluation learning in organizational policy making, strategy formulation, or planning. Evaluation trends, findings, and recommendations are not embedded in existing moments of the organization calendar at delegation and headquarter levels that would benefit from a dedicated review of evaluation learning, such as annual seminars, strategy development, annual planning, revision of policies, etc. There is also no practice or history of producing synthesized 'meta evaluation' learning reviews on key trends in evaluation findings and recommendations, as commonly practiced in other humanitarian organizations.

8.5 *Limited systematic application of evaluations for accountability.* Concrete ways in which evaluations can be systematically applied for internal, external, and beneficiary directed accountability are not defined in any guiding document nor actively practiced. This includes a lack of clarity on ways evaluations can complement audit findings and the role of evaluations in the AAP framework. There is no practice of sharing evaluation findings with the Assembly or senior leadership as part of the organization governance.

B. Desired Status (To Be)

This section identifies ICRC objectives for evaluation performance as indicated by the 2019-2022 Strategy and as expressed by ICRC staff during the Assessment and within key organization documents.

Application of ICRC 2019-2022 Strategy. *Expectations for and the practice of evaluation at ICRC enables organizational learning and greater operational impact. Evaluations connect with accountability to affected populations, partnership, programmatic integration, and multi-year programming in protracted crisis. Evaluations are managed in a way that is consistent with ICRC's mission statement and Fundamental Principles.*

Systems for evaluation identification, planning, and management support evaluations conducted with partners, evaluations that investigate 'transversal' program questions, and evaluations that assess longer-term impact.

Objectives expressed by ICRC staff and other key documents. Organizational performance objectives articulated by staff during the Assessment or as documented by ICRC[18]. Information source is in brackets:

Governance. (Re)create an evaluation function that fosters an enabling environment for evaluation identification, production, and use. This includes aligning with industry evaluation standards in a manner that is adapted to the structure, role, and ambition level of ICRC. Evaluation function has capacity to manage and leverage evaluations to respond to ICRC information needs. (Workshop, Interviews, Document Review)

Evaluation Purpose and Scope. The learning and accountability purposes for evaluations are clearly defined and positioned to succeed. There is a defined set of potential evaluations, including transversal / inter-disciplinary / multi-delegation regional scope. This is supported by vision, awareness, resources, and methods. (Workshop, Document Review)

Selection Criteria. Staff concerns about how to apply 'timeliness' or address issues of 'evaluability' for the specific operational structure of ICRC metiers are addressed through technical support on the evaluation process. Evaluation timing is defined in a multi-year evaluation plan (Interviews, Focus Groups)

Evaluation Coverage. Evaluation coverage improves to include metiers with limited or absent evaluation (e.g. Protection, Movement, etc.). Evaluating Protection includes protection-specific work and the protection benefits or mainstreaming efforts of non-protection metiers. Delegation-wide and thematic evaluations of strategic importance to ICRC are identified and commissioned. (Interviews, Focus Group Discussions, Document Review)

Evaluation Management. ICRC follows a clear and consistent process for evaluation management, including defined roles and responsibilities and the application of quality criteria. (Workshop)

Evaluation Utilization. Evaluations are visible, accessible, promoted, disseminated, and used for both internal and external purposes in a manner that is consistent with ICRC policies for Data Protection and Access to Information. Evaluations are positioned to feed into the analysis of Strategy Monitoring (SM). (Interviews, Workshop)

A proposed set of industry standards for evaluation performance that ICRC should prioritize based on these objectives is found in Section 4, 'Recommendations'.

---

[18] See the 'Evaluation Function DIRGEN Action Plan' 2017, the 'OCP – outcomes of workshop – Jan / Feb 2019', and the DIR2541 Annexe 1, 'Institutional Monitoring Framework '.

C. Gaps Analysis

This section presents a summary table (Table 4) of key issues identified in the As Is findings of the Assessment on organizational performance. They are matched against their most relevant To Be objectives as described above in Section B. The table presents the primary gaps to be addressed and where this links to the recommendations.

The Gaps Analysis focuses on priority recommendations for establishing an evaluation function. Additional actions that are required in a second or third phase of a phased approach to establishing an evaluation function are addressed within the proposed Evaluation Strategy, under a proposed timeline for actions in Years 1-2, 3-4, and 5-beyond.

Table 4: Organizational Performance: *Gaps Analysis*

Existing Practice (As Is)	Desired Status (To Be)	Primary Gaps: *What is missing*	Linked Recommendation
Finding 1: Evaluation Governance			
Weak performance on standards for institutional structures that ensure independent management of evaluations.	(Re)create an evaluation function that aligns with Movement Principles of neutrality, impartiality, and independence. Evaluation function has capacity to manage and leverage evaluations to respond to ICRC information needs.	Defined evaluation function, including scope and location. Defined evaluation framework and accountabilities for standards in the evaluation strategy.	**Recommendation 2:** Purpose and structure of the evaluation function **Recommendation 6:** Evaluation Strategy
Independent source of evaluation funds is not determined or guaranteed.	(Re)create an evaluation function that fosters an enabling environment for evaluation identification, production, and use. Function structure and scope aligns with Movement Principles of neutrality, impartiality, and independence.	Defined approach to securing funds for evaluation, including budget holding arrangements, budget source, and budget amount.	**Recommendation 3:** Resourcing the evaluation function
Expectations for evaluation practice are unclear in the absence of an active ICRC evaluation strategy or organization-level guidelines.	Develop an evaluation strategy that defines the purpose and scope of evaluation and presents a framework for different categories of evaluation. Strategy includes position of evaluation topics (e.g. programme, strategic, thematic, etc.) according to centralized or decentralized management.	Defined priorities for evaluation practice and standards. Defined and actively applied evaluation strategy, including a clear evaluation framework.	**Recommendation 1:** Proposed application of industry standards **Recommendation 6:** Evaluation Strategy
Finding 2: Purpose and Scope			
Lack of organization-level definitions on the distinct purpose and scope of evaluation vs other like functions.	The learning and accountability purposes for evaluations are clearly defined and positioned to succeed. A comprehensive evaluation strategy defines the purpose and scope of evaluation as distinct from other learning and accountability-oriented functions.	Clarification on the role of evaluation as compared to other learning and accountability functions (e.g. Internal Audit)	**Recommendation 6:** Evaluation Strategy

Inconsistent understanding and application on the purpose and scope of evaluation vs. like activities.	Develop an evaluation strategy that presents a framework for different categories of evaluation. Strategy includes position of evaluation topics (e.g. programme, strategic, thematic, etc.) according to centralized or decentralized management.	Defined purpose and scope of evaluation vs. like activities, including levels of centralized or decentralized management.	**Recommendation 6:** Evaluation Strategy
Variable performance on the clarity of evaluation purpose and scope as defined in the ToR or Inception Report.	Develop an evaluation strategy that defines the purpose and scope of evaluation. A defined set of potential evaluations is supported by: vision, awareness, resources, and methods.	Defined purpose and scope of evaluation vs. like activities. Defined responsibilities for technical support and quality assurance.	**Recommendation 6:** Evaluation Strategy **Recommendation 2:** Purpose and structure of the Evaluation Function
Finding 3: Identification and Commissioning			
Limited engagement on evaluation identification or commissioning from Delegations or at the Departmental / Org level.	Systems for evaluation identification support evaluations conducted with partners, evaluations that investigate 'transversal' programming models and questions, and evaluations that respond to questions of longer-term impact. Develop a prioritized evaluation plan that presents systematic coverage of the ICRC.	Defined and enforced responsibilities for evaluation identification and planning. Defined and proactively supported process for creating an organizational-level evaluation plan.	**Recommendation 6:** Evaluation Strategy **Recommendation 2:** Purpose and structure of the evaluation function
Reduced opportunities for teams to identify 'transversal' evaluations			
Reinforced evaluation practice according to staff familiarity and favourability.			
Finding 4: Selection Criteria			
Evaluation selection criteria are not defined.	Selection criteria and use of OECD/DAC criteria for evaluation questions are addressed in the evaluation strategy.	Stated expectations for the type of selection criteria applied by ICRC for evaluation planning. Support on the use of criteria in the process of identifying and selecting evaluations.	**Recommendation 6**: Evaluation Strategy **Recommendation 2:** Purpose and structure of the evaluation function
'Risk' or 'Size of Operation' was not mentioned or observed in reports as an applied or considered selection criteria.	Selection criteria and use of OECD/DAC criteria for evaluation questions are addressed in the evaluation strategy.	Stated expectations for the type of selection criteria applied by ICRC for evaluation planning.	**Recommendation 6:** Evaluation Strategy

Selection according to 'timeliness' is a point of confusion for staff.	Staff concerns about how to apply 'timeliness' for the specific operational structure of ICRC metiers are addressed in ICRC evaluation guidelines or through technical support. Evaluation timing defined on a multi-year evaluation plan	Available technical expertise to guide the use of criteria when identifying and selecting evaluations. Defined and proactively supported process for creating an organizational-level evaluation plan.	**Recommendation 3:** Resourcing the evaluation function **Recommendation 2:** Purpose and structure of the evaluation function
Concerns about 'evaluability' are drivers of 'non evaluation'.	Staff concerns about how to address issues of 'evaluability' are addressed in ICRC evaluation guidelines or through technical support on the evaluation process.	Available technical expertise to guide the use of criteria when identifying and selecting evaluations.	**Recommendation 3:** Resourcing the evaluation function
Finding 5: Frequency and Coverage			
Lack of policy and systems to ensure evaluation coverage.	The strategy describes requirements and assurance mechanisms for coverage. ICRC develops a prioritized evaluation plan that presents systematic coverage of the organization. Application of partnership, integration, and multi-year programming reflected in the practice for determining evaluation coverage. Delegation-wide and thematic evaluations of strategic importance to ICRC identified and commissioned. Evaluation coverage includes metiers with limited or absent evaluation engagement.	Defined coverage expectations or requirements. Defined and proactively supported process for creating an organizational-level evaluation plan.	**Recommendation 6**: Evaluation Strategy **Recommendation 2:** Purpose and structure of the evaluation function
Low comparative frequency for the size and scope of the organization.			
Imbalanced geographic coverage			
Imbalanced coverage between 'metier' vs other operational areas / between Assistance and other metiers.			
Finding 6: Management of individual evaluations			
Significant under-resourcing of evaluations for the size and scope of ICRC.	Ensure a dedicated and independent budget line for centralized and decentralized evaluations.	Defined approach to securing funds for evaluation and the evaluation function, including budget amount and budget holding.	**Recommendation 3:** Resourcing the evaluation function
Inconsistently applied and inappropriate ToR parameters by evaluation type.	Develop an evaluation strategy that presents a framework for different categories of evaluation. The strategy describes quality assurance requirements and mechanisms for centralized and/or decentralized evaluations.	Defined and actively applied evaluation strategy, including a clear evaluation framework. Defined responsibilities for technical support and quality assurance.	**Recommendation 6:** Evaluation Strategy **Recommendation 2:** Purpose and structure of the Evaluation Function
No defined or enforced protocol for required steps per evaluation type.	ICRC follows a clear and consistent process for evaluation management, including defined responsibilities and the application of quality criteria.	Defined and actively applied evaluation strategy, including key steps in evaluation management.	**Recommendation 6:** Evaluation Strategy
No clearly defined or consistently practiced procedure to evaluation quality control.	The strategy describes quality assurance requirements and mechanisms for centralized and decentralized evaluations.	Defined responsibilities for technical support and quality assurance. Creation and use of systems to support quality assurance for evaluation management.	**Recommendation 6:** Evaluation Strategy **Recommendation 5:** IT systems and tools

Finding 7: Evaluation Content			
Inconsistent use of appropriate methodologies for a given scope and evaluation type.	Develop an evaluation strategy that presents a framework for different categories of evaluation.	Defined purpose and scope of evaluation activities. Defined responsibilities for technical support and quality assurance.	**Recommendation 6:** Evaluation Strategy **Recommendation 2:** Purpose and structure of the Evaluation Function
Lack of a clearly defined or applied approach to inclusion of protection questions in evaluations.	Evaluations of ongoing or completed programs ensure accountability for actions taken to address protection concerns. Evaluations include questions on the protection benefits or mainstreaming efforts of non-protection metiers.	Defined intent for evaluations to include standard lines of inquiry relating to protection.	**Recommendation 6:** Evaluation Strategy
Misunderstood definition of the DAC 'impact' criteria.	Develop a comprehensive evaluation strategy that draws on international best practice. A defined set of potential evaluations is supported by: vision, resources, and methods.	Clarified meaning of 'impact'. Available expertise to guide use of 'impact' criteria.	**Recommendation 6:** Evaluation Strategy **Recommendation 3:** Resourcing the function
Lack of contribution-oriented approaches to measuring results achievement in protection programs.	Evaluations of protection programming focus on issues of contribution (vs. attribution) to a desired result(s).	Defined intent for protection evaluations to include contribution approaches for measuring results. Available expertise to guide use of contribution-oriented methodologies.	**Recommendation 6:** Evaluation Strategy **Recommendation 3:** Resourcing the evaluation function
Finding 8: Evaluation Utilization			
Limited practice of documented management responses and no systematic tracking of implementation of recommendations.	Ensure evaluation and like activities include a documented management response (MR) with clear responsibilities and a timeline for implementation of recommendations.	Defined requirement for documented management responses for evaluations and like activities. Creation and use of systems to track implementation of recommendations.	**Recommendation 6:** Evaluation Strategy **Recommendation 5:** IT systems and tools
Limited evaluation visibility / knowledge management.	Evaluations are visible, accessible, disseminated, and used in a manner consistent with ICRC policies.	Creation and use of evaluation report library.	**Recommendation 5:** IT systems and tools
Inconsistent practice on evaluation dissemination	Clarify expectations for report dissemination in a manner consistent with ICRC policies.	Defined process and responsibilities for internal vs. external report dissemination.	**Recommendation 6:** Evaluation Strategy
Limited systematic application of evaluations for organizational learning.	Evaluations are positioned to support organizational learning.	Available technical expertise to promote the use of evaluation at all levels of ICRC.	**Recommendation 3:** Resourcing the evaluation function
Limited systematic application of evaluations for accountability.	Evaluations feed into strategy monitoring analysis and reporting. Evaluations support organizational commitments to accountability to affected populations.	Available systems to enhance evaluation use.	**Recommendation 5:** IT systems and tools

II. Organizational Capacity

This section explores ways organizational capacity drives evaluation performance. It reviews the capacity of ICRC to produce and utilize evaluations *as required to meet organizational evaluation performance goals and strategic objectives* outlined above in Section I.

A. Existing Practice (As Is)

FINDING 9: **Evaluation leadership.** Evaluation leadership is fragmented in the absence of a defined evaluation function. There are no guiding documents locating responsibility for leadership on evaluation culture or evaluation performance and staff consistently cited 'weak leadership and ownership on evaluations' as a driver for 'non evaluation' (see Annexes III and IV). While the PME Sector has a loosely defined responsibility for evaluations – supporting evaluations as requested by the commissioning teams and promoting the identification and use of 'transversal' evaluations in the past three years - it is not referred to as a leading authority for evaluation practice nor is it capacitated for this role at present. In practice, evaluation leadership is primarily demonstrated by select Geneva-based metiers with an existing culture of and favourability towards evaluation. EcoSec, in particular, is a champion of evaluation practice. *Key issues include:*

9.1 *'Interest driven' evaluation leadership reinforces evaluation practice according to staff familiarity and favourability.* In the absence of an evaluation function driving evaluation performance across ICRC, the actual practice of evaluation leadership emerges in areas of the organization with an established evaluation culture – for example, in EcoSec. Structurally, there are few opportunities for these pockets of evaluation leadership to drive evaluation outside their metier. Linked to finding 3.3 under Section I, the effect is to reinforce evaluation where evaluation is already present while leaving the rest of the organization, where staff may be unfamiliar with or sceptical of evaluation, without direction.

9.2 *Opaque roles and responsibilities around evaluation initiation or participation (e.g. steering committee member, etc.).* Responsibilities for evaluation leadership are not defined in guiding documents (e.g. job descriptions, annual planning procedures) for individual positions (e.g. delegation management such as Head of Delegation, Deputy Head of Delegation, etc.) nor in role mandates for organizational levels, departments, or units. Practice and understanding varies between units, organizational levels (Geneva vs. field delegations), and organizational roles (red line vs. blue line responsibility) on who can or should lead evaluation practice, when it is appropriate for them to do so, and from whom they must inform or secure approvals.

9.3 *Limited promotion or encouragement of evaluation at the management level.* Senior leaders at ICRC are engaged with evaluation at the strategic level of (1) supporting its placement in the 2019-2022 strategy, (2) supporting the (re)creation of the evaluation function, and (3) messaging the need for improved evaluation performance at ICRC in response to the organization's commitment to learning and accountability, the changing context of ICRC operations (including increased multi-year programming in protracted crises), and demonstrating added value to remain competitive in a landscape of shrinking donor budgets. ICRC staff, however, raise concerns that evaluation is not visibly promoted at the management level in practice (see Workshop Summary Report Annex IV). This is evident in the relatively few evaluations commissioned directly by Head of Delegations at the field level (2 reports since 2016) or directly identified by senior leadership in Geneva as a priority (Annex I). This is also evident from the Case Study, where survey participants listed 'lack of stakeholder interest before or after the exercise' as the cause for evaluation 'failure' by 60-80% of respondents across cases (Annex IIa). One explanation provided by staff: stakeholders are not given the signal or space from their managers to make evaluation a priority (Annex IV).

FINDING 10: Staffing and staff capacity for evaluation. There are no fully dedicated evaluation positions active at ICRC. Positions with evaluation responsibilities are combined with planning, monitoring, or research and analysis responsibilities. Positions are largely decentralized and are few in number. The Head of the PME Sector is presently the only person in that sector providing internal and external engagement on evaluation issues. The decision to create M&E / RBM positions elsewhere in the organization is demand-led. Without a centralized, well-equipped support function for evaluation, metiers without their own dedicated M&E/ RBM staff are left without any guiding expertise on the topic. Among ICRC staff more generally, expertise on evaluation is limited and inconsistently present within the metiers. The majority of ICRC staff are not exposed to or familiar with the range of evaluation methodologies available to address difficult questions (e.g. 'what is our collective impact in a field delegation?') or respond to challenging evaluation contexts (e.g. situations without not clear performance benchmarks against which to evaluate or common objectives to anchor a transversal evaluation). Staff note that confusion about how to proceed in these situations is a primary driver of 'non evaluation'. *Key issues include:*

10.1 *Inefficient use and insufficient amount of human resources for the work required to produce and use evaluations (staff time and availability gaps, with consequences for efficiency).* In the absence of a centralized evaluation function, staff commissioning evaluations are tasked with the entire workload of evaluation management in addition to their existing responsibilities. This often prolongs what should be a relatively straightforward process (e.g. lengthy timelines around producing a ToR / RFP) (Case Study Summary Report Annex IIa). These siloed efforts are then duplicated in every metier that commissions an evaluation. Rather than working from common templates, staff holding multiple responsibilities spend a significant amount of time 'reinventing the wheel' on key evaluation steps. Staff note 'limited capacity and available time' as a key driver for non-evaluation. Notably, the metier with the most established practice – EcoSec – seems to produce evaluations more efficiently compared to other sections of ICRC.

10.2 *Confusion and fragmentation between existing M&E and RBM staffing (efficiency gap, with consequences for analytical capacity).* The decentralized M&E and RBM staffing at ICRC do not operate with a blue and/or dotted line to the PME Sector. They operate according to the requirements shaped by their hiring manager or within their team. While this can enhance the 'utility' of the position by ensuring it is crafted for the needs of the hiring team, it also creates the following challenges in the absence of a staffing structure that coherently aligns with a 'way of working' on evaluation:
- Fragmentation in vision, language / definitions, and application of evaluation standards
- Confusion on roles and responsibilities of staff with overlapping job titles and descriptions
- Siloed activity leading to disjointed datasets, limiting response-, thematic-, or delegation-level analysis
- Inequity of M&E capacity between teams, leading to inconsistent quality of evaluation products

These challenges will increase as ICRC moves towards joint results measurement, integrated program models, and multi-year programming with a desire to investigate or measure collective impact.

10.3 *Inconsistent understanding of how to shape evaluation for ICRC's programmatic model and according to ICRC's planning tools (knowledge and expertise gap, limiting demand for evaluation).* Related to findings 4.1 and 4.2, staff are unclear how to structure or apply evaluation for ICRC's program models. This includes identifying the right plan or set of commitments against which to evaluate at both the sub-programme level and for transversal topics that may not have an articulated performance framework against which to base judgements of 'effectiveness' or 'impact' (e.g. evaluating collective impact on a target population if there is no defined common objective). This is primarily an issue within Movement, Prevention / Communication, and Protection metiers. It also affects ICRC's ability to articulate a clear path forward for commissioning delegation-wide, regionally or globally thematic, or otherwise transversal evaluations.

FINDING 11: **Funding and financially resourcing evaluations.** ICRC does not maximize its funding structure for evaluating transversal issues and longer-term impact. ICRC's approach to evaluation funding does not proactively ensure funds are available and adequate in amount to conduct evaluations in the manner implied in the 2019-2022 strategy or at a standard that is commensurate to the size and function of the organization. Existing practice reinforces an overall weak and fragmented evaluation culture. *Key issues include:*

11.1 *Budget opportunities are not incorporated into budgeting process.* Within the core budget, ICRC is not beholden to project-based funding cycles or requirements[19]. It can allocate money towards evaluation at any level, covering any topic, revisiting any program or area of intervention, covering any period of time as deemed beneficial to the organization. This advantageous position, however, is not reflected in the current process of funding evaluations.

11.2 *Process for planning budgets reduces evaluation identification.* Evaluation budgets are determined and secured at the time of evaluation planning, frequently during the annual planning cycle. ICRC's process for evaluation planning and budgeting is entirely demand-led, reinforcing 'non-evaluation' in areas of the organization that are unfamiliar with or sceptical of evaluation. This limits the coverage and type of evaluations commissioned at ICRC (see Section I, Finding 5). This is linked to Findings 3.3 in Section I and 9.1 above. Additionally, the metier-oriented structure of evaluation budgeting reduces opportunities for teams to identify or fund 'transversal' evaluations (e.g. thematic files at HQ, integrated program responses in field delegations, collective impact measurement etc.).

11.3 *Process for securing budgets creates uncertainty and undermines evaluation commitments.* Evaluation budgets are largely allocated from within the commissioning unit's budget and must be approved by the relevant manager. Approval for funding is not guaranteed. Approved funds can later be reallocated or cut by budget holders according to interviews. The ability to reallocate evaluation funds and the case-by-case approval process creates competition between funding evaluations and funding other priorities. For Field Operations this is particularly acute: evaluations will not be prioritized if pitted against - instead - funding life-saving humanitarian assistance. The uncertainty of funding approval de-incentivizes interest and limits capacity to evaluate long-term impact, which may require consistent funding and staff commitment over several years of data collection and analysis. This is especially true if the evaluation requires revisiting an area where a program closed and the learning is not of immediate interest to the budget holder.

11.4 *Limited ability to track evaluation spending.* Although guidance for costing evaluations exist in the ICRC Budget Instructions, evaluations do not have a dedicated budget code in the accounting architecture. This makes it impossible to accurately track the level of evaluation spending across the organization[20].

FINDING 12: **IT systems and initiatives.** ICRC IT systems and initiatives are not yet targeting evaluations in depth and overall knowledge management for evaluations is weak (see Finding 8, Section I). There is, however, significant investment in digital transformation at ICRC and existing platforms could be leveraged to improve and promote evaluation management and use. *Key issues include:*

---

[19] e.g. in project-fixed funding, the availability of funds for evaluation ends when the project ends. Funds are also often tied to the project and cannot be used outside it. This limits the ability of an organization to return to a closed project and review longer term impact, as well as measure impact on issues sitting above the project level (e.g. integrated program effects or collective impact).

[20] Evaluations working with consultants will use the consultancy budget line, which can provide a limited degree of visibility. It does not provide a way to monitor budgets for internal reviews or internally conducted evaluations (e.g. where consultants are not hired). Likewise, consultants may be hired for a range of additional purposes, which are then included in the budget analysis on the consultancy line.

12.1 With the addition of relatively simple database and warehousing tools, ICRC's two primary platforms - SharePoint and Tableau – can be leveraged to improve evaluation practice. Three primary areas of IT application are identified for evaluations at ICRC:

- **Management of individual evaluations**. This covers the process of managing centralized or decentralized evaluation workflow.
- **Evaluation function operations.** This covers the fulfilment of responsibilities by a centralized evaluation function charged with steering ICRC's evaluation strategy, organizational-level guidelines, and overall quality control of evaluation practice. It also involves direct management of centralized evaluations.
- **Evaluation use** (*beyond individual evaluation response*). This covers the use of data on and from evaluations in a range of functional areas and by positions across the organization.

The development and maintenance of data warehousing tools that link Tableau and SharePoint, plus configuration of SharePoint would be required for leveraging data across evaluations for evaluation function operations and improved evaluation use. Configuration of SharePoint would also be necessary for its use in improving the management of individual evaluations.

Details and a full body of recommendations is provided under Recommendation 5, Section 4 and Annex VIII 'Information Technology and Evaluation Recommendations Report'.

FINDING 13: **Incentive or compliance systems to promote evaluation production and use.** There are no active incentive or compliance systems focused on evaluations, nor is evaluation embedded in larger organizational systems that could promote evaluation practice. This is related to findings in Section I of the Assessment that ICRC does not currently define or enforce requirements for evaluation. There are, however, existing mechanisms that could promote evaluation practice *if* evaluations were integrated into their operations. *Key issues include:*

13.1 Incentive or compliance opportunities for evaluation practice are not incorporated into evaluation planning, management, or use. A range of incentive and compliance mechanisms exist across ICRC. Evaluations, however, are not embedded within these systems. This includes missed opportunities to include expectations or requirements for evaluation practice within the protocol of Internal Audit, Assembly and Directorate meetings, staff performance management, the PfR, archives, and the AAP self-assessments. Details on ways this can be done are included in the proposed Evaluation Strategy, Annex IX.

13.2 Weak incentive and compliance systems for IT procedures will impact ability to leverage IT for improved evaluation management and use. Existing compliance with document management procedures is reportedly and observably low for evaluations and more broadly across the organization. This could impact ICRC's ability to leverage its IT platforms (see Finding 12) for improved evaluation management and use, with particular consequences for establishing an evaluation report library, securing visibility over evaluation activity at multiple levels of operation, and systematizing quality assurance. The Assessment did not find any plans to improve IT compliance that could be linked to this concern.

B. Desired Status (To Be)

This section identifies objectives for organizational capacity as indicated by the 2019-2022 Strategy and as expressed by ICRC staff during the Assessment and within key documents.

