STOCKHOLM
EVALUATION
UNIT

MEDECINS SANS FRONTIERES
LÄKARE UTAN GRÄNSER

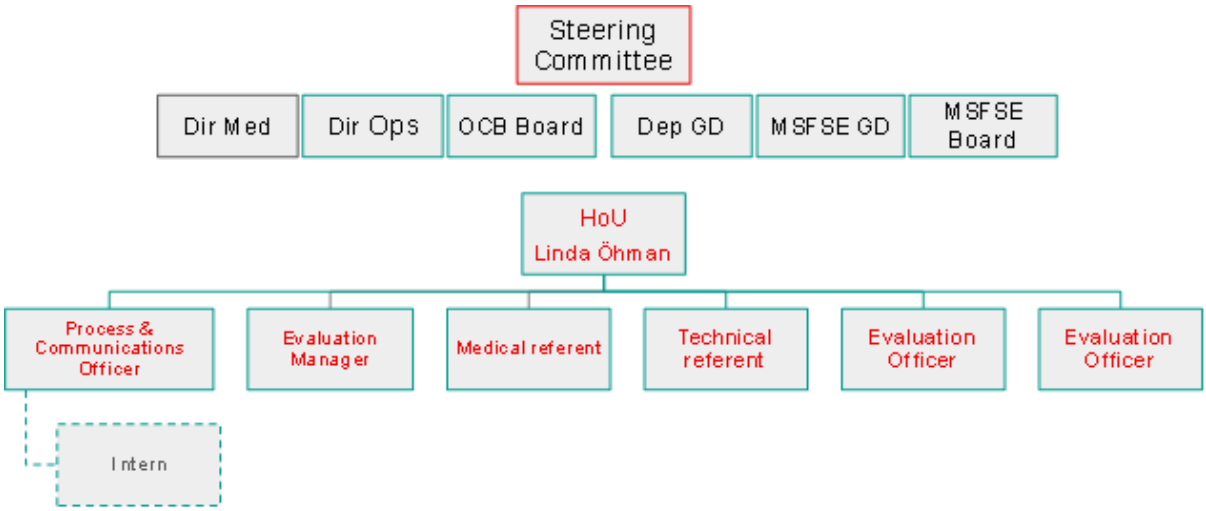
# ANNUAL

# REPORT

# 2025

The Stockholm Evaluation Unit (SEU) is the primary mechanism to support the MSF Operational Centre Brussels (OCB) in conducting project evaluations. The unit is based at and largely financed by MSF Sweden.

It is accompanied by a Steering Committee, which includes representatives of OCB Operations, Medical Department, the General Director office, the OCB Board, MSF Sweden Board, and MSF Sweden. The Steering Committee is chaired by one of its members and meets quarterly.



# CONTENTS

Click to jump to a section

# The SEU in 2025: Beyond Reports

Hi there,

It is time for another annual report from the Stockholm Evaluation Unit. We are happy to share not only key evaluation information and overarching findings, but also reflections and thoughts on how the processes went and what we learned in 2025.

In 2025, we completed eleven evaluative exercises – but worked on 30 throughout the course of the year. While we are still refining how we manage our resources – time and budget – both improved overall in 2025, with faster and more cost-efficient evaluations.

We want to continue making evaluation more relevant to MSF, and so we continue to address these issues. In 2025, we reinforced our focus on designing evaluations as user-focused, engaging and empowering processes. We see this as key to making evaluation useful – particularly for OCB – and will continue to build on it. In this Annual Report, colleagues reflect on how well we achieved this, and what we learned along the way.

We also explore engaging people beyond our own walls – where we have a lot to learn. Yet this does not mean we need such engagement in every evaluation. The way the evaluation is designed needs to be aligned to how the project is designed and implemented. If, for example, driving community co-ownership is not a project objective, it will be hard and not necessarily useful to make it a key element of an evaluation.

Ensuring the use of evaluations goes beyond the individual project. To make findings more accessible, we continue to produce transversal analyses in various formats, including papers and briefs (in 2025, on handover and on capitalizations). As in previous years, this annual report includes an overview of reoccurring themes across evaluations, now also linking them to the new Strategic Ambitions and identifying what should be replicated, scaled up and what requires (urgent) attention.

AI has made this type of analysis much easier and allows us to crunch far more data. Before we mainly relied on the evaluations' recommendations – now we can draw from entire reports. Much better according to us, as the most valuable insights lie in the findings. AI also speeds things up, enabling us to quickly put together reflections based on the evaluations – with the caveats and limitations that it entails. But AI can only source from what has been evaluated, and there are areas and priorities where evaluation remains limited. If you have questions that evaluation could help unpack, don't hesitate to reach out. And for inspiration, check out the questions AI answered for us in this report.

Finally, and not everyone knows this, besides the evaluations we manage, we also serve as a sounding board for those seeking input and support in developing different inquiry processes. This can be formal, like our contribution to the SPARC M&E framework or the presentations we did on the difference between capitalizations, operational research and evaluations (Hint: don't put the cart before the horse, start with the purpose!). But it can also be ad hoc, through a quick call or email. So, if you are thinking about something, reach out. It may not be an evaluation, but we are always happy to help figure things out!

