

Comparative study of WHO evaluation function with selected UN entities

Executive Summary



Acknowledgements

This report is issued by the WHO Evaluation Office. It is based on the independent evaluation conducted by the Evaluation Team from Euro Health Group comprising Clare Dickinson (Team Leader), Maiken Mansfeld Jacobsen (Deputy Team Leader), Erin Ferenchick, Matthew Cooper and Finja Daegling.

This evaluation was managed and quality controlled by Marie Bombin, Senior Evaluation Officer, with contributions from Ye Li, WHO Evaluation Office.

Cover photo credit: WHO

Acronyms

ADG	Assistant Director-General	IEO	Independent Evaluation Office (UNDP)
AFRO	Africa Regional Office	ILO	International Labour Organization
AIDA	Artificial Intelligence for Development Analytics	IOS	Internal Oversight Services
AMSTG	Agile Member States Task Group	JIU	UN Joint Inspection Unit
CEQAS	Centralized Evaluation Quality Assurance System (WFP)	JPO	Junior Professional Officer
CPD	Country Programme Document	LTA	Long Term Agreement
CPE	Country Programme Evaluation	MOPAN	Multilateral Performance Assessment Network
CRM	Compliance and Risk Management and Ethics	OECD	Organisation for Economic Co-operation and Development
CSP	Country Strategy and Programme	OED	FAO's Office of Evaluation
DAC	Development Assistance Committee	OEV	Office of Evaluation
DEQAS	Decentralized Evaluation Quality Assurance System (WFP)	PAHO	Pan-American Health Organization
DWCP	Decent Work Country Programme (ILO)	PBAC	Programme Budget and Administration Committee (WHO Executive Board)
DE	Decentralized evaluation	PHC	Primary Health Care
EB	Executive Board	PSA	Planning Resource Coordination and Per Monitoring
ECD	Evaluation Capacity Development	QA	Quality Assurance
EISI	Evidence Information Systems Integration	RO/	Regional Office
EMCP	Evaluation Manager Certification Programme	SDG	Sustainable Development Goals
EQAA	Evaluation Quality Assurance and Assessment system (UNFPA)	SEARO	Southeast Asia Regional Office
EQAS	Evaluation Quality Assurance System (WFP)	UN	United Nations
ERC	Evaluation Resource Centre (UNDP)	UNDAF	United Nations Development Effectiveness Framework
ERG	Evaluation Reference Group	UNDP	United Nations Development Programme
EURO	European Europe Regional Office	UNEG	United Nations Evaluation Group
EVL	Evaluation and Organizational Learning (WHO)	UNESCO	United Nations Educational, Scientific and Cultural Organization
EVO	Evaluation Office	UNFPA	United Nations Population Fund
FAO	Food and Agriculture Organization	UNHCR	United Nations High Commissioner for Refugees
FENSA	Framework for Engaging with Non-State Actors	UNICEF	United Nations Children's Fund
GEHR	Gender Equity and Human Rights	UNSWAP	UN system-wide Action Plan
GEROS	Global Evaluation Reports Oversight System	USD	United States Dollars
GNE	The Global Network on Evaluation	WHA	World Health Assembly
GPW13 / 14	Thirteenth / Fourteenth General Programme of Work	WHE	World Health Emergency
HLE	High Level Evaluation	WHO	World Health Organization
IDP	Internally displaced person Displaced Person	WPRO	Western Pacific Regional Office
IEE	Independent Evaluation of the Evaluation Function		

Executive Summary

This study has compared the WHO evaluation function, both at global and regional level, with eight UN agencies (FAO, ILO, UNDP, UNESCO, UNFPA, UNHCR, UNICEF and WFP) that were selected purposively based on comparability to WHO in size or function, or as agencies recognized (e.g. by UNEG Peer Review/MOPAN) to have strong and mature evaluation functions.