Application of ICRC 2019-2022 Strategy. *Each element of organizational capacity for evaluation is positioned to enable the use of evaluation for organizational learning and greater operational impact. ICRC strengthens its capacity to address strategic questions on partnership, programmatic integration, and multi-year programming.*

ICRC strengthens its capacity to evaluate the outcomes of its activities and learn from its successes and failures. Evaluation leadership champions the purpose and value of evaluations in this regard. The structure for evaluation staffing and, more broadly, ICRC staff capacity, together with the structure for evaluation funding, enable evaluations that (1) address questions of partnership, programmatic integration, and multi-year programming in protracted crisis and (2) leverage a variety of evaluation methods. ICRC embeds evaluation in its planning and RBM systems. Existing mechanisms that support learning and accountability and ICRC IT systems are leveraged to improve the production of evaluations that meet ICRC's standards and expectations.

Objectives expressed by ICRC staff and other key documents. Organizational capacity objectives articulated by staff during the Assessment or otherwise documented by ICRC. Information source is in brackets:

Evaluation leadership. A defined set of evaluations (including transversal, inter-disciplinary, multi-delegation regional scope, etc.) is supported by organizational vision for why ICRC evaluates and how ICRC evaluates. Manage 'territoriality' within metiers or delegations through sustained evaluation leadership. (Workshop).

Evaluation staffing and staff capacity for evaluation. People access necessary evaluation expertise. Scoping of evaluation responsibilities at all levels of the organization; and the capacity to fulfil demands for technical expertise at all levels. Evaluation function supports evaluation practice in areas of the organization that are unfamiliar with evaluation. Evaluation identification and planning process is proactively supported. Mechanism for quality assurance backstops areas of ICRC where evaluation capacity is weak. (Workshop, FGDs, Interviews)

Evaluation funding. Evaluation function supports learning-oriented evaluation in areas of the organization that are unfamiliar with it. Increase visibility of evaluation activity and spending through the evaluation budgeting process. Incentivize interest in evaluation through budget structure and availability. (Workshop, Interviews)

IT systems and initiatives. Leverage existing IT tools to produce visuals based on data about and from evaluations. Articulated expected information flow (evaluation visibility and use) between delegations, métier, central evaluation function, and the wider institution. Ensure a strong change management process for IT solutions. (Workshop, Interviews)

Incentive or compliance systems to promote evaluation. Integrate evaluation into existing organizational systems that promote learning and accountability. Evaluation is considered at the start of programming planning. Evaluation system addresses support on outcome monitoring / measurement gaps, development of a management response, action on management response recommendations, and integration of evaluation feedback into ICRC ways of working. (Workshop, FGD, Interviews)

C. Gaps Analysis

This section presents a summary table (Table 5) of key issues identified in the As Is findings of the Assessment on organizational capacity matched against their most relevant To Be objectives (Section B). It highlights the primary gaps to be addressed and where this links to the recommendations.

The Gaps Analysis focuses on priority recommendations for establishing an evaluation function. Additional actions that are required in a second or third phase of a phased approach to establishing an evaluation function are addressed within the proposed Evaluation Strategy, under a proposed timeline for actions in Years 1-2, 3-4, and 5-beyond.

Table 5: Organizational Capacity: *Gaps Analysis*

Existing Practice (As Is)	Desired Status (To Be)	Primary Gaps: *What is missing*	Linked Recommendation
Finding 9: Evaluation Leadership			
Interest-driven evaluation leadership reinforces evaluation practice according to staff familiarity and favourability.	Leadership champions the purpose and value of evaluations. Evaluation identification and planning is proactively supported. Evaluation function supports evaluation in areas of ICRC that are unfamiliar with or skeptical of evaluation.	Defined evaluation function, including its scope for evaluation leadership and its location relative to organizational leadership / governance.	**Recommendation 2**: Purpose and scope of the evaluation function
	A defined set of evaluations is supported by organizational vision for why ICRC evaluates and how ICRC evaluates.	Defined purpose of evaluation, including levels of centralized or decentralized management.	**Recommendation 6**: Evaluation Strategy
Opaque responsibilities around evaluation initiation or participation.	A defined set of evaluations is supported by organizational vision for why ICRC evaluates and how ICRC evaluates.	Defined and enforced responsibilities for evaluation identification, planning, and process participation.	**Recommendation 6**: Evaluation Strategy
Limited promotion or encouragement of evaluation at the management level	ICRC leadership consistently champions the purpose and value of evaluations for learning from successes and failures.	Defined evaluation function, including its scope for evaluation leadership and its location relative to organizational leadership.	**Recommendation 2**: Purpose and scope of the evaluation function
		Defined purpose of evaluation, including levels of centralized or decentralized management.	**Recommendation 6**: Evaluation Strategy
Finding 10: Evaluation Staffing and Staff Capacity for Evaluation			
Inefficient use and insufficient amount of human resources for the work required to produce and use evaluations	People access necessary expertise on evaluation. Evaluation function supports evaluation practice in areas of the organization that are unfamiliar with or skeptical of evaluation.	Sufficient and effectively coordinated human resources to support evaluation practice.	**Recommendation 3**: Resourcing the evaluation function
		Defined evaluation framework, including 'light' vs. more robust evaluation types.	**Recommendation 6**: Evaluation Strategy

Confusion and fragmentation between existing M&E and RBM staffing.	The structure for evaluation staffing and, more broadly, ICRC staff capacity enables evaluations that address questions of partnership, programmatic integration, and multi-year programming. Clear scope of evaluation responsibilities at all levels of the organization.	Defined staffing requirements and associated responsibilities for the evaluation function, including connectivity with existing M&E and RBM staffing	**Recommendation 3:** Resourcing the evaluation function
Inconsistent understanding of how to apply evaluations within ICRC's programmatic model and according to ICRC's planning tools.	The structure for evaluation staffing and, more broadly, ICRC staff capacity enable evaluations that address questions of partnership, programmatic integration, and multi-year programming and leverage a variety of evaluation methods. ICRC strengthens its capacity to evaluate outcomes. Evaluation function supports evaluation in areas of ICRC that are unfamiliar with or skeptical of evaluation.	Proactive and available technical expertise to guide evaluation practice according to ICRC program models / planning tools.	**Recommendation 3:** Resourcing the evaluation function
Finding 11: Funding Evaluations			
Budget opportunities are not incorporated into budgeting process	The structure for evaluation funding enables evaluations that address questions of outcomes/impact, partnership, programmatic integration, and multi-year programming.	Approved budgeting process that reflects opportunities within ICRC funding structure, inclusive in the budget holding arrangements, budget source, and budget amount.	**Recommendation 3:** Resourcing the evaluation function
Process for planning budgets reduces evaluation identification.	Incentivize interest in evaluation through evaluation budget structure and budget availability.	Approved budgeting process that secures funding availability to all potential evaluation commissioners.	**Recommendation 3:** Resourcing the evaluation function
Process for securing budgets creates uncertainty and undermines evaluation commitments.			
Limited ability to track evaluation spending.	Increase visibility of evaluation activity and spending through evaluation budgeting architecture, e.g. securing budget code for evaluation and running queries on the use of that code.	Defined budgeting requirements for the evaluation function, including visibility of evaluation spending.	**Recommendation 3:** Resourcing the evaluation function
Finding 12: IT Systems and Initiatives			
Existing platforms - SharePoint and Tableau - can leverage evaluation practice.	ICRC IT systems are leveraged and built upon to improve the production and use of evaluations that meet ICRC's standards and expectations. Expected information flow between delegations, métier, central evaluation function, and the wider institution are articulated.	Use of existing IT systems together with database and data warehouse tools for improved evaluation practice, including defined data dependencies and information flow between different areas of ICRC.	**Recommendation 5:** IT systems and tools

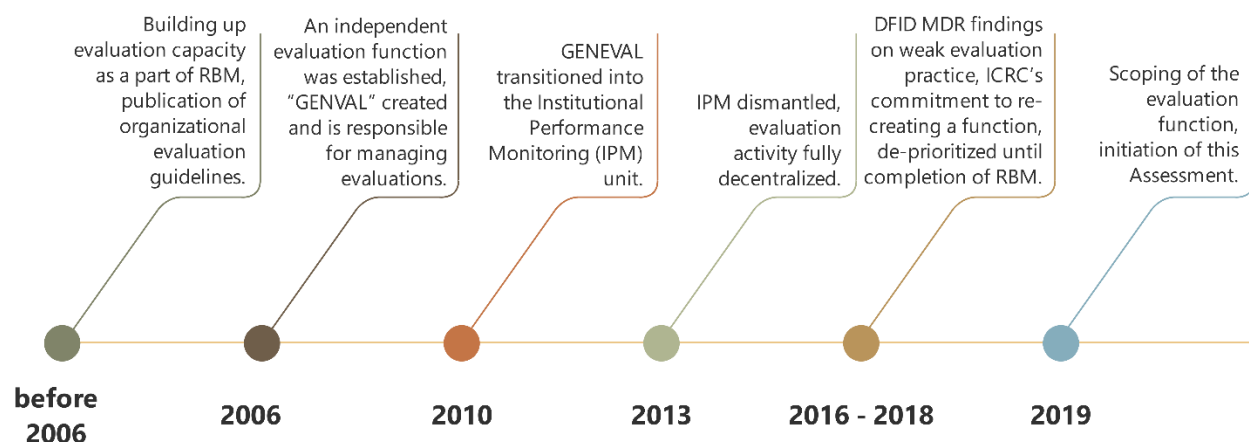| Finding 13: Incentive or Compliance Systems to Promote Evaluation			
Incentive or compliance opportunities for evaluation practice are not incorporated into evaluation planning, management, or use.	ICRC embeds evaluation more firmly in its planning and result-based management systems and integrates evaluation into other existing organizational systems that promote learning and/or accountability. Evaluation system addresses these areas of 'use' specifically: development of a management response, action on management response recommendations, and integration of evaluation feedback into ICRC ways of working.	Clarified opportunities to embed evaluation into existing RBM systems or leverage existing incentive or compliance mechanisms.	**Recommendation 6:** Evaluation Strategy
Weak incentive and compliance systems for IT procedures will impact ability to leverage IT for improved evaluation practice.	ICRC IT systems are leveraged to improve the production and use of evaluations that meet ICRC's standards and expectations. Ensure a strong change management process for IT solutions.	Clarified requirements for IT procedure compliance to ensure IT systems can adequately support evaluation management and use.	**Recommendation 4:** Information requirements of evaluation function **Recommendation 5:** IT systems and tools

III. Organizational Culture

This section explores ways organizational culture drives decisions and actions regarding evaluations and a revived evaluation function. It reviews challenges and opportunities in key areas of organizational culture as it influences ICRC's *interest in and capacity to produce and utilize evaluations* (Section II) as required *to meet organizational evaluation performance goals and strategic objectives* (Section I).

A. Existing Practice (As Is)

FINDING 14: Evaluation function history and current motivation for re-establishment. ICRC established an independent evaluation function in 2006 with the creation of 'GENEVAL'. Prior to this, the organization invested in building evaluation capacity as part of a push towards improved results-based management, including the publication of organizational evaluation guidelines in 2002. GENEVAL was responsible for managing 'independent evaluations.' Its role did not extend into other evaluative work, most notably 'reviews'. By 2010, GENEVAL had transitioned into the Institutional Performance Monitoring (IPM) unit. The Assessment did not locate documentation about the reason for this transition or the practical difference in scope between the two units. In 2010, IPM produced a report on the practice of independent evaluation from 2006-2009[21]. Following the report, a revised strategy was drafted in an attempt to address issues related to a significant drop in the number of proposed evaluations coming through the annual planning process, but it was never finalized. By 2013, IPM was dismantled and evaluation activity became fully decentralized. The Assessment did not locate a documented explanation for this decision. By 2016, the issue revived and ICRC committed to re-creating a function as part of its response to the 2016 DFID MDR findings on weak evaluation independence, coverage and frequency. Additional motivations for reviving the function included increasing interest in strategic- and impact-level evaluations and a desire for a more coherent evidence base for organizational learning and decision making. Action on this commitment was de-prioritized until the completion of the Planning and Monitoring for Results Reform initiative, understood to provide the foundation for a next phase of investment in evaluation. Scoping the function launched in 2019, with the initiation of this Assessment. *Key issues include*:

before 2006	2006	2010	2013	2016 - 2018	2019
Building up evaluation capacity as a part of RBM, publication of organizational evaluation guidelines.	An independent evaluation function was established, "GENVAL" created and is responsible for managing evaluations.	GENEVAL transitioned into the Institutional Performance Monitoring (IPM) unit.	IPM dismantled, evaluation activity fully decentralized.	DFID MDR findings on weak evaluation practice, ICRC's commitment to re-creating a function, de-prioritized until completion of RBM.	Scoping of the evaluation function, initiation of this Assessment.

---

[21] 'Overview of Evaluations 2006-2009 as per evaluation strategy', IPM Note to the Directorate on Independent evaluations, DIR1478 Annexe 2010

14.1 Findings in the 2010 report provides relevant considerations for the establishment of a function today. During the Assessment, and as documented in the OCP Project consultation[22], staff questioned the decision to terminate GENEVAL/IPM and note the need to consider lessons from previous functions. The 2010 report on evaluation practice highlights four trends that remain relevant to the re-establishment of an evaluation:

- **Low frequency and coverage of evaluations compared to ICRC commitments.** Between 2006 and 2009 there was a significant drop in proposed evaluations coming through the annual planning process. Even from 2006, evaluation coverage never reflected the ambitions of the strategy – in particular, a lack of field-level transversal evaluations (context, target populations) and an absence of evaluations for support functions or services. Frequency was also below target, with ICRC producing around 2-3 evaluations per year as compared to its institutional commitment of 6-8 evaluations a year.
- **Under-resourcing of the function.** The evaluation function was under-staffed to provide services in the manner described in the Strategy (e.g. function managed all independent evaluations). As a result, conducting an independent evaluation was perceived as slow. Staff opted for 'reviews' instead.
- **Lack of commitment.** Planned evaluations would be cancelled due to a 'lack of commitment of the budget holder'. The report highlighted the inability of the evaluation function to change that decision.
- **Increasing rate of reviews.** Departments increasingly 'sidestepped independent evaluations' through the use of 'reviews' that fell outside the scope of the evaluation function, thus 'avoiding due process and quality control'. The reasons for the shift towards reviews were listed as (1) a desire to improve the speed of the process and (2) intention to keep the findings of the review strictly internal, as by avoiding the label 'evaluation' the exercise was not under policy to share its findings externally (or internally).

14.2 Lack of independent authority over evaluation planning or decision making. Despite the role of GENEVAL and IPM in managing evaluations, the teams did not have authority to enforce evaluation coverage or frequency commitments. That authority sat within the departments, divisions, or the Directorate holding responsibility over the programs/projects to be evaluated. Likewise, the function struggled to enforce the practice of management responses to evaluations. Staff commented that management was not always receptive to critical, external perspectives on ICRC's work or willing to share this information with donors as committed in the Strategy. When enthusiasm for evaluation waned and practice shifted towards 'reviews', the evaluation function was unable to ensure continued evaluation frequency or coverage.

14.3 Past experiences remain relevant to the motivation for establishing a function today. From the information gathered by the Assessment, the current motivation for re-establishing an evaluation function can be summarized as: (1) desire to secure support and a driving team for transversal, strategically oriented evaluations; (2) recognition that the contexts where ICRC works are increasingly protracted and the organization needs to be able to assess its impact, successes, and failures in the context of multi-year programming; (3) the need for a more coherent and structured evidence base for organizational learning and decision making; (4) the need for a more coherent and structured process of ensuring organizational accountability to itself (internal accountability), to affected populations, and to donors, partners, and the wider public. For each of these points, ICRC's experiences with previous functions remain relevant. In particular, ensuring that the function is adequately resourced for the scope of the organization's evaluation ambitions, as well as ensuring the function is equipped with sufficient authority to drive commitments on what ICRC evaluates, where, and how.

---

[22] See 'The 'OCP – outcomes of workshop – Jan / Feb 2019'

FINDING 15: **Attitudes about evaluation.** Exposure to evaluation is limited in ICRC, found primarily in pockets of the organization with an established history of evaluation practice that often reflects wider sectorial trends (e.g. economic security has long tradition of evaluation). In this context, attitudes about evaluation can be summarized as 'constructive and positive', 'negative', and 'divergent' across topics of evaluation purpose, practice, and use. These attitudes are important drivers for both interest in re-establishing an evaluation function, as well as drivers of evaluation and 'non evaluation' in current practice. *Key issues include:*

15.1 *Positive attitudes regarding improved evaluation practice, a commitment to using evaluation for organizational learning, and a desire for ICRC to ensure evaluation 'integrity.'* Across discussions with staff during the Assessment, there is consensus that ICRC needs to improve its evaluation practice as part of its commitment to being a 'learning organization' (as stated in the 2019-2022 strategy). The precise path towards this improvement, however, is less consistent in staff attitudes about evaluation (see points 15.2 and 15.3). There is a concern regarding the 'integrity of process' in evaluation. Integrity is expressed as (1) motivation for evaluation is about improving efficiency, relevance, and effectiveness (versus politicized aims), (2) collaboration and internal transparency during evaluation management, and (3) the independence of the evaluator from the direct line management of the action being evaluated. If the integrity of the evaluation is seen as valid, evaluations themselves are more widely accepted. Staff also agree that evaluations help obtain an 'outside' view of what the action is doing and what it will achieve.

15.2 *Negative attitudes regarding the time requirements for evaluation and fear of accountability.* Across staff engagement during the Assessment, there are consistent messages that evaluations are perceived as time and budget-heavy exercises and that staff avoid evaluation out of fear of accountability. Negative attitudes about the time commitment for evaluation are linked to 'initiative fatigue', where staff have less space to do their core work as other demands on their time increase in response to new initiatives or reforms. Concerns about accountability include the possible use of evaluations for punitive action internally (e.g. loss of positioning or resources). These concerns are noted as drivers for the use of 'reviews.'

15.3 *Divergent attitudes regarding definitions or requirements for achieving 'independence', the benefit of external consultant expertise, and external publication or sharing of evaluations.* Across staff consultations, there was divergence in attitudes regarding requirements for achieving 'independence' in evaluation and the benefits of using external consultants as evaluators. Understandings of 'independence' in evaluation varied in regards to how it is achieved (not if it is important, a point on which there was wide agreement). There was no consensus in attitudes regarding whether evaluations could be conducted by staff internal to ICRC (but external to the action) and still be considered an 'independent evaluation'. This attitude correlated with perspectives on the purpose of evaluation, namely learning or accountability. Perspectives on independence linked with attitudes about external consultants. Although staff consistently raised concerns that external consultants may be unable to contextualize findings and recommendations to ICRC's operational structure, attitudes diverged about whether this actually manifests as a problem. Additionally, preferences diverged between using external consultants to provide expertise on evaluation that is missing within ICRC or whether to invest in building expertise in staff. Staff also expressed both practical questions or fears about the external transparency requirements of evaluations. Questions centered on whether evaluations should be shared with the full public, only with donors, or kept fully internal, and if shared, the extent to which information is made available (e.g. the full report or only parts of the report). There is also uncertainty on how to best judge these decisions at a policy level or on a case-by-case basis for each evaluation report. While a range of opinions were expressed on these topics, there is a consensus that it is an area on which ICRC is sensitive. Concern persist about the implications of the AIP, which includes the publication of Executive Summaries of independent evaluations and a number of qualifying situations where information may remain confidential.

FINDING 16: **Adaptability.** Through the Planning and Monitoring for Results Reform initiative and current discussions on reviving the evaluation function, ICRC demonstrates that it can adapt and is willing to adapt on the topic of monitoring and evaluation. It is also experienced with investing in the organizational change process. Staff noted, however, that 'initiative fatigue' (see Finding 15.2) and the significant growth experienced in the organization in recent years may present challenges to building an organizational-wide system for evaluation. In particular, staff noted anticipated challenges in building a shared vision and securing sufficient collective interest to fund the system in the context of a large, segmented organization that lacks a consistent understanding of what evaluations are and how they should be managed. *Key issues include:*

16.1 *Anticipate consistent interest in phasing the approach and building on existing systems.* During interviews, focus group discussions, and the workshop, there was consistent messaging across ICRC on the need to approach the development of a function in 'digestible phases.' Staff also consistently noted the importance of building on existing systems and initiatives to maintain coherence and reduce resistance to change. This included exploring ways to integrate evaluation planning into the PfR, and ensuring connectivity with the shift towards outcome monitoring when defining evaluation methods, requirements, or resource levels.

16.2 *Anticipate divergent interests in the scope and purpose of the function.* While there is a wide consensus in supporting the (re)creation of an evaluation function, there is disagreement on the scope of authority the function should hold. Units investing in evaluation capacity want to maintain or deepen their investment, while units eager to increase their evaluation work, but lacking resources, view the evaluation function as a source of support. Perspectives also vary on where the function and related staff should sit.

16.3 *Consider budgeting culture when determining how evaluation budgets are formed.* Evaluation budgets are formed within a wider organizational budgeting culture. While ICRC has a large degree of *de facto* flexibility on how to allocate core funds, there is resistance to 'cannibalizing' existing budget envelops to fund centralized evaluations or the evaluation function. ICRC does not commonly ring-fence budgets, direct budget amounts, or centrally direct budget priorities. The organization places a high value on maintaining independence from donors in the way it manages funds, which may affect decisions on where to source the funding (e.g. core budget vs. earmarked funds). There is also a freeze on HQ budget growth, which drives the way people imagine the evaluation function (staffing, cost of centralized evaluations) can be funded.

FINDING 17: **Information use and evidence-based decision making.** Information production at ICRC is widespread, including information framed in narrative, policy-type documents and information structured around quantitative data. Documentation of decisions and process are more systematic than is practiced in most humanitarian organizations. Despite this significant experience in information production, key documents suggest and staff report that ICRC struggles to reconcile data for wider trends analysis and decision making. *Key issues include:*

17.1 *Differential levels of evidence-based decision making.* According to interviews and as observed in the key documents (or by their absence), most of the information produced by ICRC is narrow in its intended audience or user base (e.g. it is specifically collected for one metier). Interviewees reported a mixed experience with information use by the intended user. Success factors include timing (the information was gathered at the right time for a decision or action to be taken), intentionality (information was intentionally collected to inform a specific decision), and stakeholder involvement in the process to ensure the relevance of and sustained interest in the information. At the delegation, departmental, and institutional levels, staff expressed difficulty in reconciling existing data into an analytical framework that would support decision-making at these levels.

17.2 Explore information harmonization between learning and accountability functions. In a focus group discussion with a cross-section of staff from accountability and learning functions, the issue of 'information harmonization' and 'shared accountability frameworks' were leading concerns. This includes harmonizing the way information is produced (establishing common 'ways of working') and used between the functions, as well as syncing protocol for tracking action on recommendations (see KII FGD Summary Report Annex III).

B. <u>Desired Status (To Be)</u>

This section identifies objectives for organizational culture as indicated by the 2019-2022 Strategy and as expressed by ICRC staff during the Assessment and within key documents.

Application of ICRC 2019-2022 Strategy. *Organizational culture is leveraged to adapt evaluation practice to meet ICRC's strategic objectives and commitments to learning and accountability.*

ICRC's mission statement will continue to provide the framework for ICRC's development, including especially the Fundamental Principles of impartiality, neutrality, and independence and engagement with stakeholders about ICRC's actions and their impact[23]. This includes ensuring that evaluation supports bringing together humanitarian activities and institutional initiatives independent of their organizational labels or structures[24].

Areas of organizational culture that hinder evaluation are identified and mitigated in the approach used to re-establish the evaluation function and promote evaluation practice.

Objectives expressed by ICRC staff and other key documents. Organizational culture objectives articulated by staff during the Assessment or as otherwise documented by ICRC[25]. Information source is in brackets:

Evaluation function history and current motivation for re-establishment. Lessons learned from the GENEVAL experience, including what worked in the process followed by GENEVAL, inform the development of a new evaluation function. (Document Review, Workshop).

Attitudes about evaluation. Attitudes shift about evaluation in the 'pockets of resistance' on the reasons *why* evaluations are done (e.g. staff feel threatened) and *how* evaluations are done. 'Bias' towards reviews is addressed. Demonstrate that evaluations are possible and that this is not new to ICRC or the humanitarian sector. Center ICRC principles in the discussion of 'independence' of evaluation. (Workshop)

Adaptability. ICRC staff appreciate the purpose of the evaluation function and willingly engage in the process to support it. Sustained, consistent support of leadership for evaluations during and after the change process. Evaluation framework provides direction and consistency of vision (Workshop, Interviews)

Information use and evidence-based decision making. Evaluation contributes to institutional data that is better harmonized to shape policy (internally) and ensure accountability to affected populations and donors. Institutional data is better harmonized for improved sharing of lessons learned within ICRC and with other organizations.

---

[23] ICRC 2019-2022 Strategy, Page 6.
[24] Ibid, Page 10
[25] See 'The 'OCP – outcomes of workshop – Jan / Feb 2019' and the Analysis and Evidence Strategy, 2019-2022, p7. Similar intention described in the 'Overview of the CORE's Working Model' under its 'CORE-coherence' initiative (p1).

C. Gaps Analysis

This section presents a summary table (Table 6) of key issues identified in the As Is findings of the Assessment on organizational culture matched against their most relevant To Be objectives (Section B). It highlights the primary gaps to be addressed and where this links to the recommendations.

The Gaps Analysis focuses on priority recommendations for establishing an evaluation function. Additional actions that may be required in a second or third phase of a phased approach to establishing an evaluation function are addressed within the proposed Evaluation Strategy, under a proposed timeline for actions in Years 1-2, 3-4, and 5-beyond.

