Canadian International Development Agency | Agence canadienne de développement international

CIDA ACDI

Development Effectiveness Review of the World Health Organization

2007–2010

FINAL REPORT

DECEMBER 2012

Canada

Canadian International Development Agency (CIDA)
200, Promenade du Portage
Gatineau, Québec K1A 0G4
Canada

Telephone:
819-997-5006 / 1-800-230-6349 (toll-free)
For the hearing- and speech-impaired:
819-953-5023 / 1-800-331-5018 (toll-free)

Fax: 819-953-6088
Website: www.cida.gc.ca/evaluations-e
E-mail: info@acdi-cida.gc.ca

Cette publication est aussi disponible en français sous Revue de l'efficacité du développement de l'Organisation mondiale de la Santé

Printed in Canada

Development Effectiveness Review of the World Health Organization

2007–2010

FINAL REPORT

DECEMBER 2012

Acknowledgments

CIDA's Evaluation Division wishes to thank all who contributed to this review for their valued input, their constant and generous support, and their patience.

Our thanks go first to the independent team from the firm, Goss Gilroy Inc., made up of team leader Ted Freeman, and analysts Danielle Hoegy and Tasha Truant. We are also grateful for the support of the Department for International Development of the United Kingdom and the Swedish Agency for Development Evaluation, which provided analytical support during the reviews of the World Health Organization (WHO) and the Asian Development Bank.

The Evaluation Division would also like to thank the management team of CIDA's Global Initiative Directorate (Multilateral and Global Programs Branch) at Headquarters in Gatineau for its valuable support.

Our thanks also go to the representatives of the WHO for their helpfulness and their useful, practical advice to the evaluators.

From CIDA's Evaluation Division, we wish to thank Vivek Prakash, Evaluation Officer, for his assistance with the review. We also thank Michelle Guertin, CIDA Evaluation Manager, for guiding this review to completion and for her contribution to the report.

Caroline Leclerc
Director General
Strategic Planning, Performance and Evaluation Directorate

Table of Contents

Figures

Tables

List of Abbreviations

CAH	Child and Adolescent Health and Development
CIDA	Canadian International Development Agency
DAC-EVALNET	Network for Development Evaluation of the Development Assistance Committee
EPI	Expanded Programme on Immunization
HAC	Health Action in Crises
MDG	Millennium Development Goals
MO	Multilateral Organization
MOPAN	Multilateral Organization Performance Assessment Network
NGO	Non-Governmental Organization
OECD	Organisation for Economic Co-operation and Development
OIOS	Office of Internal Oversight Services
RBM	Results-Based Management
UN	United Nations
WHO	World Health Organization
USD	United States Dollars
VPD	Vaccine-preventable diseases

Executive Summary

Background

This report presents the results of a development effectiveness review of the World Health Organization (WHO). Founded in 1948, the WHO is the directing and coordinating authority on international health within the United Nations System with the overall goal of achieving the highest level of health for all. It does not directly provide health services, but instead coordinates global health-related efforts and establishes global health norms. The WHO employs over 8,000 public health experts, including doctors, epidemiologists, scientists, managers, administrators and other professionals. These health experts work in 147 country offices, six regional offices and at the headquarters in Geneva.

While poverty reduction is not the primary focus of the WHO's mandate, it does contribute to poverty reduction through its global leadership—for example, establishing global health standards and norms which are used by developing countries and by supporting humanitarian coordination—and through its technical assistance in developing countries.

Health Canada has the overall substantive lead for the Government of Canada's engagement with the WHO, and is head of the Canadian delegation to the World Health Assembly. The Canadian International Development Agency (CIDA)'s main engagements with the WHO include policy dialogue and development assistance programming in infectious diseases, child health, and humanitarian assistance. More specifically, the WHO also plays a key role in developing health indicators and data collection in support of the G8 Initiative on Maternal, Newborn and Child Health championed by Canada.

With 284 million Canadian dollars of CIDA support in the four fiscal years from 2007–2008 to 2010–2011, the WHO ranks eighth among multilateral organizations supported by CIDA in dollar terms. In the area of health, only the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) ranks higher with 450 million Canadian dollars of support from CIDA in the same period.

Purpose

The review is intended to provide an independent, evidence-based assessment of the development effectiveness (hereafter referred to as effectiveness) of WHO programs to satisfy evaluation requirements established by the Government of Canada's *Policy on Evaluation* and to provide the CIDA's Multilateral and Global Programs Branch with evidence on the development effectiveness of the WHO.

Approach and Methodology

The approach and methodology for this review was developed under the guidance of the Organisation for Economic Co-operation and Development (OECD)'s Development Assistance Committee (DAC) Network on Development Evaluation (DAC-EVALNET). Two pilot tests, on

the WHO and the Asian Development Bank, were conducted in 2010 during the development phase of the common approach and methodology. The report relies, therefore, on the pilot test analysis of evaluation reports published by the WHO's Office of Internal Oversight Services (OIOS), supplemented with a review of WHO and CIDA corporate documents, and consultation with the CIDA manager responsible for managing relations with the WHO.