The objectives of the study are to:

- Identify good practices applicable to WHO in terms governance, coverage and resources which contribute to ensuring 1) independence, 2) credibility, and 3) use.
- Compare current policy & systems, practice and resourcing of WHO evaluation function with selected UN entities' organizational setting, mechanisms, processes and tools.
- Make recommendations on the three dimensions (policy & systems, practice, and resourcing) to Member States and the Secretariat to enhance the ability of the WHO evaluation function to stimulate learning and promote accountability, transparency, and effectiveness through independent, credible and useful evaluations.

Whilst there is no direct comparator to WHO in structure, size or mandate, this study has sought to identify lessons learnt and best practices from comparator agencies which may inform WHO's approach to strengthening its evaluation function going forward.

The below table outlines some of the key features of WHO's evaluation function and that of comparator agencies, showing it (alongside UNHCR) has the lowest financial resources as a percentage of organisational budget, and a relatively low number of staff and volume of evaluations.

Table 1: Key Features of Comparator Evaluation Functions

Organization	Evaluation Budget (million USD)	Evaluation spend as a percentage of organization budget	Volume	Number of evaluation staff
WHO	\$7.7	0.1%	36 (11 corporate, 25 decentralized) (2022-23)	12
FAO	\$10.2	0.8%	103 (2021-22)	20
ILO	\$7.3	0.8%	72 (2 Corporate, 70 Decentralized) (2022)	12
UNDP	\$50	0.77%	416 (20 corporate, 396 decentralized) (2022)	35
UNESCO	\$1.1	3%	36 (5 corporate and 31 decentralized) (2023)	6
UNFPA	\$9	0.83%	23 (2022)	12
UNHCR	\$6.5	0.1%	26 (12 corporate and 14 decentralized) (2023)	13
UNICEF	\$65	1%	199 (20 corporate, 179 decentralized) (2023)	40
WFP	\$32.85	0.31	54 (27 corporate and 27 decentralized) (2022)	86

Key points are:

- A well-functioning evaluation system is integral for WHO to achieve its strategic goals, including those of GPW13 and GPW14 and its role in relation to the health-related SDGs and health emergencies; and to function as a knowledge-based, learning organization with strong accountabilities.
- Over the last 10 years (since the 2014 JIU report), the 8 entities considered for this study have made substantial investments in their evaluation functions, in relation to evaluation policies, systems, practices and resourcing.
- Meanwhile WHO's evaluation function has also evolved in key aspects since the creation of a separate evaluation unit in 2014, and after a pause during the COVID-19 pandemic, there is now more recently additional impetus which has come with new leadership of the function. Strengths include delivery of high quality, relevant centralized evaluation helped by:
 - Bringing in evaluation specialists with relevant professional expertise from other agencies into EVL.
 - Developing more detailed evaluation guidance in key areas.
 - Working with regional leads by relaunching the global evaluation network and developing the partnership between the different levels of WHO's evaluators (global, regional, country).
 - Planning evaluations with the regions and moving to an LTA for procurement of support from evaluation providers.
 - Recent progress at regional level in developing capacity, although this has a long way to go both at regional and country level.
 - Creating a roster of evaluation specialists.
 - Developing a costed evaluation workplan.
- However, despite this good work, in many key respects (such as decentralized evaluation, coverage, financial and human resourcing, and evaluation use) there are still significant gaps in WHO's evaluation function at different levels.

In summary, WHO has not in recent years been able to keep pace with the steadily advancing standards for evaluation among comparable UN organizations. The recommendations in this study and the examples of best practice from the comparators are intended to suggest how this can be remedied. There is substantial scope to learn from the innovations that have been introduced in the 8 comparators, see the overview of their functions in the main body of this report, and the agency summaries in Annex 3.

Policy

WHO's 2018 evaluation policy and accompanying implementation frameworks have some of the key elements required for a strong enabling environment for evaluation, but there are also many gaps and awareness of the policy requirements is highly variable. All the comparators have reviewed and updated their policies, informed by UNEG peer reviews or independent evaluations; and then made progress in key areas such as coverage norms, use of evaluation and resourcing. The ILO approach is one interesting example of how results against evaluation policy norms are tracked in detail. The WHO 2018 policy needs updating and fleshing out in line with UNEG norms and standards and best practice in comparators, with a clear roadmap for how it will be delivered.