Table 6: Organizational Culture: *Gaps Analysis*

Existing Practice (As Is)	Desired Status (To Be)	Primary Gaps: *What is missing*	Linked Recommendation
Finding 14: Evaluation function history and current motivation for re-establishment			
Findings in the 2010 report on evaluation provide relevant considerations for the establishment of a function today.	ICRC's mission statement provides the framework for development, including the Fundamental Principles and engagement with stakeholders about ICRC's actions and impact.	Defined independence and authority of the evaluation function over coverage and planning.	**Recommendation 2:** Purpose and structure of the evaluation function
Lack of authority over evaluation decision making.	Lessons learned from the GENEVAL experience inform the development of a new function. Sustained, consistent support from leadership for evaluations.	Defined and secured level of staffing and funding necessary for the evaluation function to fulfil its purpose.	**Recommendation 3:** Resourcing the evaluation function
Past experiences remain relevant to the motivation for establishing a function today.	Areas of organizational culture that hinder evaluation are identified and mitigated in the approach used to re-establish the evaluation function and promote evaluation practice.	Defined evaluation framework, inclusive of reviews and appropriately levels for centralized or decentralized management.	**Recommendation 6**: Evaluation Strategy
Finding 15: Attitudes about evaluation			
Positive attitudes regarding improved evaluation practice, a commitment to using evaluation for organizational learning, and a desire for ICRC to ensure evaluation 'integrity.'	Organizational culture is leveraged to meet ICRC's commitments to learning and accountability.	Defined scope of the evaluation function, including role for improving evaluation coverage and quality	**Recommendation 2:** Purpose and structure of the evaluation function
	ICRC's mission statement provides the framework for ICRC's development, including the Fundamental Principles and stakeholder engagement about ICRC's actions and impact. Demonstrate that evaluations are possible and not new.	Defined approach to 'independence' within evaluation framework, including where evaluation is managed and how evaluation teams can be composed per type of evaluation	**Recommendation 6**: Evaluation Strategy
		Available technical expertise to guide identification of evaluation in challenging evaluation contexts	**Recommendation 3:** Resourcing the evaluation function

Negative attitudes regarding the time requirements for evaluation and fear of accountability.	Areas of organizational culture that hinder evaluation are identified and mitigated when re-establishing the evaluation function and promoting evaluation practice. Attitudes shift about evaluation in the 'pockets of resistance' on the reasons *why* evaluations are done (e.g. staff feel threatened) and *how* evaluations are done. 'Bias' towards reviews is addressed. Demonstrate that evaluations are possible and not new to ICRC or the humanitarian sector.	Defined independence and authority of the evaluation function, including over evaluation planning and role in supporting evaluation use for accountability. Sufficient and effectively coordinated human resources to support evaluation practice, including available technical expertise to guide evaluation in challenging evaluation contexts. Defined evaluation framework, including 'light' vs. more robust evaluation types.	**Recommendation 2:** Purpose and structure of the evaluation function **Recommendation 3:** Resourcing the evaluation function **Recommendation 6:** Evaluation Strategy
Divergent attitudes regarding definitions or requirements for achieving 'independence', benefit of external consultant expertise, and transparency requirements for evaluations.	ICRC's mission statement provides the framework for ICRC's development, including the Fundamental Principles and engagement with stakeholders about ICRC's actions and their impact.	Defined independence and authority of the evaluation function over coverage and planning. Defined approach to evaluation 'independence' within evaluation framework, including where evaluation is managed and how evaluation teams can be composed per type of evaluation. Defined approach for evaluation publications, as per the parameters of the AIP.	**Recommendation 2:** Purpose and structure of the evaluation function **Recommendation 6:** Evaluation Strategy
Finding 16: Adaptability			
Anticipate consistent interest in phasing the approach and building on existing systems	ICRC staff appreciate the purpose of the evaluation function and willingly engage in the process to support it. Evaluation framework provides direction and consistency of vision.	Clarified phasing for the development of the evaluation function and implementation of the evaluation strategy.	**Recommendation 6:** Evaluation Strategy
Anticipate divergent interests in the scope and purpose of the function.	Evaluation supports bringing together institutional initiatives independent of their organizational labels or structures. ICRC staff appreciate the purpose of the evaluation function and willingly engage in the process to support it. Sustained, consistent support of leadership for evaluations during and after the change process.	Defined independence and level of authority of the evaluation function over evaluation practice at centralized vs. decentralized levels. Clarified space for demand-led evaluation practice and responsibility at decentralized levels.	**Recommendation 2:** Purpose and structure of the evaluation function **Recommendation 6:** Evaluation Strategy
Consider budgeting culture when determining how evaluation budgets are formed.		Clarified options for financing evaluations, linking evaluation practice with ICRC's strategic objectives.	**Recommendation 3:** Resourcing the evaluation function

Finding 17: Information use and evidenced based decision making			
Differential levels of evidence-based decision making.	Evaluation supports bringing together institutional initiatives independent of their organizational labels or structures. Institutional data is better harmonized to shape policy, to ensure accountability, and for improved sharing of lessons learned within ICRC and with other organizations.	Available technical expertise to promote the use of evaluation at all levels of ICRC. Clarified opportunities to anchor evaluation information in ICRC's calendar and connect different information products across the organization. Clarified opportunities for using IT systems to support evaluation use.	**Recommendation 3:** Resourcing the evaluation function **Recommendation 6:** Evaluation Strategy **Recommendation 5:** IT systems and tools
Explore information harmonization between learning and accountability functions.			

⤧⤦ IV. External Influencing Environment

This section reviews opportunities and threats in the external environment as they influence ICRC's *culture around evaluation* (Section III) and *interest in and capacity to produce and utilize evaluations* (Section II) as required *to meet organizational evaluation performance goals and strategic objectives* (Section I).

A. Existing Practice (As Is)

FINDING 18: **Administrative and legal environment.** The Assessment did not find administrative or legal influences on evaluation from the external environment. In this regard, ICRC operates with a significant degree of independence over its evaluation agenda.

FINDING 19: **Political and economic environment**. The funding environment and politics of donor institutions do influence ICRC's decisions about evaluation and perspectives on the evaluation function. *Key issues include:*

19.1 Significant flexibility in the relationship between evaluations and donor funding. Related to Finding 11.1, the majority (~70%) of ICRC's funding derives from donor contributions to a core budget. Around 30% of funding is earmarked at a delegation level. There are no project-level donor evaluation requirements. The only example of evaluations that were conducted as part of a contracted donor commitment are evaluations that ICRC itself proposed for the performance framework signed with DFID in 2017[26]. It is worth noting, however, that DFID expected the inclusion of evaluations into the performance framework, even if it did not dictate the subjects that ICRC chose to evaluate.

19.2 Increasing donor and partner expectations for evaluation. Across ICRC, staff report pressure to provide a more evidenced account of program-level and institutional impact to remain competitive for funding in a landscape of shrinking humanitarian budgets[27]. Likewise, as ICRC moves to secure more development-oriented funding for operations in protracted crises, it is expected by development donors and partners to provide an account of its longer-term impact. Differences between expectations were noted as follows:

- **DFID**. Criticisms of ICRC's evaluation performance in the 2016 MDA led to two key developments: (1) ICRC initiated a series of three transversal evaluations that were supported by PME, reportedly the first time it had initiated and supported institutional evaluations since the closure of the IPM unit in 2013, and (2) ICRC committed to sharing an evaluation framework with DFID by the end of 2017. ICRC anticipates that DFID will revisit this in the next MDA. Main points of critique from DFID included limited strategic evaluations, low evaluation frequency, and the absence of an evaluation function or evaluation framework. DFID also requested that ICRC develop its approach to Value of Money.

- **World Bank**. Requests to conduct 'third party monitoring' of ICRC operations that are funded through the Bank sparked discussions in ICRC about the potential for evaluation as a 'substitute' to the Bank's monitoring approach. The evaluation function would need to be sufficiently independent in location and directive over the evaluation agenda that the Bank would accept that there is no operational influence on the reports. If this were accepted, then ICRC would consider including evaluation as a standard clause in any project with the Bank and thus eliminate the pressure for third party monitoring. The discussion about evaluation with the Bank is primarily oriented towards transparency and compliance issues.

---

[26] Including the 'Evaluation of Diversity, Inclusion, and AAP in ICRC Operations' (2018), Accountability to Affected Populations, 'Sexual Violence Field Evaluation' (2017), and the 'ICRC Health Care in Danger Evaluation Report' (2017)

[27] See the Analysis and Evidence Strategy, 2019-2022, page 7.

- **BPRM (US government) and the German MFA.** The GMFA and BPRM were interviewed as external stakeholders for the Assessment. Between them, there is a consistent message that donors want more information from ICRC on what it is doing, what it is achieving, and what it is learning. The GMFA, for example, stated that it wants ICRC to put evaluation on the agenda of the annual meeting or to find a slot for evaluation learning in a policy forum. BPRM expected ICRC to share evaluation reports. Both donors stressed using evaluation for learning, versus compliance assurance - a stark contrast to the reported discussions with the Bank. Like DFID, the two funders expressed an interest in evaluations on questions of 'strategic' importance, including the link between development and humanitarian aid (GMFA) and humanitarian diplomacy (BPRM). They were less interested in delegation-level metier-focused evaluations – again, a notable difference from discussions with the Bank, which is focused on project-level evaluation as a replacement for third party monitoring. According to donors, evaluations are an important source of information in their internal reporting systems and mechanisms for positioning funded partners for continued support. BPRM and the GMFA are enthusiastic about the development of an evaluation function at ICRC, but stressed that they support ICRC achieving this in a way that 'makes the most sense' for the organization. Unlike the Bank or DFID, they did not hold strong opinions on the function's independence or how it should operate.

19.3 Maintaining institutional independence and control over the evaluation agenda. In response to the increasing number and varied perspectives of donor and partner expectations on ICRC's approach to evaluation, the organization wants to 'get out in front' of the issue. Namely, ICRC crafts its own approach to evaluation and leads with this in its discussions with donors and partners (vs reactive commitments to external expectations).

FINDING 20: **Social or outside normative environment**. The Assessment did not find many concerns about the influence of its wider social or normative environment on the organization's evaluation practice or agenda. During interviews the following considerations were noted:
- **Alignment within the Movement on definitions and practice for evaluation.** Staff and interviewed Movement actors noted that ICRC should consider areas of alignment with other Movement actors in an effort to promote coherence in language/definitions and expectations of National Society (NS) partners. ICRC is already participating in a Movement MEAL working group, including representation with IFRC and several NSs. As expressed by ICRC staff, there is a desire to better understand collective 'Movement' impact and explore opportunities for joint evaluation with NSs in relevant contexts. A question was also raised about how to coordinate with IFRC on evaluation plans with NSs.
- **Comparative achievement or structure with peer agencies.** Although ICRC staff express a desire to see an evaluation function tailored to the specifics of the organization, there is also a concern about the degree to which ICRC aligns with industry standards and the best practice followed by other agencies. This request helped inform the approach for the Health Check exercise as a type of benchmarking and the inclusion of MSF and UNHCR standards for evaluation.
- **ALNAP Steering Committee representative.** ICRC is an active member and steering committee representative at ALNAP, a leading authority for the practice of evaluation in humanitarian action. As a member of the committee, ICRC is able to influence discussions about humanitarian evaluation priorities and contribute to the work of ALNAP in providing guidance on the topic.

FINDING 21: **Technological environment**. The Assessment did not find technological influences on evaluation from the external environment. If ICRC explores *new* IT systems to leverage evaluations going forward, it is recommended that it look specifically into systems compatibility, path dependency risks, and the ownership model of the new system (e.g. licencing, subscription, open source, bespoke). At this time, the Assessment focused on the use of existing systems (SharePoint and Tableau) and didn't note any concerns with these topics.

B. Desired Status (To Be)

This section identifies objectives for external influence as indicated by the 2019-2022 Strategy and as expressed by ICRC staff during the Assessment and within key documents.

Application of ICRC 2019-2022 Strategy. *The opportunities and threats from the external environment on current and future evaluation performance, capacity, and culture are understood and managed in a way that is consistent with ICRC's mission statement, including the Fundamental Principles of impartiality, neutrality, and independence and engagement with stakeholders about ICRC's actions and their impact[28]. Strengthening ICRC's work on evaluations should support donor diversification, while serving to provide an evidence base for demonstrating the impact of more sustainable models of humanitarian financing.[29]*

Aspects of the external environment that pose a threat to evaluation are identified and mitigated in the approach to re-establishing the evaluation function and promoting evaluation practice.

Objectives expressed by ICRC staff and other key documents. External environment objectives articulated by staff during the Assessment or as otherwise documented by ICRC. Information source is in brackets:

Administrative and legal environment. *Not applicable: no expression of objectives from staff on the topic.*

Political and economic environment. ICRC responds to the demand for better evidence on its programs as part of its commitment to being a learning organization. ICRC commits to updating its evaluation strategy, including the evaluation framework, and establishing an independent evaluation function. ICRC publishes the Executive Summaries of Independent Evaluations on its website. ICRC maintains its independence and proactively defines and maintains its own evaluation agenda[30]. (FGD, Interviews, Document Review)

Social or outside normative environment. ICRC aligns with accepted standards for evaluation practice and quality, including the OECD-DAC criteria and ALNAP guidelines. ICRC identifies points of coordination or alignment with other Movement actors regarding terminology or definitions used in key guidelines or evaluation documents, as well as in process for planning evaluations with NSs (FGD, Interviews)

Technological environment. *Not applicable: no expression of objectives from staff on the topic.*

C. Gaps Analysis

This section presents a summary table (Table 7) of key issues identified in the As Is findings of the Assessment on external environment matched against their most relevant To Be objectives (Section B). It highlights the primary gaps to be addressed and where this links to the recommendations.

The Gaps Analysis focuses on priority recommendations for establishing an evaluation function. Additional actions that may be required in a second or third phase of a phased approach to establishing an evaluation function are addressed within the proposed Evaluation Strategy, under a proposed timeline for actions in Years 1-2, 3-4, and 5-beyond.

---

[28] ICRC 2019-2022 Strategy, Page 6.
[29] Ibid, Objective 3.6
[30] See the 'Evaluation Function DIRGEN Action Plan' 2017 and the 'Access to Information Policy' (2019)

Existing Practice (As Is)	Desired Status (To Be)	Primary Gaps: *What is missing*	Linked Recommendation
Finding 19: Political and economic environment			
Significant flexibility in evaluations and donor funding	ICRC responds to the demand for better evidence on its programs as part of its commitment to being a learning organization and its objectives for donor diversification, while maintaining independence and proactively defining its own evaluation agenda.	Defined evaluation strategy and evaluation plan that reflects ICRC's objectives for organizational learning and improvement. Approved budgeting process that reflects opportunities within ICRC funding structure, inclusive in the budget holding arrangements, budget source, and budget amount.	**Recommendation 6:** Evaluation Strategy **Recommendation 3:** Resourcing the evaluation function
Increasing donor expectations for evaluation at ICRC.	ICRC responds to the demand for better evidence on its programs as part of its commitment to being a learning organization and its objectives for donor diversification, while proactively defining its own evaluation agenda. ICRC publishes the Executive Summaries of Independent Evaluations as per the AIP. ICRC commits to establishing an independent evaluation function.	Defined priorities for evaluation practice and standards. Defined evaluation strategy and evaluation plan that states process and responsibilities for external report dissemination. Defined evaluation function, including scope and location.	**Recommendation 1:** Proposed application of industry standards **Recommendation 6:** Evaluation Strategy **Recommendation 2:** Purpose and structure of the evaluation function
Maintaining institutional independence and control over the evaluation agenda	ICRC responds to the demand for better evidence on its programs as part of its commitment to being a learning organization and its objectives for donor diversification, while maintaining independence and proactively defining its own evaluation agenda.	Defined evaluation strategy and evaluation plan that reflects ICRC's objectives for organizational learning and improvement.	**Recommendation 6:** Evaluation Strategy
Finding 20: Social or outside normative environment			
Alignment within the Movement on definitions and practice for evaluation.	ICRC identifies points of coordination or alignment with other Movement actors regarding terminology or definitions used in key guidelines or evaluation documents, as well as in process for planning evaluations with NSs	Clarified areas of alignment with Movement actors.	**Recommendation 6:** Evaluation Strategy
Comparative achievement or structure with peer agencies. ALNAP Steering Committee representative.	ICRC aligns with accepted standards for evaluation practice and quality, including the OECD-DAC criteria and ALNAP guidelines.	Defined priorities for evaluation practice and standards. Defined and actively applied evaluation strategy, including application prioritized standards.	**Recommendation 1:** Proposed application of industry standards **Recommendation 6:** Evaluation Strategy

3. CONCLUSION AND DISCUSSION

This section of the Assessment provides conclusions to the framing questions (see Table 2, Framing Questions) posed about the existing practice, desired end state, and gaps and action analysis for (1) organizational performance, (2) organizational capacity, (3) organizational culture, and (4) the external influencing environment. It starts with a summary of ICRC's current standing against its objectives for evaluation.

Summary Conclusions

ICRC's current approach to evaluation encourages flexible, self-directed learning for managers and teams across different levels of the organization. It also, however, serves to reinforce an overall weak and fragmented evaluation culture, with consequent impacts for organizational learning and accountability.

The absence of a defined and consistently applied approach to evaluation independence, purpose and scope, identification and coverage, management, and utilization undermines the credibility and utility of the evaluation and like exercises that ICRC produces. It also incentivizes 'non evaluation' among staff who are unfamiliar with or skeptical of humanitarian evaluation. The end result is an evaluation portfolio that neither reflects ICRC's strategic direction nor provides an evidence base commensurate to the size and function of the organization.

From an accountability perspective, ICRC struggles to deploy evaluations as part of its institutional strategy management or governance systems. The strength of evaluations to assist humanitarian organizations in understanding the unintended consequences and effectiveness of their actions and presence, and to structure a transparent response to any issues that are discovered, goes under-utilized with detrimental implications for ICRC's accountability to affected populations. These risks affect a large proportion of ICRC's key operations – including a significant number of field delegations, thematic files and global initiatives, and nearly entire métiers, most notably Protection, Movement, and Prevention. From a learning perspective, initiatives, thematics, approaches, policies, programs or projects that the organization wants to evaluate are difficult to secure (or enforce) under the existing system for evaluation planning and budgeting. Without adequate in-house technical expertise, units that want to engage in evaluation opt-out, believing their questions of interest are unanswerable or from confusion about the applicability of evaluation criteria or techniques to their program model or context.

Gaps that need to be addressed include (1) the purpose, scope, and location of the evaluation function and its relationship to evaluation leadership and authority, (2) the staffing and funding arrangements necessary for the function to fulfill its purpose, including number of staff and technical expertise, budget ownership, funding sources, and budget levels, (3) the information requirements of the function to perform its role of evaluation manager, quality controller, promoter, and its responsibility for the implementation of the evaluation strategy, (4) an identified set of ways existing IT systems can be leveraged to promote more systematic evaluation management, quality assurance, and utilization across several levels of the organization, and lastly (5) an evaluation strategy that sets forth a defined and relevant evaluation framework and establishes parameters for expected responsibilities, assurance mechanisms, and standards for evaluation content and workflow.

In addressing these gaps, ICRC should seek to maintain the learning-oriented and flexibility benefits of its current approach and leverage opportunities in its existing capacities, culture, and from its external environment. Pockets of evaluation culture – such as EcoSec - should be leaned on as an example and source of encouragement for areas of the organization newly embarking on evaluation.

3.1 Organizational Performance

Is ICRC currently effective in producing and utilizing evaluations that meet or exceed organizational expectations?

The Assessment found that, currently, ICRC **is not producing or utilizing evaluations in a way that meets organizational expectations** as implied in the 2019-2022 strategy or expressed by staff during the course of the Assessment and in key organizational documents.

Is ICRC effective in producing and utilizing evaluations as benchmarked against accepted industry standards?

The Assessment found that, currently, ICRC **is not effective** in producing or utilizing evaluations as benchmarked against industry standards. This was assessed against (1) donor expectations as expressed through the MOPAN criteria (inclusive of UNEG and OECD-DAC criteria standards), (2) evaluation in humanitarian action standards expressed through ALNAP and The Professional Standards for Protection Work standards, (3) evaluation standards of the Movement as expressed through the IFRC evaluation framework, the Fundamental Principles of the Red Cross and Red Crescent Movement, and the Code of Conduct for International Red Cross and Red Crescent Movement, and (4) evaluation standards of humanitarian organizations of a similar size and scope as expressed through the MSF evaluation guidelines and the UNHCR evaluation policy and evaluation strategy.

What standards should be in effect to promote improved operational impact, organizational learning and accountability, and reach the objectives of the 2019-2022 ICRC strategy? How are these prioritized for action?

The Assessment identifies a set of industry standards that ICRC should **prioritize for action on the basis of the standards' ability to target primary causes of weak or fragmented evaluation practice and build a foundation for future improvements**. They respond to key issues identified in current practice, including a lack of documented definitions, requirements, or expectations for evaluation performance and fragmented evaluation practice. These standards are proposed in under 'Recommendations', Section 4.

What organizational performance gaps exist between current evaluation practice and the desired end-state?

Benchmarked against all industry standards, ICRC is performing well below average at 45% alignment to standards (see Health Check Report Annex V).

When benchmarked only against standards identified as priorities for action (see column D in the Health Check and also listed under Recommendation 1, Section 4) in achieving the desired end-state as defined by the ICRC 2019-2022 strategy and expressed objectives of ICRC staff, current performance changes to 40.3%, reflecting a significant number of critical performance gaps between existing and targeted performance levels.

The primary area of weakness is 'consistent, independent evaluation of results', which scores a 25% when benchmarked against prioritized standards. This is linked to the absence of a guiding evaluation strategy and weak practice around evaluation identification, coverage assurance, and planning. Performance on the remaining standards ranges between the low 40s to 50% alignment.

3.2 Organizational Capacity

Does ICRC currently have sufficient capacity to produce and utilize evaluations?

The Assessment found that, currently, ICRC **does not have sufficient capacity** to produce or utilize evaluations in a way that meets organizational expectations as implied in the 2019-2022 strategy or expressed by staff during the course of the Assessment and in key organizational documents. Capacity is also insufficient to produce or utilize evaluations as benchmarked against a proposed set of prioritized industry standards.

What areas of organizational capacity need to be strengthened, changed, or introduced in order for ICRC to meet its evaluation performance goals and strategic objectives?

Areas of organizational capacity that need to be strengthened or introduced to meet ICRC's evaluation performance goals and strategic objectives include:
- Sustained evaluation leadership from ICRC managers and the evaluation function that proactively drives evaluation practice across the organization, including areas that have limited exposure or interest in evaluation. Responsibilities for evaluation initiation and participation require clarification.
- Improved efficiencies in use of staff time and resources when producing and using evaluations. Evaluation needs to be sufficiently staffed in number of people and according to required expertise.
- Increased predictability and assurance of evaluation funding that proactively secures funds at an amount adequate to conduct evaluations at a standard commensurate to the size and function of the organization, as well as in the manner implied in the organization's strategy.
- Application of existing IT platforms to leverage evaluation management, governance, and evaluation use at multiple levels of organizational learning, accountability, and decision making.
- Application of existing incentive and compliance systems to leverage how evaluations are used for decision making, learning, and accountability. This needs to include consideration for how ICRC can, or will, enforce adherence to the IT procedures.

Ways in which these areas can be addressed to best impact improvement in evaluation performance are highlighted under 'Recommendations', Section 4 and in the Annexes VI-IX.

3.3 Organizational Culture

How does organizational culture influence evaluation production and the utilization of evaluations?

Organizational culture influences the practice of evaluation production and utilization in following ways:
- How people understand the scope and structure of an evaluation function and the issues they want to see addressed in its re-creation.
- Attitudes about the purpose of evaluation and the best way to structure evaluation for differential purposes of accountability and learning, with consequences on evaluation resourcing (staffing, funds).
- Levels of enthusiasm or interest in evaluation vs. suspicion, worry, or skepticism, with consequences on evaluation coverage.
- Understandings of what is required to achieve 'independence' in evaluation, with consequences for the role of the evaluation function, structure of the evaluation framework, the way in which evaluators are staffed / recruited and managed.

- Adaptability to change, with consequences for how staff react or support a shift in the degree to which evaluation is formalized, centralized, and funded.
- Belief in the utility and potential of evaluations to respond to differential levels of evidenced-based decision making, with implications for how evaluations are anchored as an information sources in key moments of the organization calendar and the degree to which ICRC invests in information harmonization.

Are there areas of organizational culture that can be leveraged for improved evaluation practice? Are there areas of organizational culture that could be a hinderance or an obstacle when seeking to improve evaluation practice?

Areas of organizational culture that can be leveraged include interest in information harmonization across learning and accountability functions at ICRC. The challenges faced by the previous evaluation functions will remain obstacles if not addressed in the approach taken with the new function. Additionally, initiative fatigue and budgeting culture may present obstacles to how the function is resourced and launched. Attitudes about evaluation can serve as either opportunities or obstacles, depending on how decisions about the evaluation function and strategy are communicated. Likewise, divergent interests in the scope of the function can provide opportunities or act as obstacles depending on how decisions are communicated.

3.4 External Influencing Environment

How does the external environment influence the current production and utilization of evaluations?

The **political and economic environment** of donor and partner expectations influence ICRC decisions about what to evaluate, how or whether to share evaluations externally, and staff perspectives about the role of an evaluation function. This influence is exercised primarily at the institutional level, not through project-based funding. The World Bank is the exception and influences both institutional perspectives on the role and structure of an evaluation function and decisions to evaluate specific projects. Attitudes vary between DFID, the World Bank, the German MFA, and BPRM regarding desired influence over ICRC's evaluation agenda. The **social or outside normative environment** has limited influence. Concerns include alignment within the Movement on evaluation definitions and practice and an interest to understand ICRC's comparative performance against fellow humanitarian agencies. ICRC can influence the normative environment of evaluation through its participation on the ALNAP steering committee. No influences were found in **legal or administrative environment or technology**.

What opportunities or threats do these influences present to evaluation performance, capacity, and culture?

Areas of the external environment that serve as opportunities include flexibility in the relationship between donor funding and evaluations, the supportive and flexible positions of GMFA and BPRM (in particular the use of evaluation for learning), and the desire to maintain institutional control over the evaluation agenda. All aspects of the social and normative environment provide ways for ICRC to improve how it defines, manages, and uses evaluations. No direct threats were noted from the external environment. Increasing donor expectations may often feel like a threat, but it can also present opportunities for funding an evaluation agenda and create incentives to improve evaluation practice. Donor interests are not widely opposed to ICRC's own interests for evaluation. There are, however, differences in the Bank's perspective compared to ICRC's perspective about the purpose of evaluation. DFID is also less flexible in how it expects ICRC to address evaluation improvements, compared to the GMFA and BPRM.

4. RECOMMENDATIONS

Recommendations are framed in response to the Assessment objectives (see page 1, 'Objectives'), with additional sections included to address Framing Question 1.3 and address areas for improvement that do not fall within the specific scope of the Assessment objectives. A more detailed presentation of options surrounding recommendations for the purpose and structure of the evaluation function and its required resources are provided in Annexes VI and VII. A more detailed account of the proposed system for IT application in evaluations is provided in Annex VIII. The Evaluation Strategy is found in Annex IX, matching the recommendations of this Assessment.

RECOMMENDATION 1: Proposed application of industry standards.

ICRC should align with industry standards that promote clarity and consistency in the production and use of quality evaluations. Industry standards can be applied in phases, first prioritizing actions that serve as foundational layers for the improved performance of the entire system. They should respond to the key issues identified in current practice, including a lack of definitions, requirements, or expectations for evaluation performance and fragmented evaluation practice. The following areas are prioritized and reflected in the subsequent recommendations:

Governance. The structure and scope of the function aligns with Movement Principles of neutrality, impartiality, and independence. (Health check criteria 1.1, 1.8, 3.12)

Evaluation Purpose and Scope. Develop an evaluation strategy that defines the purpose and scope of evaluation as distinct from other learning and accountability-oriented functions and that presents a framework for different categories of evaluation. The strategy draws on international best practice, including the positioning of evaluation topics (e.g. programme, strategic, thematic, etc.) (Health check criteria 2.1, 2.3)

Evaluation Identification, Selection, and Coverage. Develop a prioritized evaluation plan that presents systematic coverage of the ICRC. (Health check criteria 2.4, 2.5). Selection criteria and use of OECD/DAC criteria for evaluation questions are addressed in the evaluation strategy. The strategy states coverage requirements and assurance mechanisms (Health check criteria 2.2, 3.6).