Linda Öhman

Head of the Stockholm Evaluation Unit

# A Year in Review – 2025 Evaluations

| Evaluation of | Description |
|---|---|
| **Decentralization Initiative in Bangui, Central African Republic** | Mid-term evaluation exploring how decentralized SRH/HIV services were implemented in Bangui, how they were experienced by clients, and the extent to which health outcomes and system integration improved over the period. |
| **Integrated Community Case Management (iCCM) activities in MSF-OCB settings** | Thematic evaluation exploring the implementation, performance, and strategic recommendations for scalability and sustainability of iCCM activities in the MSF-OCB portfolio as per end of 2024. |
| **Pool d'Urgence Congo (PUC), DRC** | Evaluation assessing the PUC's suitability to address emergency needs in the DRC, analysing its relevance, strategic positioning, and operational effectiveness. |
| **The Basic Healthcare Project for Artisanal and Small-Scale Gold Miners in Gwanda, Zimbabwe** | Mid-term evaluation assessing the project's relevance, effectiveness, efficiency, and sustainability, aiming to generate practical learning to inform adaptation and transition planning for the remainder of the project cycle. |
| **Operational Centre Geneva Partnership Framework** | Evaluation assessing the added value and relevance of the OCG Partnership and Framework Agreement, providing an understanding of its design, implementation, and results, and generating recommendations to guide future collaboration. |
| **The Couffo Project, Benin** | Mid-term evaluation analysing the relevance, effectiveness, efficiency, and sustainability of the SRH/maternal health interventions carried out in Couffo, Benin, since 2022. |
| **The OCB Emergency Response in the Nuba Mountains, Sudan** | Evaluation examining the strategic drift from emergency response to hospital support for accessible populations, the institutional dynamics that enabled it, and the lessons for organizational decision-making under operational ambiguity. |
| **"A New Era of Working with Community Organisations"** | Evaluation documenting the project's evolution, added value, relevance, and contribution to strengthening CBOs after nearly four years, informing its future direction. |
| **OCB International Mobility Unit processes** | Evaluation assessing the human resource processes of the OCB International Mobility Unit (IMU) and Field Admin teams related to the deployment of international mobile staff. |
| **Capitalization of** | **Description** |
| **The Sexual Violence Portel project in Marajó, Brazil** | Capitalization analysing the MSF Portel project in Marajó (2023–2025), focusing on its strategic innovation: the development of a comprehensive care model for sexual violence survivors within formal health system contexts. |
| **The OCB Yanomami Project, Brazil** | Capitalization of MSF's intervention in Yanomami Indigenous Territory (Feb 2023–Nov 2025), following the national humanitarian emergency declared in January 2023. |

## ALIGNMENT WITH OPERATIONAL PRIORITIES

| THEMES | Bangui | iCCM | PUC | Gwanda | Couffo | Nuba | Marajó | Yanomami | TOTAL |
|---|---|---|---|---|---|---|---|---|---|
| **Epidemics** | | | X | | | | | | 1 |
| **Conflict and Violence** | | | X | | X | X | X | | 4 |
| **Migration and Detention** | | | | | | X | | | 1 |
| **Sexual Reproductive and Women's Health** | X | | | X | X | X | X | | 5 |
| **Child Health and Nutrition** | X | X | X | X | X | X | | X | 7 |
| **Trauma Care** | | | | | | | | | |
| **Chronic Infections: HIV, TB, Hepatitis** | X | | | X | | | | | 2 |
| **NCD** | | | | X | | | | | 1 |
| **Continuum of Care** | X | X | | X | X | X | X | X | 7 |
| **Clinical Care** | X | | | X | X | | X | X | 5 |
| **Antibiotic Resistance** | | | | | | | | | |
| **Environmental health** | | | | X | X | | | X | 3 |

Not included as not relevant: OCG Partnership Framework, OCB IMU and TIC Project Working with CBOs.
Find more information about the operational priorities here.

### EVALUATIONS IN 2025 | FACTS AND FIGURES

**Where**: Central African Republic, Democratic Republic of Congo, Zimbabwe, Benin, Sudan, Brazil, HQ (Belgium, Switzerland)

**What topics**: Sexual Reproductive and Women's Health, Child Health and Nutrition, HIV/TB, NCD, Environmental health, Outbreak response, Community Engagement.

**Which evaluators**: Individuals, teams and/or consultancy firms representing country contexts such as Benin, France, Canada, Zimbabwe, Italy, USA, Cameroon, Spain.

**Dissemination**: An average of 55 participants attended the final webinars, nearly a third of whom were based at the project level, while a quarter were from other operational centres.

**Total average cost** per evaluation completed in 2025: 33,499€

# Evaluations' Recurring Themes, Good Practice and Persistent Gaps

As a standard deliverable of the SEU Annual Report, here we include an analysis of the main recuring themes across 2025's evaluations. They generate institutional learnings covering learning, project design, community engagement, and – as in previous years – exiting a project and handing over.

The SEU has, as in previous years, conducted transversal analysis based on 2025's evaluations. This analysis draws on six evaluations: the thematic evaluation of integrated community case management (iCCM) activities across eight OCB projects; the Pool d'Urgence Congo emergency response project in the DRC; the Couffo maternal and neonatal health project in Benin; the basic healthcare project for artisanal and small-scale gold miners in Gwanda, Zimbabwe; and the TIC project on working with community organisations across Southern and Eastern Africa. Findings from evaluations not yet publicly available at the time of preparing this report are integrated into the synthesis without direct attribution.

Together, these evaluations span emergency and stable contexts, urban and remote settings, fragile and functional health systems. These are different operational logics, different health system contexts, and different population groups. Yet the SEU has identified six themes that recur throughout.