Recommendations

R1. Update the 2018 evaluation policy, addressing the gaps noted in this study. Keep it regularly updated, informed by UNEG peer reviews/independent evaluation, every 5 years.

R2. Develop explicit coverage norms in key areas of WHO work, and track progress annually, relating them to strategic outcomes for WHO. These should include:

- Balanced coverage of WHO's intended strategic outcomes based on the SDG3 goals, GPW14, key WHA resolutions and policies.
- 100% coverage (through inter-agency evaluations) of global health emergencies and level 3 humanitarian crises.
- 100% coverage of all country strategies in the year prior to their revision, for those countries with off track health SDGs and/or high levels of risk identified in programming and audits. Coverage at least once every 2 programme cycles for all other country strategies.

R3. Set out a fully developed roadmap on how the evaluation policy will be delivered, accompanied by a detailed results reporting arrangement including key performance indicators around quality, credibility, capability, resourcing and use.

Systems and processes

WHO's evaluation systems also include some of the elements one would expect to see, including a consultative approach to work planning with a newly developed costed workplan, guidance and support from EVL, reporting to the governing body via PBAC, management response systems etc. Again, however, there is a need to take this to the next stage and strengthen systems -this is important particularly for decentralized evaluation¹.

Recommendations

R4. Further develop and embed an institutional approach to work planning (including with WHE) and consulting on selecting evaluation topics, informed by the coverage norms as per R2.

R5. By agreement with Management and the Governing body as appropriate, set out an explicit budget line for evaluation so that the costed workplan and the supporting function can be delivered; and to underpin structural independence of the function.

R6. Strengthen evaluation oversight and visibility by:

- informal sessions with Member States where EVL presents the evaluation report and management present their response².
- an internal evaluation committee in which WHO senior management discusses and approves the evaluation and the management response.

¹ An influential JIU report (JIU/REP/2014/6) on evaluation functions looked at their maturity in the UN systems and covered many aspects including a finding that policies, norms and standards for decentralized evaluation were inadequate. Since then, all the comparator agencies considered here have strengthened their decentralized evaluation functions. WHO's decentralized structure and focus on results at the level of member states means that progress on this aspect is particularly important.

www.unjiu.org/sites/www.unjiu.org/files/jiu_document_files/products/en/reportsnotes/JIU%20Products/JIU_REP_2014_6_English.pdf

² The internal evaluation committee should also receive annual reports from EVL on progress against results outlined in the evaluation roadmap (i.e. evaluation planning, implementation, resourcing, quality, evaluation culture) and ensure necessary actions are required to stay on track.

R7. Monitor the coverage and quality of decentralized evaluations and take steps to map the existing coverage more reliably and enhance the quality of DE over time.

- This should be supported through a well-developed approach to training and capacity building learning from the approach adopted in ILO, UNDP, WFP and others.
- Ensure that specific guidance is available to inform quality of humanitarian and evaluation of emergencies.

R8. Consolidate existing elements such as QA checklists and templates already developed by EVL, into a fully developed QA system, with annual reporting including ex-post QA scoring for centralized and decentralized evaluations (including WHE evaluations). A good model here is the approach developed in WFP. A staged approach is recommended, starting with a system at central level closely followed within 2 years by expanding to cover decentralized evaluations.

Financial and human resources

The resourcing of the evaluation function is not clear and is significantly below what is needed given the size, scope and complexity of WHO's work. There have been sharp increases in resourcing for evaluation in many of the comparators, but not so in WHO. A first essential step is (as per recommendation 5 above) to have an explicit budget line for evaluation. The second key step is to identify and implement a target level of resourcing overall. The 2018 policy includes a reference to the norm of 0.5% to 1.0% established by the JIU report, but this has never been implemented and resources are closer to 0.1% currently.

The human resources at the level of EVL and capacity at regional and country level to support evaluation are and order of magnitude below what is needed for a well-functioning evaluation function in an organization like WHO. This requires additional staffing both in EVL and at regional and country level (see R12 below).