Evaluation Management. Ensure a dedicated and independent budget line for centralized and decentralized evaluations (Health check criteria 1.4). The strategy states quality assurance requirements and mechanisms for evaluations. (Health check criteria 2.2)

Evaluation Content: Evaluations of ongoing or completed programs ensure accountability for actions taken to address protection concerns. Evaluations of protection programming focus on issues of contribution (vs. attribution) to a desired result(s)[31]. (Health check criteria 3.9, 3.10)

Evaluation Utilization. Ensure evaluation and like activities include a documented management response with clear responsibilities and a timeline for implementation of recommendations (Health check criteria 4.1, 4.2, 4.3). Develop a complete evaluation report library and clarify expectations for internal and external report dissemination in support of accountability and learning. (Health check criteria 5.1, 5.2, 5.3)

---

[31] Attribution-focused evaluations determine the extent to which an action directly caused an observed set of results. The approach is primarily quantitative and statistically based. Contribution-focused evaluations determine the extent to which an action contributed to or influenced results in relationship to other factors such as the environment, the work of other actors, etc. This approach is largely qualitative. While the methodologies differ, both approaches aim to answer (1) if change happened on a result of interest and (2) how the action is linked to that change.

Evaluation function purpose and scope: *What would an evaluation function do?*

The Assessment reviewed three models for the evaluation function purpose and scope: Command / Control, Center of Excellence, and a Hybrid of the two. Based on the Assessment findings, **ICRC should adopt a 'hybrid' approach to the evaluation function and structure**, wherein the function combines elements of the service-oriented 'centre of excellence' model with more accountability-oriented 'command and control' models.

In addition to providing a technical support and expertise role for evaluation across the organization, there are key areas on which the evaluation function would also have direct authority and a control function. This includes:

- Oversight on evaluation strategy and all guidelines produced by ICRC. Function has control over evaluation strategy and organization-level guidelines. Guidelines may be adapted by the métiers.
- Function defines the process for evaluation planning at central and decentralized levels. Function steers development and approval of centralized evaluation plans. The function provides technical support, as required, during decentralized evaluation planning, but does not directly lead the process. Function steers approval of a final evaluation plan, inclusive of centralized and decentralized evaluations.
- In its authority over the evaluation plan, the function reviews and approves ICRC's use of different evaluation types according to their appropriateness for the scope of the exercise, this includes decentralized evaluation plans for reviews and any proposed evaluative 'lessons learned' exercises.
- Function either directly manages <u>or</u> sits on the steering committee for all centralized evaluations. It may provide 'help-desk' guidance on decentralized evaluations, as requested. Function has full quality 'control' over centralized evaluations and provides quality 'assurance' for decentralized evaluations.
- Function has capacity to join evaluation team or develop methodology for central or (if requested). decentralized evaluation. There is, however, limited involvement of the function in directly conducting evaluations. Function staff ensure adherence to quality standards and evaluation utility when working with consultants, academic partners, and potential staff rosters who conduct evaluations.
- This scope implies a split decentralized / centralized evaluation framework.

Additional details provided in Annex VI, 'Evaluation Function Purpose and Location Recommendations Report'.

Evaluation function placement: *Where is the evaluation function located?*

The Assessment reviewed three models for the location of the evaluation function: Fully within the Governing Body / Assembly, Fully within the Executive / Director General's Office, and a Split Responsibility vs. Location approach. Based on the Assessment findings, **ICRC should consider adopting a 'split responsibility vs. location' approach to the placement of the evaluation function**, wherein the function is located in the Director General's (DG) office and responsibilities for aspects of the evaluation function are split between the Assembly and the DG to maximize the learning-oriented benefits of an 'executive office placement' model with the more accountability-oriented 'governing body placement' model supported by donor evaluation standards (reflected in MOPAN criteria) and UNEG.

It would have the following characteristics:
- The function is located in the DG's office.
- Key areas of responsibility for the evaluation function sit with the Assembly or with the Director General (DG) *following approval from* the Assembly. Assembly responsibilities and approvals include:

- ➢ Hiring Head of Sector / Unit for the function
- ➢ Validation of the final evaluation plan and selection criteria
- ➢ Approval of the function budget and staffing structure
- ➢ Approval (following presentation) of key function products, such as an annual report or review on the learning or main findings / trends highlighted by evaluations
- The function is situated under the line management of the DG for all other activities.
- There is a process for presenting information from evaluations directly to the Assembly (without DG involvement), as well as to the DG and the Directorate.
- Function may be a fully stand-alone office or co-located with other executive management functions.

Additional details provided in Annex VI, 'Evaluation Function Purpose and Location Recommendations Report'.

RECOMMENDATION 3: **Resourcing the evaluation function**

Staffing: *Who works in the function? What do they do?*

The Assessment reviewed staffing models according to the options explored for the evaluation function purpose and scope (see Recommendation 2 and Annex VI), including staffing implications for a function that is designed according to Command / Control, Center of Excellence, or a Hybrid approach. Based on the Assessment recommendation to structure the evaluation function according to a 'hybrid model', ICRC should adopt the following structure for staffing the central office of the function:
- Head of Sector / Unit with significant experience in humanitarian organization 'center of excellence' models, significant experience in humanitarian evaluation systems, and a track record of successful team or agenda management. Ability to network and persuade is critical.
- At least three staff with experience in organizational evaluation systems to develop organizational-level guidelines and to support: evaluation planning, evaluation management, evaluation quality control procedures, and evaluation use. Two of these positions dedicate time towards organizational capacity building and change management in the first two years of launching the function.
- One or two of these positions has evaluation expertise in qualitative methods. This assumes the function is also staffed with the positions that have quantitative expertise currently sitting with PME.
- Equip the function with the technical expertise necessary to demonstrate the possibility of evaluation across different types of ICRC work and contextual challenges. Consider hiring a dedicated 2-year post with expertise in contribution analysis, outcome harvesting, and outcome mapping to work specifically with the protection, prevention, and movement / cooperation métiers in developing and testing measurement approaches for evaluating programs with long time horizons or unclear outcome / impact level objectives, 'impact' evaluations, or any evaluation that includes DAC criteria on effectiveness. This would complement work underway on quantitative methods for impact evaluation.
- In Years 3-4 of phasing the evaluation function, ICRC may want to consider developing a staff roster to support reviews or mixed-team evaluations. A consultancy roster should be developed within Years 1-2.
- Total number of staff: 4 to 5 (the 5th being a 2-year 'consultant') in the central office, including the Head of Sector (who may share responsibilities with planning, monitoring, reporting).

Details on staffing recommendations, including capacitating the function beyond the central office and connecting the evaluation function to existing M&E or RBM staff, are provided in Annex VII.

Due to the lack of consistent definitions or structural parameters for the regional level of ICRC operations (e.g. the Regional Model such as NAME vs. Regional Delegations, placement of regional red line vs. blue line functions, etc.), the delegation-level recommendations for evaluation staffing structure focus on the need for field capacity without specifying where this capacity should be located. When deciding where to place field-level staff capacity, ICRC should take the different regional approaches into account.

Budgeting: *Where does the money sit, where does the money comes from, and just how much money will it be?*

Budget ownership: The Assessment reviewed three models for budget ownership: Centralized / Controlled, Split Ownership / Controlled, and Split Ownership / Partial Control. Based on the Assessment findings, **ICRC should move towards a 'split-control' model for evaluation budget ownership**. This will encourage an increase in user-driven, demand-led evaluation practice across the organization, while also ensuring that a core set of strategically determined, fully quality-assured evaluations are initiated and resourced from within the evaluation function. This involves:

- Three types of dedicated evaluation funding 'pots': (1) held by the evaluation function; (2) held by the five HQ Departments; (3) a 'delegations' pot managed by the Director of Operations. It is preferable that the HQ Departments manage their 'pots' directly. However, if this is regarded as an unfavorable administrative burden, it is possible for the evaluation function to administer the budget on their behalf.
- Funds follow a 'use it or lose it' rule – if the funds are not used for evaluation, they return to a 'central pot'. The dedicated budget is ringfenced and cannot be (re)allocated.
- The evaluation function pot is for centralized evaluations that are managed or strongly supported by the function. Funds for 'decentralized' evaluations are held by the HQ Departments for Geneva-initiated decentralized evaluations and by the Director of Operations for funds dedicated to decentralized delegation evaluations.
- A clear budget or account code ensures control over and visibility of funds and spending.

Budget source: ICRC needs to determine the source of evaluation funding, namely a re-allocation of core funds, a line-item request to donors, or a combination of the two. Based on the Assessment findings, **ICRC should consider possibilities for a 'combination' approach** while recognizing the need to be strategic in how it approaches donors for specific evaluation funds. Details on the relative pros and cons between the two are listed in Annex VII. A combination approach involves:

- Line item request to donors on specific aspects of the development of the function (e.g. 2-3 'groundwork' evaluations and a possible 2-year fixed post on contribution analysis, outcome harvesting, and outcome mapping expertise development), the visibility of which to donors would be an advantage to ICRC. Select flexible and accommodating donors for this, e.g. Scandinavian and German MFA. Avoid DFID, ECHO, World Bank.
- Additional costs are funded through core budget envelopes. This includes the proposed 'pots' of funds described above and the costs of any existing or newly hired function support staff.

Budget amount: ICRC needs to establish how it determines the amount of funding it dedicates to evaluation. Two models reviewed by the Assessment include the use of allocation benchmarks and allocation according to a number of anticipated evaluations. Based on the Assessment findings, **ICRC should follow industry standards for allocating a set proportion of its total operating budget towards evaluation**. This includes the following actions:

- Determine a realistic, but robust benchmark for the allocation. The UN JIU suggested standard is 0.5 – 3% of operational budget going to evaluation. However, benchmarking against UNHCR - the UN agency most like ICRC in terms of scope – provides a more realistic target as a way of starting in Years 1-2 of the function, which should be increased as the function is phased up in Year. UNHCR currently funds evaluation at around .07% (2018) of its operating budget. In 2017, this amounted to 3.58 million[32]in funding for evaluations.

- Agree on the required amount according to a phased approach to ensure absorption capacity and to allow time for the organization to develop the expertise in evaluation practice necessary to responsibly manage the funds.
  - ➢ **Years 1 and 2**: Benchmarking ICRC to UNHCR for Years 1 and 2 of the function suggests that a starting evaluation budget level 'commensurate to the size and function'[33] of ICRC is **1.4 million CHF.** This would cover the costs of staffing the centralized office of the evaluation function *and* the budget for individual evaluations for the two 'pots' of funds held in Geneva. It does not include the costs for staffing departments or establishing staff capacity for the field and does not include a budget for evaluations at the field level (the third 'pot'). The Assessment recommends re-shaping the structure and capacity of existing RBM and M+E staff (e.g. take what exists and organize it differently), rather than hiring additional new people (see Annex VII). If new positions at these levels are required, however, at the start of the evaluation function (e.g. ICRC is unable to re-organize existing staff capacity), then the allocation for Years 1 and 2 should increase to around 0.1% of ICRC's operating budget, or an estimated amount of 2 million CHF. This would enable hiring three to four additional full-time staff in priority placements at the departmental level[34].
  - ➢ **Year 3**: The budget allocation and amount should increase to create the dedicated delegation evaluation budget that was not included in Years 1 and 2. This assumes that ICRC has established sufficient centralized capacity to govern and support evaluation practice across the organization by the second year of the function. The amount of increase depends on whether additional staff were required at the departmental level. It is anticipated that the Year 3 allocation will sit at either 0.1% of the total operating budget (2 million CHF) if additional hiring was not required in Years 1 and 2 or at 0.13% (2.6 million CHF) if additional hiring was needed. This would enable both a sufficient amount of funds to cover decentralized delegation evaluations in Year 3, including the cost of staffing (if required) to support expanded field capacity and the cost of individual evaluations. As with the departments, it is recommended that ICRC explore ways to re-organize existing staff capacity for field support rather than hire new positions.
  - ➢ **Years 4 and 5**: In Year 4, ICRC should review the allocation amount it is using to determine if it is sufficient and adjust as needed. On balance, the organization should aim to have a stable and set allocation by Year 5 that it will continue to use in the years going forward, allocating no less than 0.1% and no more than 0.3% of total organizational expenditure towards evaluation work.

---

[32] See UNHCR 2017-2018 MOPAN Assessment, page 44: http://www.mopanonline.org/assessments/unhcr2017-18/UNHCR%20report%20[web-1a].pdf
[33] UNEG Norms and Standards, Norm 13 point 17
[34] Assuming the cost of one FTE is 160k CHF/year, as per correspondence with ICRC.

RECOMMENDATION 4: Information requirements for the Evaluation Function

Governance: *What does the function need to know to do its job?*

The evaluation function is responsible for managing ICRC's evaluation strategy and providing overall quality assurance of evaluation practice. Fulfilling these responsibilities requires transparent and consistent information sharing from units outside the function's line management control. It also requires a degree of information tracking and systems within the function itself. This includes:

- **Tracking and quality assurance for evaluations**: Access to information on evaluation planning / evaluation section of the PfR (decentralized evaluations), access to evaluation ToRs and budgets, evaluator bids and contracts, evaluation reports and other key products, evaluation management responses, and action on the commitments agreed to in the management response.
- **Manage the implementation of the strategy**: The creation of annual evaluation coverage and frequency mapping, finance tracking and spending trends on evaluation, compliance tracking for evaluation requirements (e.g. traffic lighting adherence to evaluation requirements, such as the completion of management responses), and the ability to create a roster of consultants hired for ICRC evaluations.

A semi-regular health check exercise is also recommended. This can be done internally as a self-assessment review. The template used for this Assessment (see the Health Check Report Annex V) could serve as a baseline to monitor progress over the next 2-3 years.

Promoting evaluation use: *What information does the function need to promote evaluation use?*

The evaluation function is responsible for promoting the use of data on and from evaluations in a range of functional areas of the organization. Fulfilling this responsibility requires:

- Ability to access evaluation information from all levels of the organization and to produce information based on evaluation reports.
- Capacity to develop a complete evaluation report library.

RECOMMENDATION 5: IT systems and tools to leverage evaluation practice

Digital application: *How can the function leverage technology in the service of evaluations?*

The Assessment recommends three interrelated areas of evaluation management and use where IT application would provide a significant benefit to the evaluation function:

Management of individual evaluations. Process of managing centralized or decentralized evaluation workflow. It requires working through the following steps consistently:

- Evaluation planning and decision-making (approval of evaluation initiation, individual evaluation budgets and budget source)
- Evaluation Terms of Reference (developed according to guidance/template and uploaded to central repository).
- Evaluation inception reports (developed according to guidance/template and uploaded to central repository).
- Evaluation reports (developed according to guidance and uploaded to central repository).
- Evaluation management response (developed according to template, uploaded to central repository).

Evaluation function operations. The fulfilment of responsibilities by a centralized evaluation function charged with managing and maintaining ICRC's evaluation strategy, agency-level guidelines, overall quality control of evaluation practice. It also involves direct management of centralized evaluations. This includes:

- Managing the evaluation process for *centralized* evaluations (having direct control in managing evaluation ToRs, recruiting of evaluators, overseeing the evaluation process and approval on evaluation products, steering the evaluation management response, fulfilling tasks around evaluation publication and dissemination).
- Tracking the process of and providing a degree of quality control for *decentralized* evaluations (having access to information on evaluation planning and access to evaluation ToRs, evaluator bids and contracts, evaluation reports, and evaluation management responses).
- Systematic extraction and categorization of key information (e.g. evaluation findings or recommendations, information about the evaluations, etc.) for use across the organization.
- Systematic archiving of all evaluations according to a consistent set of evaluation parameters (such as keywords, geography, program type, subject, date etc.).

Evaluation use (*beyond individual evaluation response*). The use of data on and from evaluations in a range of functional areas and by positions across the organization. This includes:

- Visualizing evaluation coverage and frequency (by subject, geography, date etc.). This is important for centralized evaluation planning and prioritization (e.g. ensuring that sufficient evaluation coverage is achieved) as well as for ensuring staff are aware of evaluations relevant to their work.
- Visualizing data on evaluation management responses (through progress traffic lighting, for example) to ensure transparency and promote action on accepted evaluation recommendations.
- Presenting and disseminating summaries of findings and recommendations from evaluations. This can involve *proactive* dissemination of evaluation information to users based on defined position attributes (level, position, area of work etc.), as well as incorporating evaluation findings and recommendations into existing dashboards on organizational performance (agency, department, unit, team, function, or delegation level).
- Linking evaluation findings to dashboards and data generated within the monitoring / MfR reporting platform (e.g. evaluation results are dashboarded with relevant MfR data according to the level of analysis, such as delegation specific, métier-wide findings, or organizational trends)
- Ensuring evaluation reports are accessible from a central report library and on the ICRC intranet in a manner that is searchable using a consistent set of search parameters (geography, métier(s), evaluation type, date, etc.). Report library should be linked to existing or planned knowledge management platforms.

With the introduction of relatively simple database and warehousing tools, ICRC can leverage existing SharePoint and Tableau software to meet most of these needs. SharePoint can be configured to manage evaluation workflows and to store evaluation files, while Tableau can be used to visualize and dashboard evaluation information. This requires the development and maintenance of basic database tools for organising and warehousing evaluation data.

Additional details and a full body of recommendations including anticipated data dependences and a proposed system for IT application for evaluations is provided in Annex VIII 'Information Technology and Evaluation Recommendations Report'.

RECOMMENDATION 6: Evaluation Strategy

Evaluation framework: *What is an evaluation, and who gets to do one?*

ICRC's evaluation strategy should address levels of centralization and decentralization in evaluation practice and describe the purpose and placement of different types of evaluation and like activities. It needs to include the appropriate scope and lines of inquiry for each evaluation type.

The following parameters are recommended:
- ICRC creates two levels of evaluation – decentralized (managed outside the evaluation function) and centralized (managed directly by the function and/or directly supported by the function including steering committee participation).
- Evaluation types include:
  - **Decentralized evaluations** commissioned by delegations, metiers, and departments and conducted by external consultants or a mixed team (ICRC staff and externals). The scope of the exercise includes main programmes (Protection, Assistance, Prevention, Movement), thematic operational dimensions (target population, response type), delegations (single delegation or multi-delegation regional scope), departmental policies or projects.
  - **Decentralized reviews** commissioned by delegations or metiers and conducted by a mixed team (ICRC staff and externals) or a fully internal team. The scope of the exercise includes main programmes, delegation-level thematic questions (target population within one country, response type within one country, etc.), departmental projects. Decentralized reviews are <u>not</u> used for: departmental policies or thematic operational dimensions cutting across more than one country or context. The purpose of reviews is to capitalize on staff learning and facilitate staff reflection.
  - **Centralized evaluations** commissioned or strongly supported by the evaluation function in collaboration with stakeholders and conducted by external consultants or a mixed team (depending on the type of evaluation). There are two types of centralized evaluations:
    - *Strategic evaluations* conducted by an external consultant or a mixed team. Scope of the exercise includes thematic questions that link to global policies or strategies. Can also include methodological development evaluations, such as piloting new measurement techniques with a metier or supporting an innovative approach to evaluation.
    - *Institutional evaluations* conducted by an external consultant only and always managed by the evaluation function. Linked to ensuring coverage commitments, donor requirements, or risk management parameters. Scope of the exercise can include programmes (Protection, Assistance, Prevention, Movement), support units, emergency response, a single delegation, or a specific initiative.
  - **Centralized reviews** commissioned, as needed, by the evaluation function in collaboration with stakeholders to respond to a specific need within a rapid, flexible timeframe and conducted by a mixed team or fully internal team. Team arrangements often involve a member of the evaluation function. Scope of the exercise includes institutional projects or organizational experiences that would benefit from rapid capitalization of staff learning. An example of this would be the Ebola Lessons Learned Report (which was more 'review' than 'lessons learned').

- Lines of inquiry are levelled according to the type of evaluation activity. DAC criteria for evaluation questions are included in the focus of every *centralized and decentralized evaluation*, extending beyond 'effectiveness' to include questions of relevance, efficiency, and, where appropriate for the length and objective of the action, impact and sustainability. Additional lines of inquiry recognized by ICRC for their strategic or operational value should be established and consistently used in evaluation ToRs. As a start, this includes a question tied to the AAP framework and accountability to affected populations and a question about protection results or quality of protection mainstreaming in the action evaluated.

  In the case of *centralized and decentralized reviews*, the lines of inquiry include: what did we plan to do, what did we actually do, what were the challenges / obstacles, what have we learned, etc. DAC criteria are not applied for reviews. Lines of inquiry on AAP and protection should be considered for incorporation into reviews according to their value in capitalizing staff learning or reflection.

As per Recommendation 1, the evaluation function will have authority to approve the evaluation plan for decentralized evaluations. This includes cancelling or shifting exercises that are outside their appropriate scope.

Evaluation standards, structure, and planning: *How are evaluations done?*

The evaluation strategy needs to address guiding principles for evaluation practice, including independence, ethical conduct, evaluation utility, transparency, and quality assurance. The policy framework for evaluation should include the definition and purpose of evaluation, selection criteria for evaluation, coverage standards and responsibilities, and requirements for evaluation workflow, documented evaluation responses, follow up and use of evaluation findings, and report publication procedures according to the AIP. Areas of potential alignment with Movement actors will be highlighted as relevant.

It will also include the considerations for the purpose, scope, and placement of the evaluation function proposed under Recommendation 1 and the resourcing of the function, in particular the budget arrangements, proposed under Recommendation 2.

A proposed approach to phasing the development and launch of the evaluation function is annexed to the end of the strategy. It includes key considerations for Years 1 through 5 of the development of the evaluation function. A critical part of this phasing is positioning evaluations to better enable ICRC's outcome monitoring / outcome measurement efforts. Of particular concern is identifying and commissioning evaluations that can help Protection, Movement, and Prevention target and test appropriate outcome indicators and develop methodologies for measuring them. As the organization solidifies its outcome monitoring, evaluations should shift focus from the outcome to the impact level for these metiers.

Ways to leverage existing ICRC incentive or compliance systems (Finding 13.3) for improved use of evaluation for learning and accountability is also annexed to the end of the strategy.

Details and the full proposed strategy are found in Annex IX, 'Evaluation Strategy'.

RECOMMENDATION 7: Additional areas.

Additional areas: *What else is there to consider for a successful evaluation system?*

In addition to the recommendations framed in response to the Assessment objectives, a number of additional areas are noted for ICRC's consideration:
- Include evaluations in the AAP Framework. This includes (1) ways in which evaluations can produce learning on AAP for the organization and complement the self-assessment exercises and (2) role of evaluations in ensuring accountability to affected populations.
- The evaluation function should work with the CORE team and Internal Audit to establish routine connection and collaboration points across all learning and accountability functions in the organization. Objectives include streamlining protocol on similar exercises (e.g. ToR development) and exploring opportunities for information and process harmonization.
- Consider areas of 'content harmonization' that can be applied across all evaluations to support improved evidence synthesis. For example, include a key question about protection in every evaluation during the strategy period. Start this approach in Years 1-2, and use the generated evidence in the development of a learning review in Years 3-4.
- Articulate the specific type of evidence from evaluations that ICRC should extract or synthesize to complement or enhance learning from MfR reports. Build ways of presenting the monitoring and evaluation learning together as part of a single performance or learning story.

5. ANNEXES

The following Annexes are included with the Assessment:

Methodological Summaries and Products:
- Annex I Evaluation Mapping 2016 to 2019 (*Attached in a separate file to the report)*
- Annex IIa Case Study Summary Report
- Annex IIb Case Study Score Sheet (*Attached in a separate file to the report)*
- Annex III Key Informant Interview and Focus Group Discussion Summary Report
- Annex IV Workshop Summary Report
- Annex V Evaluation Performance and System Health Check (*Attached in a separate file to the report)*

Recommendation Reports:
- Annex VI Evaluation Function Purpose and Location
- Annex VII Evaluation Function Resourcing
- Annex VIII Information Technology and Evaluation
- Annex IX Proposed Evaluation Strategy (*Attached in a separate file to the report)*

Assessment Process Documents: *Attached as separate files to the report*
- Annex X Assessment RFP
- Annex XI Inception Report
- Folder: Complete set of the gathered and referenced documentation reviewed for the Assessment

Annex IIa: Case Study - *Summary report and conclusions*

INTRODUCTION

Four evaluation reports were benchmarked for the quality of their process, content, and use against accepted industry standards with a view to identify gaps in the current evaluation practice and derive conclusions for the future evaluation function. In particular, the factors that enhance or impede evaluation quality, independence and utility were explored along with perceptions on 'success' or 'failure' for the evaluation. This annex provides details of the case studies along with a set of conclusions informing the final report and recommendations of the Assessment.

APPROACH DETAILS

Table 1: Case studies[35]

#	Report title, country, date	Function/ Sector	Reasoning for inclusion into the Case Study	Supposedly Poor/ Good practice
1	Evaluation of Diversity, Inclusion, and AAP in ICRC Operations, Multiple Countries, 2018 (hereafter: AAP case)	Transversal	KI noted 'good usage and follow up on the recommendations, team is happy with the process.' 10 respondents completed the survey.	Good
2	Assessment on the involvement of ICRC resident staff in the Protection Activities, Philippines, 2016 (hereafter: Protection Philippines case)	Protection	KI: There were specific recommendations regarding a single staff member that should have been dealt with outside the content of the report. 9 respondents completed the survey.	Poor
3	Healthcare in Danger, Analysis of the Field Planning for Result Exercise (PfR), Multiple Countries, 2018 (hereafter HCiD case)	Transversal	KI: 'It was an example of good usage; everyone was happy with the consultants and the report.' 4 respondents completed the survey.	Good
4	EcoSec Review, South Sudan, 2017 (hereafter EcoSec SSD case)	EcoSec	KI: 'S. Sudan evaluation led to some concrete changes in the way we do EcoSec today. Notably the observations around how we dealt with food assistance (scattered and with limited impact) led to more targeted assistance, and better integration of nutrition considerations.' 4 respondents completed the survey.	Good

---

[35] Only the reports dated no more than 3 years ago qualified for the case study. The initial case study set included eight evaluation reports. However, only these four received a sufficient survey response.