Considering the varied contexts, successes and problems documented here are not the products of challenging contexts or insufficient project teams. They are features of how OCB designs, manages, and closes projects; patterns stable enough across a wide and diverse collection of evaluations to be named as organisational. Some are accompanied by genuine good practice.

The SEU used AI to conduct the analysis presented in this paper (as was done in 2025). Since most of the reports are public, they can be uploaded into an AI tool, allowing for analysis of the whole report rather than only the recommendations, as was the case when done manually. The analysis is limited primarily by the evaluations themselves. Though the SEU works actively to produce good evaluation, there are of course some that are better than others. Likewise, the evaluations only cover so many topics and does not reflect the spread in projects that OCB is looking for (more evaluations cover projects that serve marginalized groups, rather than conflict/violence, for example).

## THE RECURRING THEMES OF 2025'S EVALUATIONS

For each of six recurring themes, the paper presents *what the evaluations* show (synthesised across the 2025 evaluations); *what OCB should build on*; *what OCB needs to address*; and *which Strategic Ambitions the findings connect to*. The full paper is available here, with the themes summarised below.

### OCB Acts Locally but Does Not Learn Globally

The evaluations document how OCB projects adapt intelligently yet fail to convert those adaptations into organisational knowledge.

### Community Engagement as Operational Architecture

In cases where community engagement (CE) is designed into the operational model from the outset, outcomes were measurably different. Where it was added around the model, or its relational demands were underestimated, it became a source of fragility.

### Exit Is Planned Last and Designed Least

No evaluation conducted in 2025 found a project with a credible, costed transition plan in place at the point of assessment. OCB designs for service delivery and defers transition planning to later phases that never fully arrive. Of course there can be positive examples, in projects not evaluated.

### Quality depends on a systemic approach

Evaluations show that quality is not determined by any single input but by the interaction of infrastructure, training, and supervision as a system.

### Partnerships – Building Presence Without Building Ownership

OCB's partnerships are consistently positive in perception and fragile in practice.

### Evidence: What we know and what we do with it

The 2025 evaluations show examples of where evidence generation stops at generation and never reaches influence, both for external influence and internal learning and adaptation.

## WHAT THE 2025 EVALUATIONS MEAN FOR OCB'S STRATEGIC AMBITIONS

The SEU matched the recurring themes to the 2026-2031 Strategic Ambitions (SAs) to identify areas where evaluations illustrate progress toward these ambitions, as well as where further effort is needed. These are intended as organizational learning, not judgments of success or failure.

### Where OCB is already delivering what the Strategic Ambitions require

- **Adaptive service delivery in changing contexts (SA 1, 3):** iCCM teams have redesigned routes based on real-time pastoralist migration data, while in Couffo the introduction of SAC was delayed by two years to prepare the social ground in order to be able to implement well. Both are examples of the contextual responsiveness the Strategic Ambitions call for.
- **Ecosystem-based community engagement generating measurable results (SA 5, 7):** the Couffo whisperer model and Gwanda's peer educators demonstrate that designing engagement into the care model from the outset changes who can be reached and what outcomes are achievable.
- **Community-owned sustainability under the right conditions (SA 7):** the Tohou banana plantation (Couffo), the Lokogba COGES budget line (Couffo) or the Gwanda peer educator network are not coincidences. They appear in projects where communities were involved in designing the services they rely on. OCB has proof of concept; it needs to make these conditions deliberate and replicable.

### Where the gap between Strategic Ambitions and current practice is widest

- **The organisation acts locally but does not learn globally (SA 12, 17):** OCB must build the infrastructure for learning to circulate rather than relying on evaluations to surface what daily practice should capture - through protected reflection time, standardized documentation, and cross-project analysis functions.
- **Transition is designed last and funded least (SA 7, 17):** None of the 2025 evaluations identified a credible transition plan in place at the time of assessment. This is a design issue rather than a timing problem. OCB needs to enforce exit decision-points at project validation, not discover their absence at evaluation stage.
- **Partnership creates presence without ownership (SA 7, 19):** OCB builds conditions within partner systems that those systems cannot sustain, then discovers the fragility at exit. The PUC's surveillance dependency and the iCCM portfolio's parallel operation to national systems are the same failure at different scales.
- **Evidence generated, advocacy absent (SA 6, 8, 17):** OCB is producing operational evidence and analysis that could influence both policy and internal decision making – but does not. The link between evidence and influence does not systematically exist. Addressing this gap is less about research capacity and more about organizational strategy.

# Consultation group members' view

Graham Inglis (Project Coordinator) and Munyaradzi Sidakwa (Project Medical Coordinator) were both part of the Consultation Group of the evaluation of the *Basic Healthcare Project for Artisanal and Small-Scale Gold Miners in Gwanda, Zimbabwe* (2025).

The evaluation reports and webinar recording are accessible on the OCB intranet – SEU library.

**How has the evaluation influenced the project, and what actions have been taken as a result?**

The evaluation has been highly relevant to ongoing decision making. Its emphasis on sustainability helped sharpen discussions with the Ministry of Health (MOH) about what is realistic to hand over. This has directly informed negotiations around establishing satellite clinics in mining locations. The MOH already has a national model for attaching such clinics to small health posts, and MSF is now planning to fund the construction of a few strategically placed sites in artisanal mining communities, alongside initial staffing and pharmaceutical continuity.

The evaluation also pushed the team to think more deliberately about the role of peer educators. As a result, MSF has intensified engagement with the MOH on how peer educators might be integrated as community health workers—or recognized as a distinct cadre. In parallel, MSF is preparing advocacy and collective action skills training so peer educators can eventually represent themselves within health networks.