Recommendations

R9. Ensure that there are dedicated 'draw down' resources available for humanitarian and evaluation of emergency response, articulated in the costed workplan. This should include identifying an agreed percentage (e.g. 1%) of voluntary contributions for emergencies that is dedicated to evaluating the emergency response.

R10. Commit to a target level of resourcing as a percentage of WHO expenditure. The target should be set out in the revised policy at the level of 1.0%, in line with the norm across the UN, and then related to progress against and resources for delivery of coverage norms³.

³ This should also be accompanied by a clear and explicit statement in the revised evaluation policy of responsibilities of Management and Member States to ensure that necessary resources are made available for evaluation to meet the agreed standards.

Decentralized evaluation

The decentralized evaluation function in WHO is significantly under-developed. This is at odds with WHO's overall strategic focus, which includes a decentralized governance structure, strong emphasis on its regional and country level work and assessing impact at the level of member states in relation to the triple billions.

Recommendations

R11. Regional directors, with advice from the director of EVL, should be asked to develop a clear strategy for WHO to invest in the decentralized evaluation function to bring it up to the standards of comparable UN organizations and address issues highlighted 10 years ago by JIU. Build a common strategic approach for evaluation at all 3 levels, using the helpful collaborative approach already established in the Global Network on Evaluation (GNE) as an entry point⁴.

R12. Build capability at regional and country level on evaluation, including regional evaluation units that are staffed at the appropriate level (P4 level as a minimum) with reporting arrangements to regional directors and with a 'dotted line' reporting to the director of EVL on professional and technical aspects. Develop a network of country level focal points for M&E in the larger WHO offices, who can also support smaller countries as required.

Demand for evaluation – evaluation culture / demonstrating value

Strengthening WHO's evaluation function cannot only be done through supply side measures to build capacity: attention to the demand side and culture of evaluation across the organization is at least as important.

Recommendations

R13. As part of revising and launching the new Evaluation Policy, take steps to build awareness of evaluation through a major communications and training initiative led by EVL and championed by senior managers.

- Through this process communicate a clear vision of evaluation across the organization, how it adds value and how evaluations are intended to be used.
- Spell out the accountabilities set out in the evaluation policy so that management and staff are aware of what is expected and why.

R13b. Ensure there is a high-level forum for discussing evaluations in depth.

- See the recommendation R5 above, which would put in place informal sessions with member states a high-level internal Evaluation Committee with representatives from senior management, chaired at ADG level.
- This would need appropriate representation at regional and country level, who can take delivery of and discuss (decentralized) evaluations, findings and recommendations.

R13c. Build demand, support and buy-in for evaluation and maximum use.

⁴ This must not be a top-down approach, given WHO's governance structures and culture.

- This needs to start with messages communicated by the governing body and senior management on the importance of evaluation to provide a platform for demand.
- It would then be reinforced by closing the learning loop through consistent production of high quality, credible and useful reports which demonstrate what value can come from the evaluation function.

Use, impact and follow up

Ultimately the test of the quality of WHO's evaluation function is the extent to which evaluations are used and inform decision making. Partly because EVL's resources are so stretched in delivering evaluations, there is limited capacity at present for ensuring use and follow up. Best practice in other entities is to have a strategic and innovative approach to disseminating evaluations accompanied by strong incentives for action.

Recommendations

R14. Modernize and deepen EVL's approach to dissemination of evaluations, using a wider range of tools and formats as part of a strategic investment in making evaluations accessible. Starting at the top, this should include the regular informal sessions with EB representatives of member states.

R15. Strengthen the systems and incentives for follow up on evaluation recommendations by clearly defining the responsibility of senior managers in this respect, moving the focus from EVL to operations and building this into performance management and compliance.

- The internal evaluation committee proposed in R6 would provide an important focus within the organization on this, reinforced by greater interest from MS. This in turn would help in creating incentives for staff to focus on the management response actions following each evaluation.
- Staff should receive guidance and training on what is expected around management response and follow up to evaluations.

Any enquiries about this evaluation should be addressed to:
Evaluation Office, World Health Organization
Email: evaluation@who.int
Website: Evaluation (who.int)