The evaluation reports detailed in Table 1 along with related documentation (ToRs, Management response files, etc) were examined and assessed against industry standards[36]. The Assessment team exercised professional judgement in determining whether the report met industry standards. A staff mini-survey complimented the study, with 27[37] ICRC staff members involved in commissioning, managing, or using the selected cases completing the questionnaire. Additional interviews and email enquiries were conducted to address gaps in documentation.

Each case study was assigned a score for 'Overall Meeting Objectives', 'Utility', 'Quality' and 'Independence' criteria in accordance with the pre-defined rubric, with a total maximum possible score of 100 (i.e. 25 in each criteria). The score was then translated into 'weak', 'moderate' or 'strong' assessment result in each criterion, as per the rubric framework.

The results are analyzed to identify patterns (i.e. current gaps), along with factors that enhance or impede evaluation utility, quality, and independence.

OVERALL OUTPUT

The following table summarizes the outputs of the case study analysis. Further details *per case* are included under 'Analysis Details – Cases'.

Table 2: Case Study Output

Finding Category	Case Study Findings
Organizational Performance	➢ Two out of four cases, namely AAP and HCiD cases, scored above 70 out of 100, with 'strong' record in 'Overall meeting objectives', 'Utility' and 'Quality' criteria. Notably, the AAP case also performed strongly in 'Independence'. EcoSec SSD case received a 'strong' reading in 'Quality', which makes three out of four cases assessed as high for quality. One common factor between AAP and HCiD cases is that they were supported by the PME Sector, particularly with the evaluation process. The EcoSec report also benefited from staffing and guideline investment within that métier. ➢ This demonstrates that some of the evaluations currently produced at ICRC are highly regarded in every aspect (i.e. perceived 'successful') by both internal users/commissioners and according to widely accepted industry standards. This happens, in part, when they are supported by a function or a team with evaluation expertise. Two leading reasons that determined success of the evaluation exercises by survey respondents were the 'Participatory approach' and the 'Involvement of the stakeholders' (both chosen by >60% of respondents). The third top reason, chosen by half of the respondents was the 'Qualification and experience of the evaluators. ➢ Utility of the evaluations varied, with the Protection Philippines case having the weakest score of 8/25. Utility was strongly linked to adherence to an evaluation process, including inception phase (or its absence), involvement of key stakeholders prior, during and at post-evaluation phase, with established governance/roles/responsibilities and a management response to recommendations formatted as an action plan. ➢ Guidelines on evaluation process and practice, with clearly defined and assigned roles and responsibilities of stakeholders, including steering committee would benefit utility and, as a result, promote operational impact. Pre and post evaluation phases should be strengthened, with

---

[36] Defined in Evaluation of Humanitarian Action Guide, ALNAP, 2016 and Evaluation Manual, MSF, 2017.
[37] The survey has been distributed to 207 potential respondents, 30 of which completed it;
27 out of 30 respondents participated in evaluations from the case study set, three respondents participated in other evaluations. The answers of the latter have been included in the overall analysis, but not in the score.

	particular attention to (1) agreement on the objectives, purpose, evaluation questions between all of the stakeholders, (2) inception phase to clarify the scope, limitation and ensure alignment of expectations/focus, (3) follow-up on recommendations/management response. ➢ Currently there are gaps in evaluation practice, resulting in some reports 'missing objectives', other involved 'self-review' without sufficient mitigation measures in place to ensure independence and safeguard against conflict of interest, lack of follow-up on recommendations, limited visualization of findings coupled with overlong documents. Further details on these points are found in the individual case study reports. All the survey respondents (100%) whose evaluation experience was <u>unsuccessful</u> stated that it was due to 'Insufficient follow-up at the post-exercise stage'. Arguably related, the 'Lack of stakeholders' interest before or after the exercise' and the 'Lack of involvement of the stakeholders' were noted as reasons for failure by 80% and 60% of the respondents. Notably, the 'Low qualification and lack of experience of the evaluators' have never been chosen among the failure factors.
Organizational Capacity	➢ On resource capacity, survey respondents were asked if resourcing in terms of budget, time, field staff availability and commissioners' availability was sufficient to manage the evaluation. Many respondents (20-40% depending on the resource type) were not able to answer the question. However, only less than 10% of the respondents could state with confidence that resourcing was not sufficient. Hence, resourcing does not appear to be a consistently present issue for hindering evaluation success once the budget is secured and the evaluation is initiated. ➢ Given the evaluation is currently conducted in a de-centralized manner, producing a quality evaluation is arguably resource draining. This is illustrated by AAP case, where the RfP had to outline all the entire evaluation process in such a manner that it effectively served as a set of evaluation guidelines (the RfP is very comprehensive, 25-page long). In the absence of centralized guidelines, these efforts are duplicated in every métier that commissions an evaluation. This needs to be streamlined and centralized to ensure rational use of resources, shortened turnaround of evaluation process, and aligned expectations for organizational best practice or strategy. ➢ Some performance gaps (especially around evaluation scope and lack of inception phase) point to the need for training on 'how to enhance evaluation utility and what to expect from evaluation'. There appears to be a knowledge gap –periodic training, information sessions, or workshops can address this issue and improve knowledge on evaluation. It could also support the development of an evaluation culture. ➢ In addition, there appears to be a gap in 'knowledge management', which hasn't been routinely practiced in the organization.
Organizational Culture	➢ Some (24%) survey respondents expressed concerns about the ability of independent consultants to understand ICRC's work (referred to as ICRC's mandate/action/internal functioning/working systems/the codes of institutional culture/way of working/institutional complexity) and develop actionable recommendations. At the same time, almost half of the respondents expressed concerns about conflict of interest and inherent biases when evaluations were conducted by internal staff or even a former staff member. This affects perceived credibility and hinders evaluation use. ➢ A solution to this can be increased use of mixed evaluation team composition, where an external consultant ensures sound methodological input, independence, and different perspective, while the internal member navigates organizational complexity and the context of the action.
External Influence	➢ Two of evaluations in this case set were organizational commitments to DFID, made in response to the 2016 Partner Assessment and included in the Performance Framework signed with the donor shortly thereafter. Each of these two evaluations received direct support from the PME Sector, including guidance on the overall process the evaluation should follow, development of the RfP/ ToR, establishment of the steering committee, active involvement of the PME sector in the evaluation steering committees, and templates / protocol for the management response. ➢ This connection to DFID and PME Sector support was not directly referenced in the evaluation reports or raised in the case study survey. It was, however, noted during key informant interviews with the PME Sector, REM, and staff representing the teams that commissioned the evaluation.

1. On development of a comprehensive Evaluation Strategy and Guidelines, including sections on context, evaluation process, governance, roles and responsibilities, objectives, purpose, report requirements, templates, ToRs, recruitment process, etc.
   ➢ Notably, the **EcoSec** unit issued a Handbook on Planning, Monitoring and Evaluation a few months after the Review included as a case in this analysis. The Handbook provides some definitions, details on the process, and useful templates. This can be a valuable source for development of a more comprehensive policies and guidelines.
   ➢ The success of **AAP case** was partially attributed to clear Role and Responsibilities, Governance structures, and report expectations set in the RfP. Development of such comprehensive RfP is supposedly resource consuming. Some of the elements need to be standardised to improve use of resources and overall turnaround of the process. The RfP of this case study can be consulted for development of the standard evaluation process and practice, as proven successful.
   ➢ **Based on Protection case:** A guideline should include a section on report requirements, including a clear report structure and enhanced links between objectives, purpose, evaluation questions, findings and conclusions.
   ➢ **Evaluation questions (EQs)** were consistently not used for all of the cases, which arguably makes the evaluation process and expectations on the result less certain. It is recommended to include EQs as a part of ToR requirements.

2. On Capacity Building
   ➢ **Based on AAP case study:** Some guidelines or training should be provided to the commissioners to educate on scope-development and expectations from the exercise.
   ➢ **Based on Protection case study:** Evaluators, especially the internal reviewers without formal qualifications in evaluation, should have access to training on evaluation report writing with a view to improve utility.

3. On Inception Phase
   ➢ **Based on EcoSec case**: The lack of inception phase and no documented management response point at gaps in the EcoSec review process. Addressing them could result in increased usability and alignment of the expectations between all of the stakeholders.
   ➢ **Based on Protection case:** To mitigate the risk of evaluation report missing the focus, a phased evaluation process should be established, with an inception phase to clarify objectives, the purpose and the scope.

4. On visualisation
   ➢ **Based on EcoSec case study:** Further visualisation can lead to increased utility and aid dissemination of results to wider stakeholders.
   ➢ **Based on AAP Case Study:** Data visualisation requirements (e.g. a printable poster with findings summary) can be considered for inclusion into the standard report requirements.

5. **On ensuring impartiality**
   ➢ The evaluation process in **HCiD case** was rather simplified due to evaluator's experience with previous exercise. However, some guidance, especially related to governance and impartiality of the evaluator would potentially benefit and add credibility/decrease risks. Hence, it is recommended to include guidelines on consultant recruitment policies, including considerations on independence and risks in instances when self-review is performed. This can be complimented by mitigation measures on applying appropriate safeguards, such as forming a steering group.
   ➢ Given institutional complexity, a **balanced evaluation team composition** is recommended to ensure that independence doesn't compromise usability.

6. **On ethical aspects of an evaluation**
   ➢ **Based on Protection case:** It's imbedded in the profession that some recommendations might be sensitive. However, professional ethics should be respected – this should also be included in the guidelines for evaluation practice.

ANALYSIS DETAILS - CASES

Case 1. Evaluation of Diversity, Inclusion, and AAP in ICRC Operations - Case Study

Source	Study highlights	Potential gaps/Concerns
Request for Proposal (RfP)	• RfP is comprehensive, clearly detailing the Context; Objectives; Process and timeline; Roles & Responsibilities, and the Governance structure; Expectations and Requirements. • Formulation of the Evaluation Questions (EQs) is intentionally shifted to the consultants, a sufficient background information is provided. • Intended report use is not explicitly addressed, though the purpose can be derived from the objectives section. • There is an outstanding level of guidance on the Process and Requirements in the RfP (can serve as a guideline). The Governance appears well-established, with a project team, Steering group and Deputy Director, Operations having roles and responsibilities and being accountable. • Visualisation and maximum lengths of report elements are stated in expectations (and not quite met in the report).	• A very wide scope; • EQs not formulated in the RfP but rather shifted to the evaluators; Should there be an established process/guideline, the RfP could omit some of the details (perhaps, refer to the evaluation policy document) and be more to the point.

Evaluation Report	• EQs are well-formulated. The report clearly addresses three (policies (EQ1), current practice (EQ2) and recommendations (EQ4)) out of four questions, with lacking explicit findings on 'factors affecting performance', i.e. EQ 3. Report structure reflects the EQs. • Well-balanced methodology with a decent coverage in all aspects – interviews, doc reviews, surveys, FGDs, etc. Field data was collected over visits to five countries and 18 counties participated in survey. • Findings visualisation is implemented to some extent but could go further, e.g. with adding a 'summary page of facts and figures' to highlight main findings and critical recommendations'; a poster format finding with infographics would also benefit. • Recommendations are comprehensive and provided separately, as requested in the RfP. In terms of actionability and acceptance – out of 20 recommendations, seven were accepted, 11 – partially accepted and two receive an 'under-consideration' status as per the management response. Notably, none of the recommendations were rejected, which is possibly a sign of a sufficient consultation process during all of the evaluation phases. For all of the recommendations, a responsible party, practical implementation steps and a deadline were assigned.	• Every component of the evaluation report is ~twice longer than the requirement specified in the RfP (e.g. total number of pages 157 vs 80 in the RfP); • Visualisation, while excellent where exists, could be improved and extended to a format of a 'summary page' of the critical points of both findings and recommendations. This would improve the overall usability, especially given the length of the report.
Survey	• Highest number of respondents (10); • Independence is perceived as 'strong' (5 out of 5) by most respondents with impartiality of the evaluation team being the main factor; • The overall 'objectives-meeting', 'evaluation quality' and its 'utility' were all assessed as 7 out of 10, which is rather strong but not outstanding. The main factors contributing to success of the evaluation were the participatory approach, involvement of the stakeholders and a good management response and follow up on recommendations, all of which are arguably interrelated. Notably, the qualification of the evaluators was highly appreciated by many respondents; • Evaluation is perceived as sufficiently resourced in terms of all the budget, time and availability of the field staff and the commissioners.	Survey response critiqued 'unclear evaluation purpose' (which correlates with the assessment of the RfP) and the 'evaluation bringing little new inputs'.

Wrap-up

The case scored 'strong' in all fours criteria, including 'Overall meeting objectives and adding value' (17/25), 'Utility' (17/25), 'Quality' (18/25) and 'Independence' (19/25). The critical success factor was the **well-developed and thought-through evaluation process** at every stage, from the RfP development to management response on recommendations. This process was detailed in the RfP and determined involvement of stakeholders into the evaluation and established a strong governance along with accountability, including at post-evaluation stage. Sufficient resourcing and involvement of both field staff and commissioners have strengthen the exercise. Independence and qualification of the evaluators were also highly appreciated.

The gaps were identified around the scope (very wide, which made impossible to address all of the objectives and led to the report twice longer and 5-months later than expected). Further data visualisation (that was requested in the RfP), especially around recommendations would help improve utility.

Case 2. Assessment on the involvement of ICRC resident staff in the Protection Activities, Philippines, 2016 - Case Study

Source	Study highlights	Potential gaps/Concerns
Terms of Reference (ToR)	- The assessment appears to have a strategic goal to maximize the impact of protection work in the Philippines. - The ToR include details on the Background, Objectives of the Assessment, suggested methodology, tentative timeframe and some of the background documents. It's noted in the ToR that the consultant would have a freedom to choose methodology and to suggest other relevant documents. - The need for the assessment is also justified by the recommendation of the Protection Division in the Resident Mapping Protection Activities document, dated 2013. - Assessment deliverables are specified as 15-20 pg. report, one-page Executive Summary, and a power point presentation.	- The assessment covers a period of 22 years of activities. Arguably, this alone makes it challenging to affirm correlation and/or causality between the results and the initial decision to involve resident colleagues in protection activities that was made in 1993; - The ToR does not detail requirements for the consultant nor specify if the assessment is internal or external; - The purpose of the assessment is not explicitly stated but can be derived from the sub-objective 4 (for decision making and risk mitigation); - Evaluation Questions are not stated (arguably could be derived from sub-objectives); - Process (i.e. phases of the assessment, roles and responsibilities, etc) is not defined.
Evaluation Report	- The report focuses on HR set-up and related issues with comments on shortcomings of identifiable individuals. Some recommendations directly suggest certain HR decisions. - The nature of the report is HR internal review, rather than assessment of strategic issues. - Length- 50 pages without annexes vs 15-20 pages required in the ToR. - Structure of the report is confusing - Abbreviations not listed. There are some issues with grammar. The flow of information is rather sporadic. - Recommendations are largely HR-focused. - Executive summary largely repeats Conclusions section and can't be used as a stand-alone document.	- Main concern is that the report addresses predominantly HR issues – this applies to both Findings and Recommendations sections. While HR strategy is one of the assessment objective, the ToR has a wider focus, most of which is not addressed. - Poor structure of the report, which is arguably linked with no EQs posed in the ToR. It appears that the evaluator was given a freedom to change the focus of the assessment and its scope during the exercise. - Communication relies predominantly on text rather than visual tools, with the report being ~three times longer than specified in the ToR.

| Survey | Second largest number of respondents (9); Most respondents had mixed feelings about the value-added potential of the assessment and its utility, whereas the perception on independence ranged from 0/5 to 5/5. The degree of evaluator's expertise was highly regarded by 6 out of 9 respondents; No issues with resourcing were identified. | The gaps pointed at the 'Weak evaluation methodology and process' and 'Poor management response and follow up of recommendations.' Some respondents explicitly stated concerns that the 'recommendations might have been grounded in bias', the evaluator 'holds own beliefs which influenced the outcome of the evaluation' and that 'the manner in which the outcomes were used by the delegation was other than the purpose of the report'. |

Wrap-up

The case scored 'moderate' in 'Overall meeting objectives and adding value' (13/25) and 'Quality' (11/25), whereas 'Utility' and 'Independence' were assessed as 'weak', with 8/25 score.

The major issue with the case appears around the mismatch of the objectives as per the ToR and the eventual focus and purpose as developed/understood by the evaluator. As a result, the assessment's Findings and Recommendations focused on HR set-up and related issues, rather than the wider strategic and operational issues of involvement of ICRC resident staff in the Protection Activities. Some gaps appeared at the early stages of the assessment, namely a long assessment period (over 20 years), an evaluation process without an inception phase (or at least it hasn't been articulated), and evaluation questions not listed in the ToR. Arguably, these factors contributed to the shifted focus of the evaluation during the exercise. Consequently, the assessment was perceived as an 'instrumentalised' study by some of the respondents, especially due to the range of sensitive HR issues raised and some recommendations related to individual staff members.

Case 3. Health Care in Danger HQ evaluation 2018 - Case Study

Source	Study highlights	Potential gaps
Terms of Reference / Interview with the head of the HCiD initiative Maciej Polkowski	- This evaluation is a second in a series of HCiD evaluations. It is performed by the same evaluator as the previous evaluation. Hence, the ToR is less detailed and focused on the objectives and reviewed tasks. - Notably, the ToR includes clearly stated objectives (which determine the report structure), responsible parties (i.e. governance), background and suggestion on methods to use. - The scope of the evaluation appears reasonable. - No sections on expertise or experience requirements of the evaluator, which is understandable given the intentions to engage the same evaluator. - Evaluation process (phases) and Evaluation Questions are not specified. The latter can be derived from the objectives.	<u>Concerns:</u> Issues with impartiality of the evaluators and his conflict of interest given he reviewed his own work. Evaluation process is not well-defined – hence, no information on the inception phase, little information on governance, requirements to the report (length, visual tools, etc).

Evaluation Report	• Clear structure around ToR objectives, with conclusions and recommendations following individual findings. • Very detailed step-by-step recommendations, actionable immediately. • Good visual representation of findings, with appropriate tables and diagrams. • Mixed methods used, which appear appropriate for context and objectives. • Concise language, easy to follow. Excellent information flow. Comprehensive management response.	<u>Concerns:</u> No major concerns. While it is a strong and actionable report, the above issues on impartiality of the evaluator remain. However, given the 'continuation' nature of the report, complexity of the context, and level of satisfaction with the previous report – the choice of evaluator appears well-justified.
Survey	• Four respondents; • Unanimously assessed the evaluation as at least 3 out of 5 in all four criteria, namely 'meeting objectives', 'independence', 'quality', 'utility'; • 'Participatory approach' and 'involvement of the stakeholders' were named as success factors by all of the respondents; • Very positive comments on utility, e.g. 'The most important thing was that findings were immediately actionable', 'The exercise guided the development of the new HCiD strategy and raised important gaps', 'It was an evaluation that was not put into an archive, meaning it was used as a reference to develop the new strategy and address important gaps', 'Strong team buy in and expectation to reform things following a period of non-consultative management.' • Evaluation is perceived as sufficiently resourced in terms of all the budget, time and availability of the field staff and the commissioners.	<u>Concerns:</u> Not many concerns raised in the survey, except for one comment that the evaluation 'seemed too informal'.

Wrap-up

The case scored 'strong' in three criteria: 'Overall meeting objectives and adding value' (19/25), 'Utility' (19/25) and 'Quality' (19/25) with 'Independence' assessed as 'moderate' (14/25). The reasonable scope, involvement and interest of the key stakeholders, 'continuous' nature of the exercise, quality of information flow and report structure, and feasible and immediately actionable recommendations determined the overall success of the exercise.

The gaps were identified around independence of the evaluator, as the assignment involved aspects of self-review and there were no signs of 'safeguard' measures, such as a steering group or a fellow evaluator.

Case 4. EcoSec Review, South Sudan, 2017- Case Study

Source	Study highlights	Potential gaps/Concerns
Terms of Reference (ToR) / EcoSec Handbook on Planning, Monitoring and Evaluation (PME)	• The ToR specifies all the major parts of the assignment - its background, scope, objectives, Roles & Responsibilities (R&R), timeline. However, it lacks provisions for the inception phase, a clearly defined purpose, requirements to the evaluator, and R&R at post-evaluation stage; • The Handbook on PME issued by EcoSec was published after the case was completed (the Handbook was released in April 2018, while the case is dated January 2018). The Handbook provides guidance on PME definitions, process, and report templates. The Handbook is well-structured and clear, however, it's more comprehensive on Monitoring and Results-Based Management, than on Review/Evaluation exercises.	• Insufficient coverage of the review process, particularly the inception phase, and responsibility at post-evaluation stage in the ToR; • Absence of requirements or guidelines to follow during an internal recruitment process, especially ways to maintain independence or safeguard against conflict of interest (notably, this was not identified as an issue in the survey responses); • The Handbook is dated three-months post-case study, hence didn't guide it. However, it is noted that the Handbook's focus is primarily on monitoring and overall planning for results, with insufficient detail on evaluation and review type of exercises.
Evaluation Report	• Good report structure, clear objectives with findings, conclusions and recommendations addressing them. Recommendations appear actionable and well-developed; • Qualitative methods, combining KIIs and document review were used – appropriate to the objectives; • Visual illustration of findings is appropriate, but could be improved; • No management response on file. Key informants stated that the report 'led to some concrete changes in the way we do EcoSec today.' • The report is performed by a team of independent evaluators, one of them being a former ICRC staff member. This is considered beneficial for the outcome and not a significant threat to independence of the consultancy.	• Gaps around visualisation – could be improved; • The inception phase is not specified in the ToR and not referenced in the report; • Insufficient (documented) follow-up at post-review phase. Management response and implementation of recommendations took place according to the key informant. However, there is not enough documentation to verify this due to no management response on file.
Survey	• Four respondents; • 'Mixed' perception on overall 'meeting objectives and adding-value' and 'utility', with survey respondents having opposite opinions; • 'Participatory approach', 'Evaluation methods' and 'involvement of staff' were highly regarded, while 'insufficient follow up' is a failure factor; • Majority of respondents regard the case as 'sufficiently independent', with a remark that 'some of the team members were ICRC former staff' (arguably inevitable given complexity of the operational context); • The overall quality was positively assessed (3 out of 5) with 'Relevance of the evaluation findings and recommendations' being the key factor.	• Lack of consensus between survey respondents makes it challenging to identify gaps with certainty. • However, this case scored the least in terms of resourcing. This especially relates to such resources as 'field staff availability' and 'time allowed for the exercise' – both were deemed insufficient by most of the respondents.

Wrap-up
The review scored 'Moderate' in four categories – 'Overall meeting objectives' (12/25), 'Utility' (12/25), and 'Independence' (14/25); and 'Strong' (17/25) in 'Quality'. Notably, the survey results had limited value due to low number of respondents and their polar opinions. The quality of the report is found high (9/10), with the review successfully addressing the objectives and providing well-detailed recommendations. The report is well-structured, the evaluators appear very knowledgeable of the organisational context and of the programme. Feasibility of the recommendations should (and could) have been established by the management. The followed actions and their completion can't be verified due to the limited records on file. Gaps found around the lack of inception phase (unless it hasn't been documented) and no established (documented) governance/action plan at post-evaluation phase. Visual communication of the report can also be improved.

Annex III: Key Informant Interviews and Focus Group Discussions: *Summary report and conclusions*

This annex reports on the Key Informant (KI) interviews held with internal and external stakeholders, both in Geneva and in the field, and the Focus Group Discussions held in Geneva as part of the methodology of the ICRC Evaluation Function Assessment ('the Assessment'). It documents the questions posed and the participants. It provides a summary of responses which informed the final report and recommendations of the assessment.

The following semi-structured questionnaire was used during in-person, telephone, or Skype interviews with key informants. Questions were selected and adapted according to the position, experience, and understood knowledge of the informant in regard to ICRC evaluation practice.

1. Background *(ICRC staff–modified for externals)*
   1.1. Please describe your current position with ICRC, length and type of experience with ICRC.
   1.2. Do you have any experience with evaluation inside ICRC?

2. General questions (*all or most Interviewees*)
   2.1. What do you see as the main value or purpose of evaluation at ICRC *(alt: your section, your unit, your delegation)*?
   2.2. Do you think evaluations should be standard practice for the organization? What would 'standard practice' look like, in your opinion?
   2.3. Can you provide an example of an evaluation that worked well and one that was challenging?
   2.4. How important is independence in evaluation? What does 'independent' mean to you?
   2.5. How could ICRC improve in terms of evaluation support?

3. Staff from sectors / métiers / delegations with <u>active evaluation practice</u> or individuals with <u>a known evaluation history</u> in ICRC:
   3.1. In the most recent evaluation you commissioned / observed / participated in, what was your role? Who managed the evaluation process? Who conducted the evaluation? Who drafted and approved the ToR?
   3.2 How are evaluations commissioned in your team/ unit / delegation? Who decides an evaluation is needed, and who approves that decision? How are funds for the evaluation allocated?
   3.3. What aspects of the evaluation process work well? And what could be improved?
   3.4. How does your team / unit / section / delegation ensure quality control during evaluations? Are there areas that you feel are strong? Areas that are weak?

3.5. Have you been involved in or benefitted from any of the following:
- o   Internal Reviews
- o   Lessons Learned Exercises
- o   Research or studies

3.6. How do the above differ from an independent evaluation, in your opinion?

4. **Questions for external interviewees**

4.1. How is the Evaluation Function set up in your organization: How are findings and learnings used? Who 'owns' the management of evaluations – determination to evaluate, funding or budgeting of evals, scope, ToR, establishing a steering or review committee, selecting the evaluators, quality control, dissemination, etc?

4.2. Are here any known points where your organization and ICRC should *have consistency in eval definitions, types, or process? If so, why?*

4.3. Are there any challenges for evaluation unique to your organization?

5. **Questions for donors**

5.1. What do you see as the existing role of evaluations and evaluation functions in the organizations you fund? Specific reflections or distinct concerns for the ICRC?

5.2. How do you, as a donor, engage with the evaluation findings or recommendations that are shared about the programs or work done by the organizations you fund?

5.3. What questions or issues do you want to see evaluations addressing in the organizations you fund? How would this be different from other types of data reported or shared as part of a funding agreement (e.g. financial reporting, narrative annual or mid-term reports, etc.). Specific reflections or distinct concerns for the ICRC?

5.4. Are there any other specific issues or topics regarding evaluation at ICRC that you would like to raise or discuss?

KI PARTICIPATION

Key Informants included ICRC staff both at headquarters in Geneva and in the field, external interviewees from international organizations and donors. See the list of participants in Table 1 below.