The evaluation also encouraged us to think about handing over knowledge and evidence in advance of the project's closure. We will soon be holding a "capitalization" discussion within the team. We want to clarify what knowledge should be captured, and for whom—since some is more for MSF's institutional memory, but some can be useful to feed into MOH processes and policy.

In terms of formally sharing the findings more broadly with our stakeholders, this remains an action point, but the evaluation itself continues to feature regularly in internal and external discussions.

**How did you experience the evaluation process while the evaluators were on the ground?**

We found the debriefs with the Evaluation Team during data collection extremely useful. These sessions provided insight into stakeholder meetings the evaluation team were conducting and enabled the project team to start interpreting the trends and acting on the evaluation insights while the evaluation was ongoing. We drew parallel conclusions to the evaluation team.

We really appreciated that the evaluation team members who conducted the data collection were local. They were attuned to both cultural dynamics and the wider public health landscape in Zimbabwe. While the team's clinical and operational experience was somewhat limited, the qualitative insights had strong value and contributed meaningfully to strategic discussions in the project.

**What are your reflections on the role and functioning of the Consultation Group?**

We appreciate the intention behind the Consultation Group, but there was uneven engagement across the group. The SEU could perhaps do more to communicate the expectations and responsibility of being a Consultation Group member. Better advance planning could also make a difference. Blocking time in calendars well ahead of meetings and requests for feedback on long documents may help ensure more consistent participation.

# Impact: The Lasting Change We Want

This paper aims to break down MSF's conflicted posture toward sustainability by examining MSF's practice to identify areas where we strive to contribute toward lasting change. It does so by elaborating the COMPASS definition of Impact in a way which retains what MSF values about sustainability and leaves out what it does not.

Does MSF "do" sustainability? That question brings to the surface tensions and questions related to MSF's mandate, principles, ethics and our core identity. On the one hand, sustainability pits the immediacy of humanitarian action and the principle of independence against the implications of engaging with longer-term and political processes that seek structural change. On the other hand, sustainability confronts us with practical and ethical questions related to our obligations to patients, partners, and communities' well-being beyond our presence.

Thus, whether MSF values sustainability, and if so how, is murky and ill-defined. The word does not appear as a programmatic priority in strategic documents, such as MSF-OCB's Strategic Ambitions for 2026-31. Yet within evaluations of MSF projects conducted by the SEU, sustainability frequently appears as a criterion – not simply because of its inclusion in common evaluative frameworks – but because MSF stakeholders ask for it. Questions about whether a Ministry of Health will continue what MSF starts, how patients will continue to access care after we leave, or the long-term effects of an intervention are of major concern. Our experience shows us they matter deeply to those who bear responsibility for MSF projects.

## MSF'S DEFAULT POSTURE TOWARD SUSTAINABILITY: AMBIVALENCE

We conducted a comprehensive review of findings related to sustainability and handover from 45 evaluation reports between 2020-2025, which further confirms this ambiguity. While variations exist across evaluations, the highest-level general finding is that, by default, MSF adopts a posture of ambivalence toward sustainability. We do not explicitly reject sustainability as a value or objective, but we do not embrace it either. We tend to pursue sustainability opportunistically as a secondary objective, when conditions allow. The result is, unsurprisingly, that MSF does not tend to perform well on metrics related to sustainability.

While MSF may or may not be OK with that general finding, many evaluations highlight specific problems related to sustainability that MSF does care about: problems we create for our patients, for our partners, and for communities after we leave. Evaluations document instances of loss of trust from communities, discontinuity of care for patients, dynamics of dependency, and strained partnerships with Ministries of Health and other actors, among other things. MSF may be ambivalent toward sustainability as a concept, but we are not ambivalent about those kinds of consequences. Being better partners, integrating patient-centeredness, strengthening continuity of care, and strengthening community partnerships are all areas for focus in MSF OCB's Strategic Ambitions.

## A DIFFERENT VOCABULARY TO REFRAME THE ISSUE

Perhaps there is a need to find new ways of talking about these issues without the conceptual baggage we associate with sustainability. The framework underpinning our work to strengthen project monitoring, COMPASS, offers a compelling framework to build from. COMPASS includes four normative criteria which describe what MSF-OCB considers to be "good" projects. A good project, according to COMPASS, should be relevant; it should be effective and offer high-quality outputs; it should be efficient, and it should contribute toward positive impact.

The last criterion is interesting to this analysis, because COMPASS defines Impact in a way that is adapted and unique to MSF:

> *Impact looks at the final scope and broader effects of our interventions. It asks whether our work contributes to real and sustained change — in people's health, well-being, and dignity. It also includes what MSF learns or influences through its presence — from changing practices to generating new knowledge or advocacy evidence.*
>
> *MSF OCB COMPASS Framework*

## METHODOLOGY: ELABORATING A DEFINITION OF IMPACT

This paper elaborates on the definition of "Impact" by identifying sub-criteria which describe more specifically the ways in which MSF-OCB seeks to effect "lasting change" with its projects, based on data from MSF SEU evaluations. To do that, we extracted all findings and recommendations relating to "sustainability," "handover," and "lasting change" from 45 publicly available evaluation reports from 2020-2025. The sample spans a broad scope of projects, thematics, medical issues, geographies, emergencies, and non-emergencies. With support from AI, we analysed this data to specify the various domains in which MSF aims to contribute to lasting change. We identified five domains related to patients and communities; and seven domains related to systems, processes and sustainability. Cumulatively, they offer a conceptualization of Impact which retains elements of sustainability that MSF cares about, while jettisoning those it does not. The result is a distinctively MSF understanding.