Table 1: KI Interview Participants

ICRC Internal			
Name	Position	Department	Location
Antoine Quellet	Head of Sector, Planning, Monitoring and Evaluation	DIR GEN	HQ
Balthasar Staehelin	Deputy Director General	DIR GEN	HQ
Barbara Hunziker	Prot/Assist Organisational Manager	OP DIR	HQ
Benno Kocher	Germany	REM	HQ
Carole Dromer	Deputy Head of Unit - Health	OPS	HQ
Carole Pittet	Co prot Mexico	PROT	FD
Caroline Khoubesserian	Head of Unit - Prot Civilians	OPS	HQ
Caroline Putman-Cramer	Sweden & Finland	REM	HQ
Christian Wabnitz	Head of Financial Institutions	REM	HQ
Christoph Luedi	Head of Dakar Regional Delegation	HOD	Field
Christophe Driesse	Head of Sector – Economic Security Unit	OPS	HQ
Christophe Harnisch	HoD Bogota	HOD	FD
Christophe Martin	HoD Lebanon	HOD	FD
Claire Van Den Heuvel	Head of Project within Operations	OP DIR	HQ
David Loquercio	Head of Accountability to Affected Population	OP DIR	HQ
Eloi Fillion	HoD Abuja	HOD	FD
Erik Tollefsen	Head of Weapon Contamination	OPS	HQ
Esperanza Martinez	Head of Unit - Health	OPS	HQ
Eva Svoboda	Deputy Director	DP	HQ
Fabrizzio Carboni	Regional Director, Near and Middle East	DIR GEN	HQ
Fatuma Mohamed Adan	Risk Management Coordinator	SOM NAI	FD
Filippo Minozzi	Analysis & Evidence Advisor	ECOSEC	HQ
Fruzsina Di Ruggiero	US Market	REM	FD
Gabriele Onorato	Supra Regional Movement Coordinator	NAI	FD
Guive Rafatian	Policy and Humanitarian Diplomacy Ops Manager	DP POL	HQ
Hugo Van Den Eertwegh	Deputy Head of Unit - SCMS	OPS	HQ
James Reynolds	HoD South Sudan	HOD	FD

Jerome Daumas	Community Contact Center Project Manager	OPS	HQ
Jonas Baumgartner	Dep. Head of Internal Audit	AUDIT	HQ
Jules Kagwahabi Amoti	Head of Unit - Economic Security	OPS	HQ
Layal Horanieh	Norway	REM	HQ
Lilli Heinrichs	Head of Data	MOUV	HQ
Maciej Maksymilian Polkowski	Head of Health Care in Danger Initiative	OP DIR	HQ
Manuel Fatana	Head of Unit - FAD CORE	FAD CORE	HQ
Mara Ponta	Head of Unit	REM	HQ
Mathias Frese	Veterinary Surgeon	NAI	FD
Milena Osorio Montealegre	Mental Health and Psychosocial Programme Coordinator - Health	OPS	HQ
Mina Mojtahedi	Disability and Inclusion Manager	OPDIR	HQ
Mohamed Scheikh Ali	Ecosec CO	ECOSEC	FD
Mohini Ghai Kramer	Chief of Staff	DIR GEN	HQ
Nathalie Klein Kelly	Co prot	ILOT	FD
Pascal Porchet	Chief of Staff	OP DIR	HQ
Philippe Spoeri	HoD Syria	HOD	FD
Sophie Sutrich	Sexual Violence Adviser	OP DIR	HQ
Stephane Du Mortier	Health Learning and Development Manager - Health	OPS	HQ
Yves Daccord	Director General	DIR GEN	HQ
Zita Crener	Co Prot	MAN	FD
ICRC External			
Name	**Position**	**Organisation**	**Location**
Josse Gillijns	Head of PMER	IFRC	Geneva
Oivind Hetland	Learning and Evaluation Coordinator	NORCROSS	Norway
Donors			
Name	**Position**	**Organisation**	**Location**
GERMANY			
Corinna Holst	M&E team division of humanitarian aid	German Ministry of Foreign Affairs	Berlin
Dora Simon	Assistant officer for Red Cross and Red Crescent Movement	German Ministry of Foreign Affairs	Berlin

Peter Hermes	Part of department in charge of evaluation within ministry, evaluation management	German Ministry of Foreign Affairs	Berlin
USA			
Becky Kinsey	Evaluation	PRM (Bureau of Population, Refugees, and Migration)	Washington
Maria Rowan	M&E advisor	PRM (Bureau of Population, Refugees, and Migration)	Washington
NORWAY			
Lars Skari	Head of Strategy and Performance	NORCROSS	Norway

The interviews were conducted by all evaluators in the Assessment team.

FOCUS GROUP DISCUSSION GUIDE

The following questions were used during Focus Group Discussions. Questions were selected and adapted according to the position, experience, and understood knowledge of the groups in regard to ICRC evaluation practice.

A. Introduction
1.1. How do you as a team experience what is happening with evaluation? What does evaluation mean to you? How do you define evaluation?

B. Evaluation Drivers and Use
1.2. What is influencing or driving the decisions around evaluation? What aspects of ICRC culture and/or the external environment influences the way your sector thinks about or conducts evaluations? What is supportive and what are the obstacles?
1.3. What should evaluation look like?
1.4. How do you share information?
1.6. At what point in organisational calendar around planning and strategy, when would be best moment to get information and learning in that process?
1.7. Evaluation resources: What capacity do you have in your team for evaluation? What is working well, and what areas would you prioritize for improvement?

Focus Group Discussions included ICRC staff from a select number of departments, as well as one cross-departmental group. See the list of participants in Table 2 below.

Table 2: Focus Group Discussion Participants

Sector	Participants Attending (Name / Position)	Date	Joining Remotely
COORDINATION	Lilli Heinrichs – Head of Movement Resources and Data Katy Attfield – Head of Movement Operational Cooperation Jerome Giraud – RBM Advisor Vincenza Mancuso – Cooperation Coordinator Kabul Wangari Kiluva – Cooperation Coordinator Addis	26.06.2019	n n n y y
ECOSEC	Jules Kagwahabi Amoti, Acting Head of Unit Sarah Wilson, Ecosec Coordinator Amann Christa Utiger, Deputy Head of Ecosec Unit Christophe Driesse, Head of Middlea East Jean Pierre Nereyabagabo, Ecosec Coordinator Niger Mohamed Sheikh-Ali, Ecosec Coordinator Lybia	27.06.2019	n y n y n y
PROTECTION-DETENTION	Vincent Ballon, Head of Unit (detention) Jean-Philippe Dross, WatHab Coordinator Valérie Belchior-Belino Captier, Nutritionist (health) Ian Michael Leslie Macdonald, Protection Coordinator (US Delegation) Francesco Bruscoli, Detention Adviser	27.06.2019	n n n y n
CROSS DEPARTMENTAL	Fiona Terry, Head of the Centre for Operational Research & Expertise Jonas Johann Baumgartner, Deputy Head of Internal Audit Guive Rafatian, Policy & Humanitarian Diplomacy Operations Manager Rebeka Reka Johnson, Campaign and Marketing Manager Maria Thestrup, Head of the Global Compliance Office Antoine Ouellet-Drouin, Head of Sector, Planning, Monitoring and Evaluation	28.06.2019	n n n y n n

The FGDs were conducted by Cara Winters and Patricia Goldschmid.

The following provides a summary of the main points raised in the KI Interviews and the Focus Group Discussions. It informs the final report and recommendations of the assessment.

Evaluation/use/value/

Evaluation was seen by most interviewees and FGD participants as positive for the organization, providing insight into what worked, what could be improved, and an opportunity to learn and better steer the projects. Several respondents referred to the learning process as essential to the organization.

- *"More and more people accept the evaluation and we should push it. It's a way of learning. We are a learning organization so we should exploit this". (Internal interviewee)*

At the same time, most respondents also indicated having little experience in evaluation. Many respondents were familiar with the concept of evaluation, they lacked clarity on exact definitions and implementation processes. Confusion was noted with variations among monitoring, compliance, evaluation, audits, reviews, lessons learned. Respondents had varying expectations about what would or could be considered evaluation work, for example the scope of inquiry, topics covered, and timing.

There was consensus about the fact that there was less measurement of impact and generally a missing culture of evaluation. The culture was seen as more reactive than proactive in this sense. If evaluation was done, it was not done systematically with some exceptions.

- ➢ *Evaluation is not the practice in ICRC and people don't understand what needs to be done before. We should push for a change on mindset on that. Too many people feel that if they give something then they are doing good. For example, someone who wants to improve hygienic conditions in a prison will provide soap and think that's enough. But they don't follow up to see if the beneficiaries are actually using the soap. (Internal interviewee)*

The main reasons identified by many respondents for the lack of evaluation use included the perceived critical aspect of evaluation and that some may see evaluation as a means for judgement or blame rather than learning. Another reason was the fact there was a lack of capacity and resources for evaluation. Frequent changes in delegations was also identified as a possible reason for the lack of evaluation, as approaches to implementations vary and there is a lack of continuity between different delegates. High turnover was seen as provoking more short-term perspectives rather than long term.

- ➢ *"Evaluation is a space at ICRC where there is discomfort and confusion. There is confusion between output, outcome, and impact. Confusion between project and programme." (FGD respondent)*

➢ *"There are clear definitions around design and monitoring, but not evaluation. For example, at the closing of a programme evaluation would be valuable for learnings but it is not done systematically." (Internal interviewee)*

➢ *"We can be very good at quick reactive / emergency response, but we are very bad at sustaining efforts…We are too much into the narrative and not enough in the data…we don't have the data to back up what we say we do." (Internal interviewee)*

Approach/scope

Participants underlined the importance of adapting approaches to varying levels of complexity within the different areas of the organization and distinguishing between métiers. For example, areas such as protection or relief activities were noted as more difficult to evaluate than others. Many also mentioned that evaluation was more appropriate for programs that are sustainable in time and not in an emergency context. Challenges linked to monitoring information and data management were also mentioned by several respondents, particularly linked to gaps in reporting and accessibility.

➢ *"We are moving into multi-year programming and increasingly we are responding in protracted crisis. Evaluations are more and more relevant to this reality (vs. pure emergency response)." (Internal interviewee)*
➢ *"Evaluation has been a process led by ICRC global, project was not defined from beginning plus very general and global. That made it difficult. How can you evaluate without a baseline or concrete actions to see if what was done was relevant? Indicators are also global, very general and not specific to countries. We implement global ICRC process without specifying the context and what we want to do in each." (Internal interviewee)*
➢ *"This is the challenge of preventative action. How do we measure what we prevented from happening? We have these high aspirations and limited impact in the short term. It's difficult to lower objectives." (FGD respondent)*
➢ *"How do we measure if we prevented something? How do we measure impact on hard to access people or impact in situations where there is not much data or large data gaps?" (Internal interviewee)*

Evaluation was also noted both as challenging due in terms of transversality and scope, questioning how to work with partners and how to define the boundaries of the evaluation.

➢ *"Challenge for an evaluation is that there are many things going on but no connection point. What is missing is organised transversality" (FGD respondent)*
➢ *"How do we create a common understanding of what we're aiming for and what we're evaluating / what are our objectives? (Internal interviewee)*
➢ *Evaluation should be at a technical level for beneficiaries, and at community, regional and global levels. All levels are important in terms of accountability. (Internal interviewee)*

Diverging feedback was also on the link of evaluation to existing tools such as the PFR process. While some felt that this was the basis for a future evaluation structure, others felt that it was not developed enough to serve as a foundation.

➤ *To properly evaluate you need well defined objectives. Currently have some through PFR, its standardized but each country interprets it in different ways. One might be happy with an outcome in one country and not in another. PFR is for one year, the GO is for 5 years." (FGD respondent)*
➤ *The PFR process doesn't reflect well the amount of monitoring and evaluation. We write down how many hygiene kits we delivered. But the evaluation process is just as important and is not covered by the process. (Internal interviewee)*
➤ *"Evaluation should be part of the PMT, PFR. Would make sense that evaluation would be done at same level as PMT and PFR." (Internal interviewee)*
➤ *"It should be a part of Planning and Monitoring tool and should be a learning tool as well. Checking where we are, what to do next if the context changes or the partnership changes." (Internal interviewee)*

While there was consensus about the need for a strong but simple evaluation structure, with a clear framework and guidelines, there were diverging responses in terms of the drivers of evaluation, the location for the evaluation function, and whether evaluations should be managed and conducted by parties internal vs. external to the action being evaluated.

➤ *"To restart with evaluation, it would have to restart small and simple, without complexity. It needs to be small and achievable. You have to rebuild the trust of the organization, also in terms of means." (FGD respondent)*
➤ *"The ICRC being a large and complex agency, it needs a simple evaluation framework." (Internal interviewee)*
➤ *"The ICRC would need a standard frame for developing ToRs and guidelines on how to do it as there is currently little knowledge on this. Could have an in-house expert with a network of evaluation experts and a budget to provide advice on how to manage an evaluation. (Internal interviewee)*

Some respondents felt that the evaluation should be driven by a function located at HQ either within DirGen or as an independent team reporting to the Assembly. Others emphasized a need to position evaluation drivers and control at the delegation level.

Some respondents felt that an evaluation should be conducted by an internal staff member who is not involved in the evaluated action or by a former delegate who understands the complexities within the organization. Others felt that a fully external perspective was essential to an objective result. There were diverging responses about the required experience of the evaluator. While some emphasized a need for the consultant to have expertise in the topic evaluated (e.g. an expert in protection), others felt that expertise in evaluation was more pertinent.

➤ *"Might have a good consultant who has experience in the humanitarian world but does not speak ICRC language." (FGD respondent)*
➤ *"Evaluations work better when it's from someone within the movement. External consultants don't always grasp the dynamics of the movement." (Internal interviewee)*
➤ *"External is always better. We want to step out and someone can come in with a different perspective." (Internal interviewee)*

A further perspective mentioned was the creation of a pool of consultants, both internal and external that could be called upon depending on the nature of the evaluation.

Divergence was also noted on who would initiate evaluation in terms of Red Line or Blue Line. Likewise, while many felt that evaluation should be centrally structured, some felt it should be initiated in the field while others felt that it should be driven by HQ.

Information management and sharing was considered a potential obstacle to improved evaluation practice by those who felt that results and other information is not shared sufficiently, particularly between HQ and the field. Others felt that this would not be an impediment for the development of a structured evaluation approach.

➢ *"There are many reports, but they are not shared and they are not accessible to the field. It is often said that they are on the "team space" but they aren't found by field. Evaluation should go beyond what we currently have which at best are reviews or monitoring." (FGD respondent)*

Some respondents identified the lack of visibility or accessibility of documents within the organization as a challenge, linked to the magnitude of documents produced.

There should be clear definitions and guidelines for evaluation/ ToR Development / establishing 'ways of working' that are common on similar exercises. Respondents highlighted interest in improving information harmonization in terms of the process around how it is produced, as well as how information is brought together and used at different levels of the agency. For example, looking at trends and accountability frameworks, how information is positioned to drive critical reflection, and establishing a process for recommendations tracking.

➢ *"How we learn as an organization needs to be strengthened. Learning from evaluations needs to be addressed. Even when evaluations are done – they are parked. There is not enough discussion happening around the learning. We should look into how we can integrate learning into the strategy process and how this informs what happens next." (Internal interviewee)*
➢ *"Evaluation is the best answer to questions that we cannot always answer linked to sustainability of programmes, environmental impact. For us evaluation is best tool to answer." (Internal interviewee)*

There were diverging opinions about the timing of evaluations, whether it should be on an annual basis, or adapted to each specific unit or delegation. Defining correct timing was a point of divergence with some respondents noting it as complicated since projects often differ in terms of length / duration. While some areas of ICRC's work operate on a shorter-term basis, other areas of work may be ongoing for years. In some cases, staff referred to the same point, but did not see it as a challenge.

- *"We should evaluate at the end of programmes, for example after a 12-month period or at the end of a crisis situation such as war." (Internal interviewee)*
- *"PfR is done on a yearly basis, which is too short. It's limited for long-term or multi-year programming. Protection is not structured as a short-term project." (FGD respondent)*
- *"An evaluation should look at the results and outcomes of what we have done not necessarily immediately at end of activity or program but a few years after completion of activity. What is done immediately after would be more a review." (Internal interviewee)*
- *There are pure emergency response actions but also structural programs planned over two years and implemented within a year that would or should have an impact over decades. For a project like that, an evaluation would make sense during the implementation as well as after the delivery. We should go back to the authorities to see the real impact we have over a certain number of years. Currently, there is speculation about impact, but it is not measured scientifically." (Internal interviewee)*

Staff availability / human resources

Staff availability and resources were mentioned by a majority of respondents as one of the major challenges in evaluation. Consensus was that in an operational context, people were overloaded with work and that evaluation should not take time away from ICRC's core mandate.

Concerns about available human resources and the time that would be required for evaluation were noted by internal interviewees mostly from the field. This was linked to the importance of having a clear and simple structure mentioned above and perceived facilitated approach for field delegations.

- *"If we could have a rapid evaluation response – a quick non-bureaucratic service that we could activate. Anything that would not burden the field too much but is of use." (internal interviewee)*
- *"Part of the struggle will be that it is not in our culture to do evaluation. Also, the organization has transformed massively over the past decade. Staff are overwhelmed." (internal interviewee)*

External influence

There was consensus on the importance of addressing donor expectations, while also maintaining integrity and uniqueness of organization.

- *"There are different donor perspectives and demands on for assistance programming vs. protection programming.– what ICRC does in assistance looks similar to what other agencies do, and the donors don't understand why ICRC cannot provide the same level of reporting and evidence on these programs as they are accustomed to receiving from their other funding partners." (FGD respondent)*
- *"ICRC is trying to balance its unique position and privilege toward donors in terms of reporting and the pressure of some donors to provide increasingly more information, but the ICRC system is not built for that. ICRC wants to keep it's unique position, but it does not have the systems to respond to or reach donor demands." (FGD respondent)*

Annex IV: As Is / To Be Workshop: *Summary report and conclusions*

This annex reports on the As Is / To Be Workshop held in Geneva as part of the methodology of the ICRC Evaluation Function Assessment. It documents the content, participation, and the outputs of the workshop and provides a set of conclusions informing the final report and recommendations of the Assessment.

WORKSHOP CONTENT

The workshop was held on July 1st, 2019, at the ICRC Geneva Headquarters. Participants reviewed and validated five key initial assessment findings on the 'as-is' state of evaluation practice at ICRC and formulated a 'to be' vision for the ICRC evaluation function and future evaluation practice. Objective (and objective-level) statements were created by workshop participants to correspond with each of the 'as-is' findings.
The format of the workshop was participatory and cross-sectional. See the agenda in Table 1 below.

Table 1: Workshop Agenda

Time	Activity	Method	Facilitator
2:00-2:15pm	Introduction of the Assessment, Assessment Team, and Workshop Agenda	Presentation	Cara
2:15-2:30pm	Brief presentation of the initial 'as-is' assessment findings.	Presentation	Cara
2:30-3:00pm	'World Café' review, discussion, and input on each 'as-is' findings	Rotating group work	Glenn (room) Patricia (online)
3:00-3:15pm	BREAK		
3:15-3:30pm	Recap main inputs to the 'as-is' findings	Plenary	Glenn
3:30-4:30pm	Formulate 'to-be' objectives and objective statements (per 'as-is' finding).	Small team group work	Cara (room) Patricia (online)
4:30-4:45pm	Cross-team review and discussion on formulated 'to-be' objectives	Small team gallery walk	Cara (room) Patricia (online)
4:45-5:00pm	Recap main inputs on the 'to-be' objectives and close workshop	Plenary	Cara

Participants in the workshop represented a cross-sectional set of experiences, concerns, and ambition on the subject of evaluation at ICRC. All staff are based in Geneva, with varying length of tenure at HQ versus time working in field delegations. See the list of participants in Table 2 below.

Table 2: Workshop Participants

Name	Position	Joining Remotely
Karen Cecilie Rogenaes-Panxha	Adviser, Performance and Accountability, REM	Yes
Antje Van Roeden	Head of Project, Health	No
Evaristo De Pinho Oliveira	Head of Unit, Water and Habitat	No
Guive Rafatian	Operations Manager, Policy & Humanitarian Diplomacy	No
Siobhan Foran	Manager, Operations Diversity Inclusion	No
Vincent Ballon	Head of Unit, Detention	No
Antoine Ouellet-Drouin	Head of Sector, Planning, Monitoring, Evaluation	No
Barbara Hunziker	Organizational Manager, PROT/ASSIST	No
Karla Levy Simancas De Marichales	Manager, Institutional Strategy & Performance	No
Julia Afton Eppts	Project Officer, Planning, Monitoring, Evaluation	No

The workshop was facilitated by consultants conducting the Assessment:
- Cara Winters
- Glenn O'Neil
- Patricia Goldschmid

Responsibilities of each are listed in the workshop agenda under Table 1.

The following table summarizes the outputs of the workshop, including the initial assessment findings, additions or comments from participants about the findings, and recommendations for the 'to be' vision.

Table 2: Workshop Output

'As Is' Finding	Participant Input on 'As Is'	'To Be' Objective	Participant Input on 'To Be' Objective
Finding 1: Performance *Limited evaluation visibility across all levels of the organization and within the PME Sector.* *Progress (in sections) on evaluation publication and storage in team workspaces.*	<u>Agreement / Consensus</u> <u>Input:</u> 'Free for all' culture when it comes to information, communications, systems. We don't see how evaluations are used (even if we do see that an evaluation occurred). The 'E' does not exist in the PME Sector. This is why it lacks evaluation visibility. No incentive for sharing evaluations or evaluation learning. Lack of 'learning' focused evaluation – more focus on accountability. It drives 'who should be informed'.	1. Evaluations are visible, accessible, promoted, and used adequately for both internal and external purposes. 2. Clear evaluation process, roles & responsibilities, and quality criteria. 3. Evaluation function has capacity to manage / leverage evaluations to respond to ICRC information needs.	Challenges: ➢ It is difficult to identify when an evaluation is coming up or has been planned. Unclear how the evaluation function can address this, and whether or when it should engage. ➢ Leadership and ownership on evaluations is weak. Tactics / Solutions: ➢ Smarter use of information technology / create evaluation database ➢ Visibility does not guarantee use. Evaluations must evolve beyond the report – e.g. management response, action on recommendations, integration of evaluation feedback into ways of working. Evaluation system should address these areas of 'use' specifically. ➢ Establish a communications plan ➢ Ensure a strong change management process for IT solutions and overall absorption capacity of increased evaluation visibility ➢ Increased visibility may be possible through the planning and budgeting process of evaluations, e.g. securing a budget code for evaluation and running queries on the use of that code. ➢ Evaluation framework must articulate what is expected on information flow (evaluation visibility and use) between delegations, métier, central evaluation function, the wider institution.

| Finding 2: Performance

*Limited **identification of potential evaluations** on transversal issues, country-level 'interdisciplinary' programming, or regional or global métier-specific thematic topics.* | Agreement / Consensus

Input:
Limited to no awareness of what evaluation can offer. Lack of capacity. People question the cost of evaluation (time, budget) vs value added.

Aversion due to limited capacity /time.

Evaluation is not included in the initial program / policy / operational formulation. Very limited identification of evaluation during planning.

No annual evaluation plan for the institution and no quarterly updates.

Organizational systems and structure are siloed (including budgets).

Evaluation team are not multisectoral teams (often, not always). | 1. There is a defined set of potential evaluations, including transversal / inter-disciplinary / regional, etc. This is supported by:
➢ Vision and awareness
➢ Resources
➢ Tools and methods | Challenges:
➢ ICRC's independence vs. becoming donor driven
➢ Balancing demands on staff, e.g. time spent on administration vs. time spent with beneficiaries.
➢ How much of the issue is about structure vs. about ways of working and mindset? The reality is that there is there is little appetite for transversal evaluation (or transversal work). This is the issue.
➢ Territoriality
➢ No vision, lack of asking 'what should we be doing?'
➢ Evaluation is not promoted or encouraged at the management level.

Tactics / Solutions:
➢ Manage territoriality through leadership and eval budget structure/availability
➢ Invest time on internal pedagogy, develop staff knowledge and experience of evaluation
➢ Building vision and awareness requires that ICRC articulates WHY we evaluation, and HOW we evaluate.
➢ Look at the success stories of the (previous) transversal set up for evaluation support (e.g. GENEVAL).
➢ M&E should part of the program / project from the beginning – bring it into planning
➢ Common objectives between métiers |
| Finding 3: Capacity
Ad hoc procedures for evaluation quality assurance during the past three years. Organization guidelines for evaluation practice exist (2006), but not widely known. Métier-level guidance exists in varying degrees of detail and applied practice. | Broadly agree, points of disagreement.

Agreement:
There hasn't been a coordinating body or 'authority' on evaluation in years (since 2011)

Focus is on procedure compliance, not learning.

Linked to issues of convenience and 'lack of integrity' in evaluation process. | 1. ICRC has a standard evaluation framework with standard criteria that can be adapted to different evaluation situations (scope, scale, specificity) at HQ and Field Delegation levels.

2. Baseline is included from the start into program / project description. | Challenges:
➢ Resource allocation / resource availability
➢ Initiative fatigue – evaluation / baseline requirement seen as 'yet another procedure'
➢ Limited capacity to collect relevant data for baselines.
➢ Limited internal expertise and culture of M&E
➢ If given the option between a 'review' and an 'evaluation', staff may opt to do a 'review' as a way to avoid the requirements of 'evaluation' even if the appropriate approach is an 'evaluation'.
➢ Lack of systems thinking in the agency. |

		3. Timing of evaluations is defined on a multi-year plan (i.e. fixed milestones)	
	Disagreement: 'Ad hoc' can also mean more freedom. Different implications. The situation is nuanced between the métiers (not all the same). Ad hoc should not be considered synonymous with a lack of capacity.		Tactics / Solutions: ➢ Consider phasing of the evaluation framework, e.g. address the 'second level' in a second phase, after consolidating the centralized function. ➢ Create a multi-year evaluation plan ➢ Clarify definitions between 'review' and 'evaluation' ➢ Consider qualitative baselines, using approaches common in needs assessment methodology (which staff may be more familiar with) ➢ Limit 'review' requirements – if there is less pressure to do 'review', then staff may find more value in 'evaluations' ➢ Criteria and methodologies for evaluation quality assurance need to be adapted to evaluation type.
Finding 4: Capacity *Limited expertise in evaluation methodology within the organization. Inconsistently present in units driving decisions to evaluate and setting evaluation scope and design.* *Staff are unfamiliar with the evaluation methodologies able to respond to challenging evaluation contexts. 'Impact' is not clearly defined, creating confusion.*	<u>Agreement / Consensus</u> Input: Need a consistent approach to capacity requirements at a centralized vs. decentralized level. Unclear roles and responsibilities. Unclear whether ICRC should hire specific evaluation capacity, vs. all staff should be capacitated to evaluate, external competencies. Limited expertise does not equal a lack of value or promotion of evaluation. Staff may be interested in being trained or positioned to work on evaluations.	1. People have access to the necessary expertise on evaluation. 2. Clear scoping of evaluation responsibilities at all levels of the organization; and the capacity to fulfil demand for technical expertise at all levels.	Challenges: ➢ This is a Movement-wide issue. ➢ Today, more effort and attention is placed on compliance around procedures and transactions. Tactics / Solutions: ➢ Options include (1) repository of information or people (e.g. helpdesk, teamspace, etc.), (2) support for accessing existing competencies, (3) train existing staff / managers on evaluation methods OR create dedicated M&E capacity. ➢ ICRC needs to map existing expertise in M&E. We should also not forget local capacities in the field delegations. ➢ Integrate evaluation into existing departmental tools and trainings. ➢ Clear mechanism for quality assurance should be created to support / backstop areas where capacity is weak. ➢ Decision to evaluate to should be at different levels, but the process supported.