## REFRAMING THE CONCEPT: FROM SUSTAINABILITY TO IMPACT

The table below elaborates the summary definition of Impact provided in COMPASS with the addition of 12 sub-criteria identified in the analysis of five years of evaluation reports. Those sub-criteria fall broadly into two categories: five related to long-term change for patients and communities; and seven related to systems, practices and sustainability.

Under patients and communities, sub-criteria examine whether lasting change is achieved after MSF's presence in terms of sustained health outcomes; access and continuity of care; patient and beneficiary well-being; dignity, trust and respectful care; and equity of lasting benefit.

Under systems, practices and sustainability, the sub-criteria examine whether lasting change is achieved after MSF's presence in terms of routinization of new practices; institutionalization and ownership of practices introduced by MSF; increased system resilience; development of knowledge products; advocacy influence; replicability and transferability of approaches or models developed by MSF; and handover quality and post-exit safeguards.

| **A: PEOPLE, PATIENTS AND COMMUNITIES**<br>***These criteria reflect the various ways in which MSF evaluation findings imply that MSF values whether benefits for patients or communities last beyond immediate project delivery.*** |
|---|
| • **A1 – Health outcomes -** Whether the intervention contributes to lasting improvements in health outcomes (e.g. morbidity, mortality, disease control), not just short-term results.<br>• **A2 – Access and continuity of care** - Whether people are able to continue accessing needed services over time, including after MSF reduces or withdraws support.<br>• **A3 – Well-being (psychosocial / protection)** - Whether improvements in psychosocial well-being, protection, or safety are maintained beyond the project period.<br>• **A4 – Dignity, trust, and respectful care** - Whether care practices that promote dignity, trust, and respectful treatment continue once MSF is no longer present.<br>• **A5 – Equity of lasting benefit** - Whether lasting benefits are shared equitably, particularly for vulnerable or marginalized groups, rather than only for easier to reach populations. |

**B: SYSTEMS, PRACTICES, AND SUSTAINABILITY**

***These criteria reflect the various ways in which MSF evaluation findings imply MSF values whether change is embedded in systems, institutions, or practices in ways that can persist without MSF.***

- **B1 – Practice change adoption (routinization) -** Whether new practices introduced by MSF are adopted into routine ways of working by staff or partners.
- **B2 – Institutionalization and ownership -** Whether responsibility for activities, systems, or practices is taken up by local institutions (e.g. Ministry of Health, local partners) rather than remaining dependent on MSF.
- **B3 – System resilience** - Whether the system is better able to absorb shocks, adapt, or continue functioning when facing stress, uncertainty, or crisis.
- **B4 – Knowledge products -** Whether MSF produces documented knowledge products such as guidelines, tools, protocols.
- **B5 – Advocacy influence -** Whether MSF's advocacy contributes to changes in policy, practice, or attention by external actors.
- **B6 – Replicability and transferability -** Whether approaches or models developed by MSF can realistically be applied elsewhere or scaled without MSF's direct involvement.
- **B7 – Exit / handover quality and post exit safeguards -** Whether MSF exit or handover processes are planned and managed in ways that protect continuity, reduce risks, and support sustainability.

## HOW DOES THIS HELP US?

The expanded definition of Impact offered above describes the various ways in which MSF projects seek to contribute toward lasting change within their contexts based on documented practice. Not every project aims to contribute toward every one of these sub-criteria. Together, however, the definition describes the breadth of MSF's aspirations to contribute towards lasting change. It provides a vocabulary adapted to MSF realities to describe the lasting changes we aim for with our projects, offers clarity to inform objective-setting and decision-making within and about programs, and provides a foundation for monitoring and evaluation.

While we have proposed to attach the sub-criteria emerging from this analysis to the Compass terminology of Impact. Ultimately the words we use are not as important as the definitions we give them. A similar analysis could have yielded an MSF definition of sustainability distinct from how it is understood more broadly in humanitarian and development circles, and that would be equally valid. What matters most is that the criteria we use are defined and adapted to capture our values and ambitions as an organization.

# Reflections on the iCCM Thematic Evaluation

In 2024, the Medical Department commissioned a thematic evaluation of Integrated Community Case Management (ICCM), covering a portfolio of OCB projects implementing this model. We spoke with Kemi Ogundipe, OCB Child Health and Paediatrics Adviser, who was closely involved in the evaluation. She shared her reflections on the process, the usefulness of the findings, and what could be improved in future thematic evaluations, which remain a relatively new area of work for the SEU.

The reports are accessible on the OCB intranet - SEU library.

### How did you experience this thematic evaluation, and what were your first impressions?

I appreciated how the evaluation helped build a clearer picture of an activity we've only recently begun implementing. Looking across several projects provided a more comprehensive understanding of the overall landscape, rather than a set of isolated pieces.

One aspect I really valued was how the evaluation immediately triggered reflection at project level. Simply asking certain questions led some teams to reassess their activities with a more critical or strategic lens. In some cases, this led to removing elements that were not actually part of ICCM; in others, it meant tightening referral gaps.

In addition, the evaluation helped create momentum within OCB, something I consider one of its most important contributions. It provided evidence that the needs are real, giving us a basis to advocate for resources. Without that, such priorities risk becoming unfunded mandates.