Finding 5: Culture	Broadly agree, points of disagreement.	1. ICRC staff appreciate the purpose and intended use of evaluations / the evaluation function, and willingly engage in the process.	Challenges:
*Staff identify **'integrity' as a primary concern** in evaluation. Evaluation integrity includes: why, how framed, acceptance of the evaluator, degree of collaboration and transparency, articulated approach to 'independence.'*	**Agreement:** Look at how evaluation was commissioned initially (e.g. under GENEVAL) – there was proper governance, integrity of process and the evaluators. Now there is suspicion around the reasons for evaluation. Bias towards 'internal reviews'. Staff unsure or don't know how to engage on impact evaluation. Yes and no. People feel threatened. **Disagreement:** Fear of accountability, vs. learning (learning was the originally foreseen objective under GENEVAL) Tendency to 'overkill' methods for credibility. ICRC principles are my main concern. Evaluation seen as 'critic of the metier' and not an opportunity to learn. A clear framework would bring back a zone of comfort.		➢ Lack of consistent support of leadership ➢ Very big pockets of lingering resistance (old school thinking) on HOW this can / should be done. (This may vary between units) Tactics /Solutions: ➢ People enjoy being in their comfort zone. Evaluation function must look for ways to bring people from the unknown to the known. ➢ Demonstrate that evaluations are possible, and that this is not new to ICRC or the humanitarian sector.

During the feedback and plenary sessions, participants identified the following **questions that need to be addressed as part of achieving the 'to be' vision**:
➢ Should the evaluation function stop at providing materials and guidance, or does it accompany and support evaluations at all levels of the agency?
➢ What is the regional delegation role? Will it also support the process and accompany? How does that relate to a central evaluation function?
➢ How to balance evaluation requirements, e.g. achieving evaluation quality without overwhelming staff or detracting from time with beneficiaries?

Conclusions

- **Review and Validation:** Participants broadly validated the initial set of findings. Places of disagreement included the negative implication or possibility for generalized application of the term 'ad hoc' under Finding 3 and points of nuance on organizational culture or re-prioritization of the issues captured under Finding 5. In particular, some participants stressed that 'integrity' concerns are another way of saying that the motivation and reasons for evaluation are not trusted, linking to their overall concern about evaluation as a politicized source of critique (vs. a tool for learning and improvement).

- **'As Is' Input:** Participants provided further detail on the drivers or causes of the issues raised per finding. Consistent messages on what is driving the situation included: degree of formalization or systematization of evaluations and evaluation responsibilities, staff capacity and availability, organizational focus and prioritization (e.g. competing pressures for staff attention), and organizational culture (e.g. leadership, incentives, attitudes).

- **'To Be' Objectives:** Objectives formulated by participants cover the following areas: Evaluation Framework, Evaluation Capacity and Responsibilities, Evaluation Planning / Timing / Connectivity with the Project Cycle, Evaluation Promotion / Use / Appreciation. <u>This suggests that the issues raised in the 'as is' findings and the participant input on those findings can be addressed through defining, resourcing, and systematizing these key areas</u>.

- **'To Be' Challenges and Solutions:** Input provided on the 'to be' objectives can be categorized as potential challenges to achieving that objective and solutions to those challenges or tactics towards reaching the objective. Input provided during the workshop on this section is consistent with the findings generated through the KIIs and the Focus Group Discussions.

Annex VI: Evaluation Function Purpose and Location: *Considerations and recommendations*

INTRODUCTION

This annex details recommended ways ICRC can define the purpose of the evaluation function and structure its location. It responds to Objectives I and II[38] of the ICRC Evaluation Function Assessment ('the Assessment') and elaborates on Recommendation 2, covered in the main report (Section 4)

Recommendations: *Options Tables*

The 'options tables' presented in this recommendations report cover the choices that ICRC can take for structuring the evaluation function and their benefits or drawbacks as linked to the recommendations covered in the main report. They address the following:

- Evaluation Function Purpose and Scope
- Evaluation Function Placement

The points above are influenced by the relationship of the evaluation function with planning and strategy, results monitoring, and compliance and risk management functions or ambitions at ICRC. This issue is addressed at the end of the brief, following the recommendation tables.

A. **Purpose and Scope:** What would an evaluation function do?

Models followed for the evaluation function in humanitarian agencies, and in particular its role in conducting evaluations, vary according to priorities on resource availability for the function, quality control in evaluation, accountability vs. learning oriented evaluation systems, and how an agency defines and assures evaluation independence. Models also vary according to the amount of time the function dedicates to activities beyond evaluation generation, e.g. developing and institutionalizing the system, integration of evaluation learning into organizational decision making, planning, and policies, staff capacity building, and coordination and alignment with other accountability and learning functions in the organization. These elements are influenced by how an agency is funded, its overall scope of work (e.g. is it programmatic, normative, both?), and the degree to which its technical expertise is defined and structured (e.g. technical singularity vs. multi-disciplinary).

Based on the findings of the Assessment, **it is recommended that ICRC adopt a 'hybrid' approach to the evaluation function and structure,** wherein the function combines elements of the service-oriented 'centre of excellence' model with more accountability-oriented 'command and control' models. This approach fosters a more active evaluation culture and emphasizes the use of evaluation for learning and improvement, while ensuring a degree of evaluation independence and structure for the accountability aims of evaluation. This model also retains space for areas of the organization that are invested in evaluation, building on their foundation of experience, encouraging them to continue their work within a common set of objectives, and promoting them as ambassadors for evaluation practice vis-à-vis units who are new to evaluation.

---

[38]'Define the business model of the evaluation function' – Assessment Objective I and 'Define organisational structure...necessary to support the future business operations of the evaluation function' – Assessment Objective II

Table 1: Evaluation Function Purpose and Scope

Options	Description	Pros / Cons
"Command / Control" Model	Evaluation function itself conducts evaluations and produces evaluation reports as its core purpose. Little to no outside consultants, academic partners, or organizational staff are used in the evaluation teams. Function also carries out a diverse range of evaluation-related tasks to support institutionalization of the function, including a focus on evaluation use to support organizational decision making. The function has full control and is entirely accountable for all evaluation reports produced by the organization. Evaluations are centralized through the function, which has an independent budget and authority over the organization's evaluation plan. This scope implies a fully centralized evaluation framework. Requires a substantial level of resources for staffing and hiring evaluation professionals. Examples include: WFP, World Bank	Pros: Works for a centralized, accountability driven evaluation framework in a top-down institution. Reflects MOPAN criteria and UNEG Norms and Standards. Strong degree of quality control and assurance of evaluation independence. Cons: Organizations with this model tend to emphasize evaluation as a function of organizational accountability – to the detriment of organizational learning. A function that is skewed towards accountability (and away from learning) can limit the value and sustainability of the function. The function must be highly resourced with technical and managerial staff to ensure its success, or its perceived (or actual) value will be challenged. Functions may struggle to integrate evaluation into organizational decision making, policy, and strategy. Loss of flexibility to use evaluation as a tool to respond to complex or rapidly changing issues in the field.
"Center of Excellence" Model	Primary role is centralized oversight and development of evaluation policy, guidelines, and system to ensure quality and uptake of learning. The function has full control over organizational level evaluation policy and guidelines. These guidelines may be adapted or further detailed by the programmatic or areas of work in the agency. Function has a budget for its activities and steers the process through which the centralized evaluation plans are developed and approved. Decentralized evaluation plans are collated or mapped by the function, but not led or approved. Function staff lead the mechanism for ensuring adherence to quality standards and evaluation utility when working with consultants, academic partners, and potentially staff rosters who conduct the evaluations. The function often acts as evaluation manager for centralized evaluations or is on a steering committee for centralized evaluations managed by other teams. It may provide 'help-desk' guidance on decentralized evaluations, as requested by the commissioning teams. Decentralized evaluations are discretionary and demand-led by the field or technical departments. This scope implies a decentralized or split decentralized / centralized evaluation framework. Examples include: IFRC, NORCROSS, MSF	Pros: Works for decentralized or split decentralized / centralized, learning driven evaluation framework in organizations with a culture of diffuse authority or bottom-up decision influencing. Can work to ensure that evaluations reflect a range of industry standards, even if the function itself diverges from meeting all UNEG or MOPAN criteria. Strong degree of flexibility. Cons: Organizations with this model tend to emphasize evaluation as primarily a function of organizational learning – to the detriment of strict quality control. The evaluation function must be adequately accepted by staff or supported by senior leadership (through enforcement of policies and function responsibilities) to exert its influence OR have a mechanism for oversight (eg approvals process for initiating evaluations). Degree of devolved decision making on evaluation plans may lead to fragmented coverage of evaluations (geography, topics, scope) – this can compromise the potential for evaluation to provide a comprehensive evidence base for corporate strategy or agency-level policy.

| Hybrid

Recommended | Role mirrors the description for the 'center of excellence' model, *except it also has the following control functions:*
• Function can directly commission evaluations that it centrally manages
• Function has full quality control over centralized evaluations and provides quality assurance for decentralized evaluations
• Function has capacity to join evaluation team or develop methodology for central or decentralized evaluation.
• Function defined and ensures process for evaluation planning at central and decentralized levels.
• Function holds responsibility for how the organization tracks action on evaluation management responses
• Function leads the process of embedding evaluation in key organizational systems that support their use

Scope implies a centralized / decentralized evaluation framework.

Examples: UNHCR, ILO | <u>Pros:</u> Works for decentralized / centralized, learning and accountability balanced evaluation framework. Reflect a range of industry standards for evaluation functions. Mirrors approach taken by larger organizations with a record of strong evaluation practice. Maintains flexibility, while promoting a degree of quality control and assurance of evaluation independence.

<u>Cons:</u> Difficult to implement or enforce in organizations with a culture of diffuse authority or bottom-up decision influencing like ICRC. |

B. **Placement:** Where is the evaluation function located?

Placement for the evaluation function can be considered according to the following parameters:
- Location of function responsibility (e.g. with the governing body of the organization or with the highest executive level)
- Degree of 'co-location' of the function (e.g. is it entirely 'stand-alone' or 'co-located' with the agency oversight function like internal audit or with executive management functions like planning and strategy and/or monitoring and research and/or with policy units).

The table below focuses **on the first point**, where does responsibility sit for the function and – by extension – where does the function itself sit. The choice of location for the function is influenced by several factors: perceptions around evaluation independence, efficiency gains, beliefs about how to secure the function's overall relevance to and impact on organizational culture, and utility of the function. The issue of co-location is not addressed in detail in the Assessment. It is, however, commented upon at the end of this report in relation to ongoing discussions at ICRC about how to best situate M&E capacity in the agency as a whole. It also touches on possible affiliation with the Internal Audit office.

The table below outlines a set of options on where to best locate the evaluation function, following the most relevant models for ICRC. Based on the findings of the Assessment, **it is recommended that ICRC adopt a 'split responsibility vs. location' approach**, wherein the function connects with the primary users of evaluation through its location in the Executive Office, but retains independence and authority over the evaluation agenda through a defined set of approvals that sit with the governing body of the organization (in this case, the Assembly). This model promotes the organizational learning and technical support aspects of the function, of particular importance for organizations that adopt split centralized and decentralized evaluation frameworks.

Table 2: Evaluation Function Placement

Options	Description	Pros / Cons
Fully within the Governing Body / Assembly	Responsibility for the evaluation function sits with the Assembly, including the role of appointing the Head of Sector/Unit and approving the plans and budgets of the function. Location of the function is situated under the direct line management of the Assembly for all activities and performance management (HR). There is a clear process for presenting / sharing information from evaluation learning / findings with the Assembly. Function may be a fully stand-alone office or co-located with Internal Audit. Examples: WFP, UNDP	Pros: Best reflects MOPAN and UNEG standards. Signals a strong focus on evaluation independence and, potentially, authority of the function. Enlarges the space for autonomy of function decision making (including the size and budget of the function). Positions the function to provide the Assembly with independent assessments of activities or strategies (e.g. a distinct voice and direct channel of information) Cons: May compromise utilization-focused evaluation, or evaluation with a primary purpose towards organizational learning. Can create an over-emphasis on evaluation for accountability. Locating the function at the same level as Internal Audit may weaken the identity of the function, as seen in UN agencies following this model[39]. Reduced visibility and potential loss of influence of the function if seen as disconnected from agency operations or work.
Fully within the Executive / Director General's Office	Responsibility for the evaluation function sits with the Director General (DG), including appointing the Head of Sector/Unit and approving the plans and budgets of the function. Location of the function is situated under the direct line management of the DG for all activities and performance management (HR). There is a clear process for presenting / sharing information from evaluation learning / findings with the DG. Function may be a fully stand-alone office or co-located with other executive management functions, such as strategy, planning, monitoring and reporting. Examples: IFRC, UNHCR	Pros: Signals a focus on evaluation independence from operations and, potentially, a preference towards shaping evaluation for learning and improvement (vs. accountability driven). Improved coordination and alignment of evaluation with other organizational functions focused on learning and evidenced-based decision making. Cons: Potential for the function to be directed more towards supporting corporate-level strategy, oversight role, and decision making to the detriment of the function's possible role in supporting decentralized evaluations / supporting the entire house. Risk of the evaluation function being compromised by the agenda of the DG and not seen as fully independent. Potential limitations on the size and budget of the function.
Split Responsibility vs. Location *Recommended*	The function is located in the DG's office. Areas of responsibility for the evaluation function that sit with the Assembly or with the DG following approval from the Assembly include: • Hiring Head of Sector / Unit for the function • Validation of evaluation plan and selection criteria • Approval of function budget / staffing structure • Receipt and approval of key function products, such as an annual report or review on the learning or main findings / trends highlighted by evaluations	Pros: Aligned with MOPAN and UNEG standards for independence. Signals a focus on evaluation independence and, potentially, authority of the function, balanced with a placement that encourages evaluation as a tool for learning and improvement. Enlarges the space for both autonomy of function decision making and alignment of evaluation with other organizational functions focused on learning and evidenced-based decision making. Cons: Split management is difficult under any circumstances – and in particular for potentially politicized functions like evaluation. Risk of the evaluation function being compromised by the agenda of the DG is not entirely eliminated. Assembly may also have undue influence over the evaluation agenda.

---

[39] The JIU 'Analysis of the Evaluation Function in the UN System' (2014)

	All other activities and routine performance management (HR) of the function sit under the line management of the DG. There is a process for presenting / sharing information from evaluation learning / findings directly to the Assembly (without DG involvement or interference) and to the DG. Function may be a fully stand-alone office or co-located with other executive management functions, such as strategy, planning, monitoring and reporting. <u>Examples:</u> UNICEF(ish)	

The option tables provided above are influenced by decisions the organization will take regarding co-location of the evaluation function with planning and strategy, results monitoring, and compliance and risk management functions at ICRC.

The issue of whether an evaluation function is co-located with other learning, accountability, or strategy functions in an agency is largely a decision of resource efficiency, potential visibility of the function, and desired areas of collaboration[40].

In the case of ICRC, the following should be noted:

- **Co-location with Internal Audit:** Evaluation functions that are co-located with oversight functions can lag in development[41]. For this structure to be successful, agency leadership and the governing bodies need to keenly understand the particular added value of an evaluation function relative to other oversight functions. Without this understanding, the evaluation function can become confused with general oversight and lose its identity in the agency, thus limiting its ability to develop and champion evaluation culture and organizational learning. Agencies may also reduce investment in the function under these models (time, staff, budgets), feeling that evaluation does not provide a unique service. *It is not advised that ICRC co-locate evaluation with Internal Audit.*

- **Co-location with Monitoring (Results):** There are several benefits to co-locating evaluation with monitoring (results) at the strategic level. First is ensuring a coherent use of evaluation learning together with corporate-level analysis of monitoring reporting. This 'marriage' is possible in forums such as agency strategy or planning sessions / process, quarterly presentations to senior leadership or governing bodies, annual learning reports, etc. Second is ensuring a coherent approach to

---

[40] The JIU 'Analysis of the Evaluation Function in the UN System' (2014) indicates that embedding evaluation information and learning in the process for organizational strategy and planning, a set calendar of presentation of evaluation findings to the senior leadership / governing bodies, and including evaluation learning in the annual report (together with other oversight or learning information) are the most significant factors in ensuring the development of the evaluation function. This holds under any structural model, indicating that an intentional and resourced plan for evaluation use can be successful, regardless of where the function is co-located.

[41] The JIU 'Analysis of the Evaluation Function in the UN System' (2014)

evaluation based on the institutional approach to monitoring, as well as the unique issues around monitoring for the areas of agency work. The strongest evaluations are linked to solid output and outcome monitoring practice, or where that is weak, are positioned to help strengthen monitoring in the future. As a challenge, co-locating evaluation with results monitoring may create a situation where function staff time and capacity are diverted towards supporting outcome monitoring – and away from evaluation - in cases where monitoring remains under-resourced in the agency.

- **Co-location with Planning and/or Strategy**: Co-locating evaluation with planning and/or strategy can better orient evaluation towards organizational learning and ensure that evaluations are consistently used / referenced in developing organizational strategy or plans. Situating evaluation with strategy in a structural set up can, however, increase the potential for the function to be directed towards supporting corporate-level strategy to the detriment of the function's role in supporting decentralized evaluations / supporting the entire house.

The current structure of the PME Sector includes monitoring, evaluation, and planning. ICRC should consider either keeping this model in the development of the evaluation function, or co-locating evaluation with either monitoring or planning / strategy if the PME Sector is split.

Annex VII: ICRC Evaluation Function Resourcing: *Considerations and recommendations*

This annex details recommended ways ICRC can resource the evaluation function in terms of staffing capacity and funding structure. It responds to Objective II[42] of the ICRC Evaluation Function Assessment ('the Assessment') and elaborates on Recommendation 3, covered in the main report (Section 4)

Recommendations: *Options Tables*

The 'options tables' presented in this recommendations report cover the choices that ICRC can take for staffing and funding the evaluation function and their benefits or drawbacks as linked to the recommendations covered in the main report. They address the following:
- Evaluation Function Staffing: The central office of the function and across the organization
- Evaluation Funding: Budget Ownership
- Evaluation Funding: Budget Source
- Evaluation Funding: Budget Amount

C. **Staffing:** Who works in the function? What do they do?

Staffing the evaluation function is a direct factor of the purpose of the function and may be influenced by the location of the function (see Recommendation 2 and Annex VI). It can be considered along two parameters:

- Staffing the central office of the function
- Staff capacity required at different levels of evaluation activity

Staffing options are best reviewed in relation to the models for the evaluation function purpose and scope presented under Annex VI. It will also be influenced by the decision taken in regards to evaluation funding.

Based on the findings of the Assessment, **it is recommended that ICRC adopt the staffing structure matched to a 'hybrid' model for the evaluation function and structure,** wherein the number of staff and their specific qualifications enable the technical capacity required for evaluation leadership and support on evaluation practice across several levels of the organization. Staffing must also be sufficient to permit direct management of centralized evaluations, as required by the evaluation plan. (Note: Under the 'hybrid' model, the evaluation function it is not obligated to directly manage every centralized evaluation. It is, however, obligated to serve in a steering capacity and quality assurance capacity for every centralized evaluation, while also providing technical guidance and quality assurance for decentralized evaluations). Staffing should address important gaps in evaluation knowledge and practice as required by ICRC programming and operational structure.

Due to the lack of consistent structural parameters for the regional level of ICRC operations (e.g. the Regional Model such as NAME vs. Regional Delegations, placement of regional red line vs. blue line functions, etc.), the delegation-level recommendations for evaluation staffing focus on the need for field capacity without specifying where this capacity should be located. When deciding where to place field-level staff capacity, ICRC should take the different regional approaches into account.

---

[42] 'Define the organizational structure, staffing levels, roles, and skills requirements necessary to support the future business operations of the evaluation function' – Assessment Objective II

Table 1: Staffing the central office of the function

Function Model	Description
"Command / Control" Model	Head of Sector / Unit with significant experience in humanitarian evaluation systems and team or agenda management. At least 2 staff with experience in organizational evaluation systems to support: evaluation planning, evaluation quality control procedures, and evaluation use. Aspects of these will require agency-level guidelines, while others will require only unit/sector-level procedures and ToRs. One of these positions dedicates time towards organizational capacity building and change management (years 1&2) on organizational evaluation use and how the agency engages with the function. At least 3 to 4 additional staff as evaluation experts at a high position level, with differential skills in quantitative vs. qualitative methods. They should have experience in evaluation approaches tailored to the different areas of ICRC work (e.g. communication, law and policy, protection, water and habitation, health, food security and livelihoods, etc.). These staff conduct evaluations. Total number of staff: 6 to 7 in the central office
"Center of Excellence" Model	Head of Sector / Unit with significant experience in 'center of excellence' models in humanitarian agencies and a track record of successful team or agenda management. Ability to network and persuade is critical. Demonstrated background in humanitarian M&E. At least 3 staff with experience in organizational evaluation systems to develop agency-level guidelines and support: evaluation planning, evaluation management, evaluation quality control procedures, and evaluation use. Two of these positions dedicate time towards organizational capacity building and change management (years 1&2) on each of the listed areas, ICRC evaluation guidelines, and how the rest of the agency engages with the function. Total number of staff: 4 in the central office
Hybrid *Recommended*	Head of Sector / Unit with significant experience in 'center of excellence' models in humanitarian agencies, significant experience in humanitarian evaluation systems, and a track record of successful team or agenda management. Ability to network and persuade is critical. At least 3 staff with experience in organizational evaluation systems to develop agency-level guidelines and support: evaluation planning, evaluation management, evaluation quality control procedures, and evaluation use. Staff will also need to adapt guidelines to evaluations managed by the function itself. Two of these positions dedicate time towards organizational capacity building and change management (years 1&2) on each of the listed areas, ICRC evaluation guidelines, and how the rest of the agency engages with the function. Recommended that one or two of these positions have evaluation expertise in qualitative methods. This assumes the evaluation function is staffed with the two 'impact evaluation' positions currently sitting with the PME Sector (e.g. staff with quantitative methodology expertise). ICRC should also consider hiring a dedicated 2-year post with expertise in contribution analysis, outcome harvesting, and outcome mapping to work specifically with the prevention, protection, and movement / cooperation métiers in developing / testing measurement approaches for 'impact' evaluations or any evaluation that includes DAC criteria on effectiveness. This would complement the work already underway in the PME sector on quantitative methods for impact evaluation. Total number of staff: 4 to 5 (the 5th being a 2-year 'consultant') in the central office. *(This number does not include existing 'impact evaluation' staff. It also counts a 'Head of Sector' as required new staff, when it remains possible that this position will share responsibilities if the evaluation function is co-located with monitoring and planning as currently is the case within the PME Sector).*

Note: The function will require dedicated time towards configuring SharePoint and BI systems to support a range of applications for evaluation (see full report for more details). This expertise may need to come through hiring a one-year IT consultant or may be available in-house through other initiatives (e.g. Data Transformation).

Table 2: Staff capacity at different levels of evaluation activity

Function Model	Levels of decision making
"Command / Control" Model	**Centralized Office:** All staff listed for this model under Table 1. **Geneva-based métiers:** Any métier-based RBM or M&E staff should have a 'blue line' to the evaluation function on evaluation related matters. Their ToRs need to be reviewed for stated authority on evaluation and adjusted to fit the new model adopted by ICRC. Their role would shift towards support and coordination of work led by the centralized evaluation function.
"Center of Excellence" Model	**Centralized Office:** All staff listed for this model under Table 1. **Geneva-based Departments:** If ICRC goes with the 'center of excellence' option for the evaluation function, it is *advisable* to create evaluation staff capacity *at the departmental level* of each HQ Department. This would be in addition to the staff situated in the central evaluation function team. This model requires a higher level of 'devolved' evaluation capacity than the 'command and control' model. Primary authority would sit with the departmental-level position(s) on departmental evaluation planning and budget procedures and approvals, quality control / assurance procedures and approvals, and leadership on evaluation use and visibility in departmental decision-making, strategy, and policy. This position would be the budget holder for departmental 'pots' of evaluation funding. At minimum this position / responsibility should be considered for Operations, but it is also recommended for the other HQ Departments. The capacity can be achieved in two different ways: a. *One cross-cutting, departmental level team and no métier specific staff*: Consolidating or streamlining existing RBM or M&E staff sitting within the métiers into single departmental-level PMEL teams. Update ToRs to reflect appropriate roles and responsibilities under this model. Teams have a 'blue line' to the evaluation function on evaluation related matters. No additional spending or hiring required. b. *One cross-cutting, departmental level adviser and métier specific staff remain/are created where absent*: Appointing a departmental-level evaluation only or PMEL/MEL adviser for each department from existing RBM / M+E staff. Where relevant, this position is located in existing departmental-level functions for RBM, Evidence/ Research, or M&E. This person coordinates and guides the evaluation efforts of existing métier-based RBM or M&E staff, and has a 'blue line' to the centralized office. In this model the remaining métier-based RBM or M&E staff have a 'blue line' to the departmental evaluation focal point on evaluation related matters. ToRs of métier-based RBM or M&E staff should be reviewed for redundancy with the cross-cutting position and any duplication of responsibilities removed. In these two approaches, no additional staff are hired (unless absent). Existing positions are formalized into expert posts and re-structured to promote coherence and efficiency. **Delegations:** Under this model, ICRC should review whether and where to place PMEL staff in the field. This would be best reviewed in a second phase of evaluation development. These positions / teams could provide cross-cutting support to the management and métier of delegations within that region. Positions would lead the process for evaluation identification and planning, while also providing leadership on evaluation use and visibility in delegation decision making and strategy. This can be achieved in a similar way as described above for the HQ departments (e.g. streamline existing positions), including the logic around 'blue lines'.
Hybrid *Recommended*	The *structure* of the 'hybrid' model for staffing at different levels of evaluation capacity mirrors that of the 'center of excellence' model, while including key distinctions in the *responsibility and scope* of staff to ensure greater authority over the evaluation process and connectivity with the central evaluation function. The 'hybrid' model should mirror for the 'center of excellence' model, with the following modifications on staff responsibilities: • **Geneva-based Departments:** Departmental evaluation plans must adhere to the process established by the evaluation function, including procedures for approval. Departmental evaluation capacity ensures ability of the central office to conduct quality control of decentralized evaluations. • **Delegation:** Evaluation plans must follow the process established by the centralized evaluation function, including procedures for approval. Field-level evaluation capacity ensures ability of the central office to conduct quality control of decentralized evaluations.

D. **Budget Ownership:** Where does the money sit for evaluations?

Based on the findings of the Assessment, **it is recommended that ICRC adopt a 'split ownership / controlled' model** for evaluation budget ownership, wherein dedicated funding for evaluation is created at centralized and decentralized levels to mirror the purpose and role of evaluation function and to foster evaluation practice across the organization. Budget ownership should proactively ensure funds are available and secured to conduct evaluations at a standard that is commensurate to the size and function of the organization. Arrangements for budget ownership should reflect and support a shift towards a coherent and deliberate approach to evaluation identification and planning, while promoting a more active evaluation culture.

Table 3: Evaluation budget ownership

Options	Description	Pros / Cons
Centralized / Controlled	Dedicated evaluation funds are held by a centralized evaluation function. All funding for all evaluation sits here. A clear budget or account code is created to ensure control over funds.	Pros: Works for a very centralized, accountability driven evaluation framework in a top-down institution. Cons: It will be difficult to achieve this at ICRC. Limits user-driven demand for evaluation.
Split Ownership / Controlled *Recommended*	Three types of dedicated evaluation funding 'pots': (1) held by the evaluation function; (2) held by the five HQ Departments; (3) a 'delegations' pot managed by the Director of Operations. Funds follow a use it or lose it rule – if the funds are not used for evaluation, they return to a 'central pot'. The ringfenced budget cannot be (re)allocated to or put into competition with other priorities. A clear budget code ensures control of funds. The evaluation function pot is for centralized evaluations that are managed or strongly supported by the function. Funds for 'decentralized' evaluations are held by the HQ Departments for Geneva-initiated decentralized evaluations and by the Director of Operations for funds dedicated to decentralized delegation evaluations. It is preferable that the HQ Departments manage their 'pots' directly. If departments regards this as an administrative burden, it is possible for the evaluation function to administer the budget on their behalf.	Pros: Ensures budget availability for a range of evaluation types and evaluation users, promoting an evaluation culture in the agency. Dedicated budget lines at different levels of operation reflect industry standards for guaranteeing evaluation resources. Creates visibility for evaluation and prompts discussion about the potential for evaluation. Dedicated budgets could also support other M&E or RBM activity to avoid perceived competition between monitoring and evaluation objectives. Cons: Feasible to lock in budgets at ICRC is uncertain. Requiring budgets for evaluation may lead to backlash, e.g. if stakeholders want the function to own the budget, in part, to protect their budgets from evaluation requirements.
Split Ownership / Partial Control	Dedicated evaluation funds are held by the evaluation function and used for centralized evaluations that are supported or managed by the function. The rest of the organization can set aside funds for decentralized evaluations *as they please*. There are no requirements or ringfencing. Budget instructions may provide guidelines for the optimal amount of evaluation spending in a given year per department. Annual planning process includes a 'prompt' for evaluation budgeting.	Pros: Quick solution for injecting resources in a newly re-constituted evaluation function, who can then use that budget to support different departments / units in building experience with evaluation. Lays the groundwork and provides buy in without backlash on budget requirements. Cons: Leaves potential evaluation users fighting for resources if they cannot access centralized funds. Limits promotion of evaluation culture and opportunity. '*As they please*' may mean not many evaluations are carried out.

E. **Funding source:** Where does the money come from for individual evaluations and the evaluation function?

Based on the findings of the Assessment, **it is recommended that ICRC consider a combination of line-item and core budget (re)allocation** for the source of evaluation funding, while not ruling out the advantages of a fully line-item request to donors for funding the function and individual evaluations. The approach to sourcing the budget is primarily a factor of possibility, followed by concerns about retaining independence over the evaluation agenda. The decision for funding source should consider longer-term sustainability together with flexibility during the different phases of the developing the function. For example, years 1 and 2 of launching the function may require more funds for targeted development of guidelines or expertise in the organization as compared to years 5 and beyond.

Table 4: Evaluation budget source

Options	Description	Pros / Cons
Core budget re-allocation	Reallocate required funds from the existing core budget (HQ, Ops, etc.) to cover the financial needs of the evaluation function and evaluations.	**Pros:** Maintains ICRC control and discretion over evaluation agenda. **Cons:** Internal discipline is required to set aside and ensure these funds, which may be difficult to secure during the initial phases of the evaluation function. Could 'crash and burn' if or when people feel that they are not 100% getting what they want from evaluations, including reaction to critical findings. Office politics are linked to the budget ownership decision and degree of reallocation between departments.
Line item request to donors	Ask donors to fund a line item for individual evaluations and the evaluation function *in addition to* the funds they currently provide through the core budget. Every operation and every department could include a new line item on evaluation in the annual appeal to donors as a way to limit visibility of the addition.	**Pros:** Ensures budget availability and locks in the funding by using donor commitments as a protective ring around the budget. Potentially less internal politics as compared to reallocation. Greater achievability. **Cons:** Potential donor interference in ICRC's evaluation agenda. Line item request may lead to increased donor expectations and reporting.
Combination line item + core budget reallocation *Tentatively recommended.*	Line item request to donors on some aspects of the development of the function (e.g. 2-3 'groundwork' evaluations, and possibly 1-2 short-term technical experts hired as 'consultants' who are staffed in the function), the visibility of which to donors would be an advantage to ICRC. Select flexible and accommodating donors for this, e.g. Scandinavian and German funders. Avoid DFID, Echo, World Bank. Additional costs are funded through core budget envelopes. This includes the proposed 'pots' of funds described in Table 3 and the costs of any existing or newly hired function support staff.	**Pros:** Quick solution for injecting resources in a re-constituted evaluation function, while avoiding issue of perceived funding grab from within the core budget for the function. Donor commitment provides a ringfence around the funding, while ICRC selects the extent to which that commitment goes within the overall evaluation agenda. **Cons:** Potentially too complicated. Leaves open the 'cons' of both core and line item approaches. Might have sustainability issues in the long run if too much of the funding comes from line-item requests.

F. **Budget Amount:** How much money?

Based on the findings of the Assessment, **it is recommended that ICRC adopt allocation benchmarks** for determining the budget amount for evaluations and the evaluation function. Allocation benchmarks proactively ensure funds are available to conduct evaluations at a standard that is commensurate to the size and function of the organization. It creates a predictable funding base around which impact evaluation can be planned and staffed, while enabling flexibility on how evaluation funds are prioritized for spending (e.g. staffing vs. technology). Allocation benchmarks mirror best practice in humanitarian organizations and industry standards for how evaluation budget amounts are determined.

Table 5: Evaluation budget amount

Options	Description	Pros / Cons
Allocation benchmarks *Recommended*	Dedicate a set proportion of the total operating budget towards evaluation. The agreed amount should follow a phased approach to ensure absorption capacity and to allow time for ICRC to develop sufficient evaluation practice expertise to responsibly manage the funds. **Years 1 and 2:** The UN JIU suggested standard is 0.5 – 3% of operational budget going towards evaluation. However, benchmarking against UNHCR - the UN agency most like ICRC in terms of scope – provides a more realistic target for Years 1 and 2. UNHCR currently has a budget for evaluation of around .07% (2018) of its operating budget. In 2017, this amounted to 3.58 million[43] in funding for evaluations. Benchmarking ICRC to UNHCR would suggest that an ICRC evaluation budget level 'commensurate to the size and function of the organization'[44] would be **1.4 million CHF.** This would cover the costs of staffing the centralized office of the evaluation function *and* the budget for individual evaluations for the two 'pots' of funds held in Geneva. It does not include the costs for staffing departments or establishing staff capacity for the field and does not include a budget for evaluations at the field level (the third 'pot'). If new positions at these levels are required, however, at the start of the evaluation function (e.g. ICRC is unable to re-organize existing staff capacity), then the allocation for Years 1 and 2 should increase to around 0.1% of ICRC's operating budget, or an estimated amount of 2 million CHF. This would enable hiring three to four additional full-time staff in priority placements at the departmental level[45]. **Year 3**: The budget allocation and amount should increase to create the dedicated delegation evaluation budget that was not included in Years 1 and 2. This assumes that ICRC has established sufficient centralized capacity to govern and support evaluation practice across the organization by the second year of the function. The amount of increase depends	**Pros:** Ensures that evaluation practice is commensurate with the size and scope of the organization. Adheres to industry best practice and industry standards. Predictable funding base around which longer-term impact evaluations can be planned and staff can be hired. Flexible prioritization of how funds are directed to line item priorities (e.g. function staff, developing a new methodology, etc.) **Cons:** Agencies using rates need to track where the spending is directed to adjust line item priorities over time as an evaluation function moves towards greater maturity and efficiency. Capacity to do this should be ensured, or the approach can yield less effective results. If evaluations are not perceived as immediately useful, this approach can be subject to attack (e.g. perceived as not 'needs based').

---

[43] See UNHCR 2017-2018 MOPAN Assessment, page 44: http://www.mopanonline.org/assessments/unhcr2017-18/UNHCR%20report%20[web-1a].pdf
[44] UNEG Norms and Standards, Norm 13 point 17
[45] Assuming the cost of one FTE is 160k CHF/year, as per correspondence with ICRC.

	on whether additional staff were required at the departmental level. It is anticipated that the Year 3 allocation will sit at either 0.1% of the total operating budget (2 million CHF) if additional hiring was not required in Years 1 and 2 or at 0.13% (2.6 million CHF) if additional hiring was needed. This would enable both a sufficient amount of funds to cover decentralized delegation evaluations in Year 3, including the cost of staffing (if required) to support expanded field capacity and the cost of individual evaluations. As with the departments, it is recommended that ICRC explore ways to re-organize existing staff capacity for field support rather than hire new positions. **Years 4 and 5**: In Year 4, ICRC should review the allocation amount it is using to determine if it is sufficient and adjust as needed. On balance, the organization should aim to have a stable and set allocation by Year 5 that it will continue to use in the years going forward, allocating no less than 0.1% and no more than 0.3% of total organizational expenditure towards evaluation work.	
Allocation according to a number of annual evaluations	Set the budget according to a predetermined number of planned central evaluations and estimated decentralized evaluations. Establish staffing level, and consequent budget, of the evaluation function based on this ambition. This would mirror aspects of the previous practice of ICRC under the 'GENEVAL' evaluation function.	<u>Pros:</u> Can provide a 'soft landing' for newer initiatives on evaluation where the funding approach is politicized or where the agency is unsure of its absorption capacity for dedicated evaluation spending or ability to translate those funds into accepted results. <u>Cons:</u> The targeted number of evaluations may be – or appear to be – arbitrary. It does not secure sufficient funds to produce a level of evaluation activity and quality commensurate with the size and scope of the organization. The budget level is vulnerable to reduction or politicization because it is not pegged to an objective standard. Does not provide flexibility on prioritization of line items.

ANNEX VIII: Information Technology and Evaluation: *Considerations and recommendations*

INTRODUCTION

This annex details recommended ways ICRC can leverage its existing IT systems in furthering the quality and use of evaluations.

It responds to Objective III[46] of the ICRC Evaluation Function Assessment ('the Assessment') and elaborates on Recommendation 5, covered in the main report (Section 4)

RECOMMENDATIONS

I.   Areas of Application

The Assessment identifies three interrelated areas of evaluation management and use where IT application would provide a significant benefit to the evaluation function (see Figure 1 below for details). This includes:

Management of individual evaluations. Process of managing centralized or decentralized evaluation workflow. It requires working through the following steps consistently:
- Evaluation planning and decision-making (approval of evaluation initiation, individual evaluation budgets and budget source)
- Evaluation Terms of Reference (developed according to guidance/template and uploaded to central repository).
- Evaluation inception reports (developed according to guidance/template and uploaded to central repository).
- Evaluation reports (developed according to guidance and uploaded to central repository).
- Evaluation management response (developed according to guidance/template and uploaded to central repository).

Evaluation function operations. The fulfilment of responsibilities by a centralized evaluation function charged with managing and maintaining ICRC's evaluation strategy, agency-level guidelines, overall quality control of evaluation practice. It also involves direct management of centralized evaluations.
- Managing the evaluation process for *centralized* evaluations (having direct control in managing evaluation ToRs, recruiting of evaluators, overseeing the evaluation process and approval on evaluation products, steering the evaluation management response, fulfilling tasks around evaluation publication and dissemination).
- Tracking the process of and providing a degree of quality control for *decentralized* evaluations (having access to information on evaluation planning and access to evaluation ToRs, evaluator bids and contracts, evaluation reports, and evaluation management responses).
- Systematic extraction and categorization of key information (e.g. evaluation findings or recommendations, information about the evaluations, etc.) for use across the organization.
- Systematic archiving of all evaluations according to a consistent set of evaluation parameters (such as keywords, geography, program type, subject, date etc.).

---

[46] "Advise on the IT systems and tools that respond to the future business operations of the Evaluation Function" – Assessment Objective III

Evaluation use (*beyond individual evaluation response*). The use of data on and from evaluations in a range of functional areas and by positions across the organization.

➤ Visualizing evaluation coverage and frequency (by subject, geography, date etc.). This is important for centralized evaluation planning and prioritization (e.g. ensuring that sufficient evaluation coverage is achieved) as well as for ensuring staff are aware of evaluations relevant to their work.

➤ Visualizing data on evaluation management responses (through progress traffic lighting, for example) to ensure transparency and promote action on accepted evaluation recommendations.

➤ Presenting and disseminating summaries of findings and recommendations from evaluations. This can involve *proactive* dissemination of evaluation information to users based on defined position attributes (level, position, area of work etc.), as well as incorporating evaluation findings and recommendations into existing dashboards on organizational performance (agency, department, unit, team, function, or delegation level).

➤ Linking evaluation findings to dashboards and data generated within the monitoring / MfR reporting platform (e.g. evaluation results are dashboarded with relevant MfR data according to the level of analysis, such as delegation specific, métier-wide findings, or organizational trends)

➤ Ensuring evaluation reports are accessible from a central report library and on the ICRC intranet in a manner that is searchable using a consistent set of search parameters (geography, métier(s), evaluation type, date, etc.). Report library should be linked to any knowledge management platforms, existing or planned.

Figure 1: Recommended areas of IT application in ICRC evaluation practice

|	Description of purpose	Primary users
3 — Use of evaluations (beyond individual evaluation response)	• Visualize what evaluations have taken place. • Visualize data on evaluation management responses. • Make accessible/present/disseminate/ summaries of findings & recommendations from evaluations. • Access evaluation reports as required.	• Senior management (for oversight and management of evaluation planning, response and prioritization) • Specified target groups (based on profile) • All staff (accessibility) • External (as needed)
2 — Management of evaluation function	• Track the process of evaluation management for decentralized evaluations. • Systematically extract and categorize key evaluation information for use across the organization. • Systematically assess and archive evaluations according to consistent evaluation parameters.	• Technical staff responsible for evaluations
1 — Management of individual evaluations (centralized or decentralized)	• Manage process for evaluation planning and decision-making. • Develop evaluation Terms of Reference. • Manage inception reporting. • Manage evaluation reports. • Develop and manage evaluation management response.	• Staff responsible for managing individual evaluations • Staff responsible for responding to evaluation recommendations

II.   IT Dependencies and Considerations

Leveraging technology to improve the management of individual evaluations (level 1), management of the evaluation function (level 2), and the use of evaluations overall (level 3) requires an understanding of the interdependence of data flow between each level.

The way each area of technology application is structured will have implications for the IT use possibilities in other areas. For example, a balance must be struck between the level of ambition for how technology can 'automate' analysis of evaluation data at the organization and flexibility on the way different types of decentralized evaluations are managed.

The main areas of interdependence are highlighted in Figure 2.

Figure 2: Data dependencies between levels of IT application
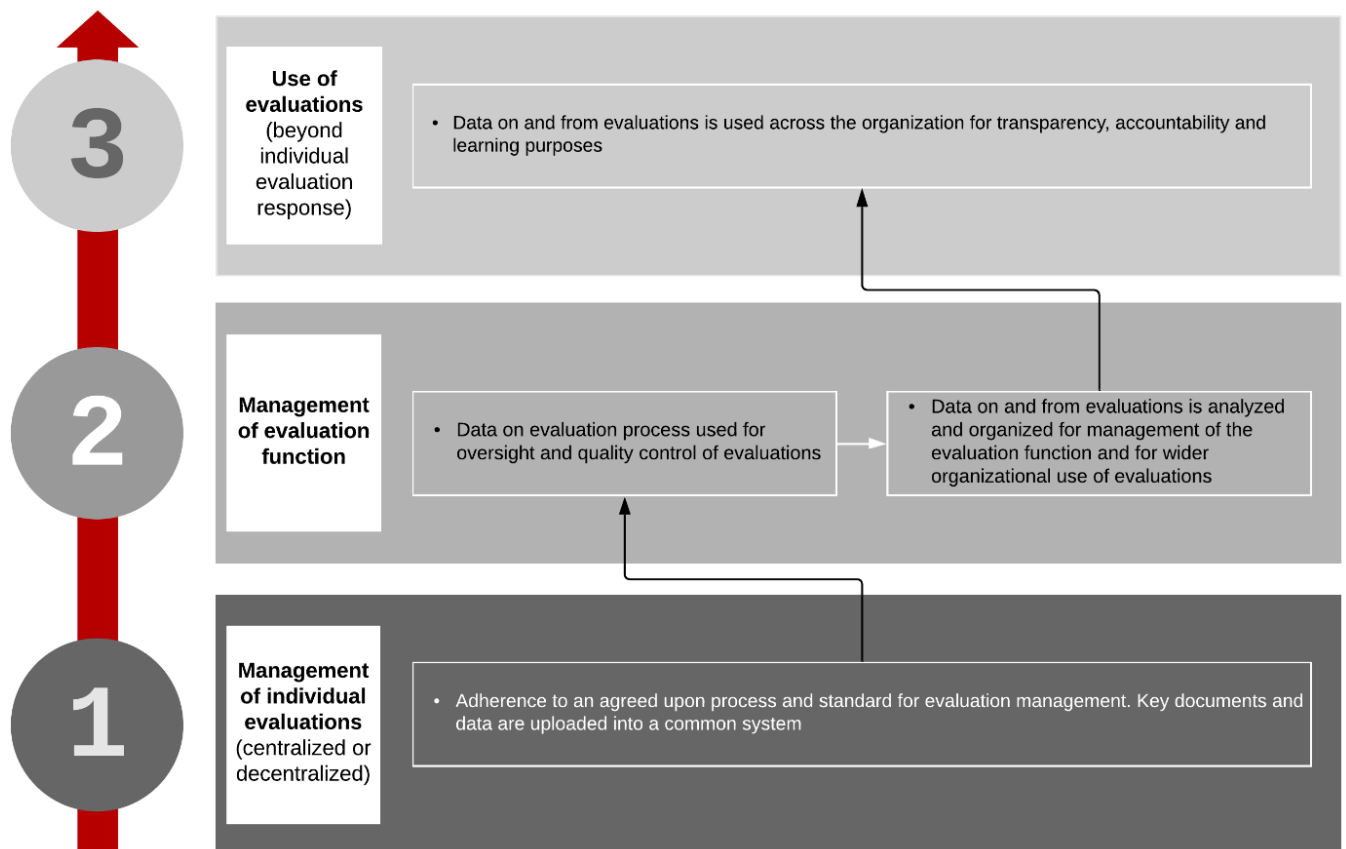
III.   Software and IT Platforms

Given the areas of IT application outlined in Section 1 and Figure 1, **the introduction of relatively simple database and warehousing tools can enable ICRC to leverage existing software (SharePoint and Tableau) to meet most needs.** This is based on a number of assumptions:

➢ Capacity (technical, staffing, and budget) exists to develop/maintain database and warehousing tools and configure SharePoint to manage the evaluation workflow (Step 1 in Figure 1 above).
➢ SharePoint is effectively linked to a system for document and file storage, archiving and retrieval.
➢ Capacity (technical, staffing, and budget) exists to link Tableau to data warehousing in SharePoint for evaluation analytics.

Areas of application where *additional* software may be necessary/beneficial are:

➢ Database and warehousing tools for organizing and storing key evaluation data
➢ Qualitative analysis of evaluation documents (e.g. Nvivo). This could include synthesizing reports for an annual 'learning review' or producing a specific piece of analysis for a set of users (e.g. responding to an area of the Institutional Performance Monitoring Framework through an analysis of findings on accountability to affected populations found across all evaluations produced in a year).
➢ Managing data on evaluation management responses. This can likely be visualised using Tableau, but requires a system for managing response data (could potentially be SharePoint).
➢ Managing a roster of evaluation consultants and/or a roster of ICRC staff that participate in evaluations (mixed team) or reviews (mixed team or fully internal). This may require a system for managing systematic feedback on the consultants or staff (could potentially be SharePoint)
➢ Collection of evaluation tools / methods used across evaluation exercises (e.g. surveys, discussion guides, etc.). This may require a system for managing systematic feedback on the utility and performance of the tool and applicability in the future (could potentially be SharePoint).

IV.   Recommended System

Factoring in issues of data dependencies and software availability for the proposed areas of application, the recommended system is presented below in Figure 3.

Note that where '?' is placed next to a platform or type of software, it indicates that ICRC would need to review whether the platform is already capable of performing this task or if bespoke configuration is necessary. In some instances, it may be that the cost of configuration outweighs the benefit of using the platform. For example, it may be the email communication is preferable to a SharePoint notification.

ICRC can address this level of detail after it has approved a final evaluation framework and launched the evaluation function.

Figure 3: Recommended system for IT application in evaluations at ICRC
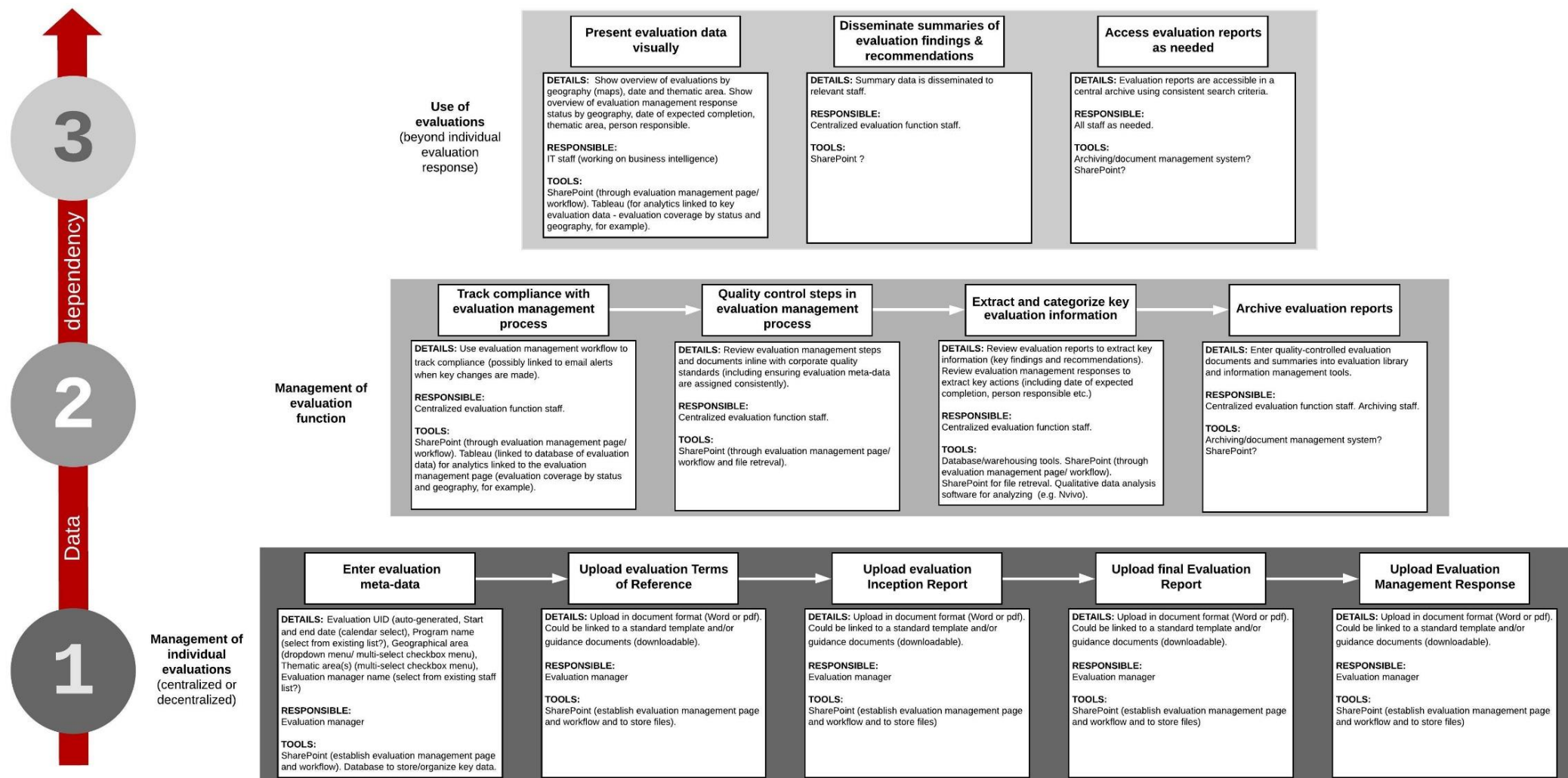
V.   Organizational Prerequisites for IT Application

Three main aspects of the evaluation system and ICRC structure need to be defined or agreed upon to address the dependencies and considerations described above. They include:

Roles and responsibilities. Using IT systems for evaluation use and management relies on defining and institutionalizing position responsibilities linked to evaluations. This requires:
- ➢ Clearly defining the evaluation responsibilities linked to certain types of positions
- ➢ Evaluation responsibilities included in job descriptions and terms of reference
- ➢ Evaluation responsibilities included in performance management processes
- ➢ Linking of workflow responsibilities with specific position roles

Structural definitions of the institution. It is necessary to establish a consistent understanding of:
- ➢ Geographical areas and groupings
- ➢ Subject/thematic areas of work
- ➢ Organizational levels and divisions of responsibility (e.g. what are the 'field', 'region', 'HQ' levels; what are the 'metier', 'unit', 'sector', 'division', 'department' distinctions)

Evaluation framework. A clear set of evaluation parameters (i.e. how evaluations are defined and categorised) need to be established and linked to the specifications for evaluation workflow. This includes:
- ➢ Evaluation type
- ➢ Required documents per evaluation type (e.g. is a documented management response required or not)
- ➢ Required steps per evaluation type (e.g. is external publication required or not)

Recommendations on the ICRC evaluation framework are covered in the main report under Section 4 and in Annex IX 'Evaluation Strategy.'