A meeting is already planned to discuss this, which is encouraging, as it creates an opportunity for political engagement and resources to finally come together.

### Did the evaluation bring anything new to you, or mostly confirm what you already knew?

Yes, it did bring something new. I know these projects quite well and visit them regularly, but having an external perspective was particularly valuable. It helped highlight where support was needed, both at the project level and more broadly.

I cannot be the only one identifying gaps, so having an independent evaluation confirm them adds weight. It also reinforced conversations that were already ongoing, which is important.

### How did you find the evaluators' engagement and approach?

I really appreciated the way the evaluators approached the work. When we clarified certain points, they listened carefully and adapted quickly.

This was not only reflected in the final report, but throughout the process. As we shared the realities on the ground, they adjusted their approach, revisited assumptions, and tailored their expectations to each context. That flexibility was essential. In other evaluations, this does not always happen, and findings can become disconnected from what we know to be true.

### Was there anything you found particularly helpful in their approach?

Yes — their use of an external ICCM benchmark as the evaluation framework (i.e. the WHO/UNICEF benchmarks for iCCM). ICCM isn't just a name; it's a defined package of services. Comparing MSF's activities against external standards helped clarify what ICCM actually includes, how it should be delivered, and what may be missing. That framing proved especially useful for guiding discussions.

What would you improve in future thematic evaluations?

For me, the main gap lies in how to operationalize the findings. The evaluation clearly identified the issues, but we need more support to understand what addressing them would look like in practice.

For example, if a recommendation involves dedicated staff, what does that mean in practice? How much time would it require? What would the role involve? What have others done successfully in similar contexts?

I can suggest ideas, but I'm not a programme organiser. It would be helpful to have a follow-up moment — once the findings have been absorbed — with people who have public health or programming expertise, to help us translate recommendations in a realistic way.

Do you have reflections on thematic evaluations specifically?

Yes — the evaluation question is crucial. For thematic evaluations, we need to ensure we are asking questions that actually require answers. If the findings are already well known, the exercise doesn't add much. That was my experience in another project evaluation I was involved in, where the results were not surprising because they confirmed what we already knew.

If you look back overall, how would you summarise your impression of the ICCM thematic evaluation?

It answered the questions we asked, for the most part. The evaluation was useful, and I appreciated the evaluators' openness throughout the process. For me, the key challenge now lies in how we translate these findings into action.

## An evaluator's Perspective - Insights from Emergency Evaluations

Paula Farias is a medical doctor with extensive operational experience in emergencies. Paula currently works as an independent consultant and has conducted several evaluations for the SEU over the past years, including the Maiduguri Nutritional Response in Nigeria, the Cyclone Idai response in Mozambique, and the emergency intervention in the Nuba Mountains, Sudan.

Having worked with the SEU on many occasions, you are familiar with our approach to evaluation. What do you think would be interesting to see in terms of evolution of our approach to evaluating emergency responses?

The SEU has begun to explore real-time assessments, which aligns with MSF's shift towards emergency response. A logical next step may be to rethink what they are and what they are for. I would start with the name. Referring to them as 'reviews' rather than 'real-time evaluations' (RTEs) is not merely a cosmetic change. RTEs still carry the methodological weight of formal evaluations: judgements, criteria, structured outputs, whereas a review suggests something more operational, closer to support than to judgement. This matters for those receiving it. Teams in the midst of an emergency have little reason to invest energy in facilitating an evaluation, but may engage with a process that helps them make sense of what they are doing, identify decisions made on the basis of unverified assumptions, and gain a coherent overview that is otherwise difficult to construct from within.

Framed this way, the primary audience becomes the team rather than headquarters.

Through an external reviewer, fragmented pieces are synthesized into a coherent narrative that helps teams better understand why certain decisions are having the consequences they are — while there's still room to act.

### How could the SEU support this in practice?

There is a structural challenge faced by any emergency evaluation: teams do not document for the sake of evaluators, but rather take action. As a result, reviews end up being based on partial reconstructions. The SEU could play a more active role in this regard, by defining what minimum information should be collected during an emergency and integrating it into existing processes: situation reports, coordination meetings and key decision-making moments. This is not about adding bureaucracy, but about anticipating what is needed to carry out a proper review later on. A simple decision log could be useful — not as a burden, but as something to help teams keep track of what has and has not been done.

And perhaps the strongest argument for this kind of review isn't methodological but political, in the most basic sense. Naming things removes the alibi of ambiguity. In emergencies, many decisions happen in the shadows — not because people are unaware of them, but because as long as they are not explicitly formulated, no one is obliged to respond. A review forces that formulation. Once an assumption is documented, continuing to act on it without verification becomes a conscious choice rather than inertia and this fundamentally changes the nature of institutional responsibility.

# From Evaluation to Learning: Advancing Adolescent Engagement in SRH

In 2025, the Mbare evaluation was presented at the MSF Paediatric Days. This is a brief summary of the evaluation and the presentation at the conference.

At OCB, project evaluations are primarily designed to generate learning that strengthens the specific project under review. This was the case for the 2024 evaluation of the Mbare Adolescent Sexual and Reproductive Health (ASRH) project in Zimbabwe, which aimed to inform a strategic roundtable and guide the project's future direction.

Yet evaluations often produce insights that extend beyond a single project. Recognizing this, colleagues involved in the Mbare evaluation - from project, coordination, and headquarters levels – jointly submitted an abstract to the MSF Paediatric Days Conference. This two-day event brings together MSF staff, researchers, and partners to advance paediatric knowledge and practice in humanitarian settings. The 2025 edition focused on paediatric care in conflict settings and on mother–childcare.

The submitted abstract was selected for an oral presentation — an opportunity to spotlight adolescent health, which is often overlooked in paediatric and maternal health discussions. Through this presentation, we aimed to convey a clear message: ***Involving adolescents and young adults (AYA) in designing, implementing, monitoring, and evaluating health interventions targeted to them is both feasible and highly beneficial.***

Dr Chido Dziva Chikwari, Assistant Professor at the London School of Hygiene & Tropical Medicine and one of the two external evaluators, presented on behalf of the team. She highlighted how meaningful adolescent engagement improves the design, delivery, and measurement of sexual and reproductive health (SRH) programs. Adolescents and young adults can effectively act as researchers and implementers when properly trained and supported. Peer educators, who are themselves AYA and often the first point of contact, play a pivotal bridging role. But they also face challenges such as power imbalances, risks of verbal or physical violence, heavy data collection burdens, and insufficient organizational support.

The evaluation found that AYA involvement moved interventions closer to the realities of adolescent lives. When adolescents co-design and co-lead services, programs become more relevant, acceptable, and trusted. However, this requires time, consistency, and a willingness from teams to adapt. Sustained AYA engagement must also account for sociocultural norms, stigma, and intersecting vulnerabilities — especially for groups such as LGBTQI youth, adolescents who use substances, or those living on the streets. A dignified, adolescent-friendly approach is thus essential for long term impact.

Macdonald Zenda, one of the young research assistants involved in the evaluation, reflected: **"***Being a young researcher during the evaluation felt like opening a door that had long been closed. For years we were the ones being reached out to, but this experience turned the tables. What surprised me was how peers spoke to us compared to how they speak to clinical staff. There is a level of trust and lack of stigma that only happens between those who live in the same community or street. I realized that making a clinic truly adolescent friendly is not about free Wi-Fi, posters on the wall or games, but about willingness to keep adjusting based on the actual lives. It's not just about ticking the boxes for reporting but ensuring the programs are trusted and relevant."*

Watch the presentation online (Day 2, from 5:52:20). Access the submitted abstract and conference poster.

# Engaged in a Project… Engaged in the Evaluation?

The SEU managed the evaluation of a project aimed at learning how to better work with and strengthen CBOs. This piece is a summary of a discussion that three colleagues involved in the process had to assess whether they were able to engage the CBOs themselves in the evaluation.

In 2025, MSF Southern Africa asked the SEU to manage an evaluation of its TIC project implemented to work with community-based organizations (CBOs). With community members central to the project, a critical question emerged: how can those communities be meaningfully involved in evaluating it? Three colleagues — Donald Zhou (MSF SnA project manager), Emmeline Kerkvliet (SEU evaluation officer), and Silva Ferretti (external evaluator) — reflected honestly on how that question played out, what got in the way, and what they would do differently. This is a summary of their discussion.

## The intention was there from the start

From the outset, there was a clear intention to include CBO voices in the evaluation. As Donald put it, the logic was straightforward: if the project aimed to support CBOs, they should have a say in evaluating it. A CBO representative was identified early for the consultation group.

For Donald, community engagement had never been an afterthought. The project was about learning what meaningful partnership with communities could look like for an organization more used to delivering services than building lasting relationships. Hearing communities reflect on whether the capacity-building was actually helping them was central to that mission, not just useful evaluation data.

But it was when Silva came on board and raised the possibility of using the evaluation process itself to serve the project's aims — not just assess them — that the team began thinking more expansively about what community participation could actually look like.

## What actually happened — and why

In practice, community engagement during the evaluation was real, but uneven, largely because of the way the project was planned and implemented. The project was in the end a series of smaller projects, each with its own context and intention. It was necessary to frame engagement in the context of each project, which in turn affected how the evaluation could frame engagement. Without such contextualization it would have been a very shallow process – yet it proved difficult.

Silva conducted a questionnaire across organizations and attempted a social network analysis (SNA) to assess whether a community of practice had emerged. In principle, the use of a SNA could have added significant value. There were hopes it could engage the communities more actively and reveal how learning and support were shared, potentially mapping interesting interactions for further exploration. In practice, however, the linkages were not particularly strong, leaving little to pursue.

In South Africa and Mozambique, Silva met directly with CBOs, but the approaches differed considerably. In South Africa, where the project had already closed, she visited separately three organizations that barely knew each other. In Mozambique, however, she was embedded alongside an active project team, which allowed for richer, more organic conversations with communities, peer support workers, and partners. Being able to have time, proximity and space proved essential.

Reflecting on past evaluations, when this was possible, Silva emphasized that in such cases, she could observe who the CBOs were, how they operated, and what the broader ecosystem around them looked like rather than arriving with a checklist. In one particularly memorable case (Kiambu), she spent extended time in a clinic shared with a local organization. As a result, what started as scheduled

interviews evolved into an ongoing exchange – and more meaningful engagement.

But not every site affords this kind of depth. In some cases, Silva only had a few hours, just enough to arrive, ask questions, and leave, leaving little space for genuine exchange. In others, the team made a deliberate choice not to engage directly with certain community members, judging that a single shallow interview would do more harm than good.

### What would be different next time

Reflecting on the process, all three identified a core tension: like the project itself, the evaluation was framed primarily as an organizational learning exercise for MSF. Communities participated, were heard and shaped the findings. Yet, the process was ultimately designed around what MSF needed to know, not around what the communities themselves might gain from it.

In retrospect, Silva reflected that it would have been very interesting to apply social network mapping to actors within individual projects (rather than the overarching CBO strengthening project). This would have clarified the different ecosystems, making them more visible and better framing things from the perspective of CBO members.

Emmeline suggested that a phased approach could have helped: a first visit to understand the landscape, followed by a more deliberate return focused on deeper engagement, informed by what had been learned. Silva pointed to an idea outlined in the inception report but never implemented: a peer exchange model, in which CBOs from different projects would visit one another, with the evaluator acting as facilitator and documenter rather than the central figure. The evaluation would have happened, but through a lens of mutual learning rather than external assessment. As she noted, drawing on similar experiences, "facilitating peer exposure can be better than evaluation, because peers come with curiosity and engage more deeply with operational tensions and challenges."

Donald offered perhaps the most pointed question of the conversation: how do you make an evaluation theirs, not yours? Not an exercise MSF designs, deploys, and extracts learning from; but something the communities themselves recognize as useful, even after the organization has moved on.

### The bits belong to the people

The conversation ended with an insight that brought the whole discussion into focus. In reflecting on how to navigate the power imbalance between an external evaluator and the communities she is engaging with, Silva described what she tries to do: observe, document, connect the dots, and give the picture back. "You have the power of putting things together, but the bits really belong to the people."

That, perhaps, is the most transferable lesson from this experience. Meaningful community participation in evaluation is not about consulting communities at the margins of a process designed elsewhere. It begins with asking whose knowledge is being collected, who it serves, and whether those whose experiences form the evidence can use what emerges.

In this way, evaluation becomes not only a means for MSF to understand change, but an opportunity for CBOs and MSF to reflect together throughout the process and/or to make explicit the reflections they already made.

The evaluation reports and webinar recording are accessible on the OCB intranet – SEU library.

# Reflections from our AI bot

We, or rather our AI bot, sat down with our 2025 evaluations for a quick chat—asking top-of-mind questions and getting straight to the point. This isn't a rigorous analysis, but a rapid pulse check. Why share it? Because it is grounded in solid evidence and even a light AI read can spark deeper thinking.

Q1. So… how are we actually doing? (Because I haven't had time to read anything longer than a WhatsApp message.)

Pretty well where it counts — reaching people others can't. But we keep stumbling on the same things: staying true to our mandate, data we can't fully trust, and staff turnover that eats institutional memory for breakfast.

Q2. Are we still acting according to our principles, or did we accidentally drift into "well… it was complicated" territory?

Somewhere in between. Our values are intact, but we have real examples of projects quietly expanding beyond what they were meant to do — often without anyone formally deciding to change course. Good intentions + no written strategy = drift.

Q3. What are the biggest things that changed around us, and did we react like pros or like "oh wow, didn't see that coming"?

Conflict, displacement, regulatory shifts, worsening access — all accelerating. When we had systems in place, we adapted fast. When we didn't, we improvised and then scrambled. Half pro, half pigeon-on-a-chess-board.

Q4. Where are we getting the most impact for our sweat, tears, and budget — and where are we basically pushing a rope uphill?

Mobile and outreach services reaching isolated or marginalised communities: genuinely high impact. Emergency response when logistics work: excellent. Long-term change projects in stable contexts without a clear exit plan: that's where the rope starts.

Q5. If patients could evaluate us, what would they praise, complain about, or politely side-eye?

Praise: free care, showing up where no one else does, genuine kindness. Complain: continuity gaps, referrals that lead nowhere, distance. Side-eye: decisions made about them without them, and sometimes not knowing we were leaving until we were gone.

Q6. What operational headaches kept popping up across missions — logistics, HR, supply, coordination — and which ones are fixable without a miracle?

High staff turnover, supply delays, fragmented data systems, slow internal approvals. The fixable ones: better inductions, pre-authorised response packages, and cleaner data flows. The hard ones require a structural conversation, not a workshop.

Q7. Which recommendations keep showing up over and over, like that one mosquito that refuses to leave the room?

Exit/transition planning before we're already packing. Data systems that talk to each other. Engaging communities and governments earlier in design. Documenting strategy before the fourth coordinator rotation. These are perennial.

Q8. Is there anything new, innovative, or surprisingly clever happening out there that we should steal — uh, replicate?

Community-based organisations being mentored as genuine partners, not handover recipients. Pre-validated emergency packages cutting response time. Peer educators sustaining health-seeking behaviour after we leave. All replicable.

Q9. What risks are coming our way in the next few years, and how worried should I be on a scale from "meh" to "better start prepping"?

Shrinking humanitarian space, access restrictions, over-specialisation in specific crisis types, and communities left mid-care when funding ends. Not existential — but they compound fast without active mitigation. Somewhere around “start prepping, but calmly.”

Q10. If I only have 30 seconds before my next call, what are the three things I absolutely need to know from all these evaluations?

One: We reach people no one else does — protect that. Two: Strategic drift happens silently; document your mandate early. Three: Communities and national partners are assets, not afterthoughts — bring them in sooner.

*Based on evaluations completed in 2025. Three evaluations remain unpublished; findings from these are reflected without direct attribution. Questions or reflections? Reach out to the Stockholm Evaluation Unit.*

Stockholm Evaluation Unit
http://evaluation.msf.org
Médecins Sans Frontières