The methodology does not rely on a particular definition of (development) effectiveness. The Management Group and the Task Team that were created by the DAC-EVALNET to develop the methodology had previously considered whether an explicit definition was needed. In the absence of an agreed-upon definition, the methodology focuses on some of the essential characteristics of developmentally effective multilateral organization programming, as described below:

1. **Relevance of interventions:** Programming activities and outputs are relevant to the needs of the target group and its members;
2. **Achievement of Development Objectives and Expected Results:** The programming contributes to the achievement of development objectives and expected results at the national and local level in developing countries (including positive impacts for target group members);
3. **Sustainability of Results/Benefits:** The benefits experienced by target group members and the results achieved are sustainable in the future;
4. **Efficiency:** The programming is delivered in a timely and cost-efficient manner;
5. **Crosscutting Themes (Environmental Sustainability and Gender Equality):** The programming is inclusive in that it would support gender equality and would be environmentally sustainable (thereby not compromising the development prospects in the future); and
6. **Using Evaluation and Monitoring to Improve Effectiveness:** The programming enables effective development by allowing participating and supporting organizations to learn from experience and uses performance management and accountability tools, such as evaluation and monitoring, to improve effectiveness over time.

Based on the above-mentioned characteristics, the review's methodology uses a common set of assessment criteria derived from the DAC's existing evaluation criteria (Annex 1). The overall approach and methodology[1] was endorsed by the members of the DAC-EVALNET in June 2011 as an acceptable approach for assessing the development effectiveness of multilateral organizations.

The review involved a structured meta-synthesis of a sample of 25 WHO evaluations completed between 2007 and 2010, at the country, regional and global/thematic level. The sampling process is described further in Annex 3. The limited number of available evaluation reports represents a limitation for this report, as discussed below.

[1] *Assessing the Development Effectiveness of Multilateral Organizations*, DAC Network on Development Evaluation Guidance Document, 2012

After being screened for quality (Annex 4 describes approach and criteria used), each evaluation was reviewed to identify findings relating to the six main criteria (and associated sub-criteria described in Annex 1) to assess effectiveness. The review team classified findings for each criterion using a four-point scale as “highly satisfactory,” “satisfactory,” “unsatisfactory” or “highly unsatisfactory.” Classification of findings was guided by a grid with specific instructions for each rating across all sub-criteria (Annex 5). The review team also identified factors contributing to or detracting from results.

Note that although no evaluations were screened out due to quality concerns, evaluations do not address all the criteria identified as essential elements for effective development. Therefore, this review examines the data available on each criterion before presenting results, and does not present results for some criteria.

The percentages shown in this report are based on the total number of evaluations that addressed the sub-criteria. However, coverage of the different sub-criteria in the evaluations reviewed varies from strong to weak. Cautionary notes are provided in the report when coverage warrants it.

In addition to the 25 evaluations, the review examined relevant WHO policy and reporting documents such as the reports of the Programme, Budget and Administration Committee to the Executive Board, Reports on WHO Reform by the Director-General, Evaluation Policy Documents, Annual Reports and the Interim Assessment of the Medium-Term Strategic Plan (see Annex 6). These documents allowed the review team to assess the ongoing evolution of evaluation and results reporting at the WHO and to put in context the findings reported in the evaluation reports.

The review team also carried out an interview with OIOS staff at the WHO to understand better the universe of available WHO evaluation reports and to put in context the changing situation of the evaluation function. Finally, the review team interviewed the CIDA manager most directly responsible for the ongoing relationship between CIDA and the WHO in order to better assess the WHO’s contribution to Canada’s international development priorities.

As with any meta-synthesis, there are methodological challenges that limit the findings. For this review, the most important limitation concerns the generalization of this review’s results to all of the WHO’s programming. The set of available and valid evaluation reports does not provide, on balance, enough coverage of WHO programs and activities in the period to allow for generalization of the results to the WHO’s programming as a whole.[2] The available evaluation reports do, however, provide insights into the development effectiveness of evaluated WHO programs.

[2] WHO evaluation reports often do not include data on the overall value of the programs under evaluation. Therefore, it is quite difficult to estimate the level of evaluation coverage for the purpose of this report.

Key Findings

Insufficient evidence available to make conclusions about the World Health Organization

The major finding of this review is that the limited set of available and valid evaluation reports means that there is not enough information to draw conclusions about the WHO's development effectiveness.

The limited number of evaluation reports that are available provide some insights into the effectiveness of those WHO programs. Results from the review of these evaluations are presented below, but cannot be generalized to the organization as a whole.

An analysis of the 2012 WHO evaluation policy indicates that while the approval of an evaluation policy represents a positive step, gaps remain in the policy regarding the planning, prioritizing, budgeting and disclosure of WHO evaluations. In addition, the WHO could further clarify the roles and responsibilities of program managers regarding evaluations, and provide guidance to judge the quality of evaluations.

A 2012 United Nations Joint Inspection Unit review also raises concerns about independence and credibility of WHO evaluations. It suggests that the WHO should have a stronger central evaluation capacity, and recommends that a peer review on the evaluation function be conducted by the United Nations Evaluation Group and be presented to the WHO Executive Board by 2014.

Based on the limited sample available, WHO programs appear to be relevant to stakeholder needs and national priorities. Evaluations reported that WHO programs are well-suited to the needs of stakeholders, with 89% of evaluations (16 of 18 evaluations which address this criteria) reporting satisfactory or highly satisfactory findings, and well aligned with national development goals (100% of 12 evaluations which address this criteria were rated satisfactory or highly satisfactory). Further, the objectives of WHO-supported projects and programs remain valid over time (100% of 21 evaluations rated satisfactory or better). There is room, however, for better description of the scale of WHO program activities in relation to their objectives (60% of 20 evaluations rated satisfactory) and for more effective partnerships with governments (61% of 18 evaluations rated satisfactory or highly satisfactory).

One factor contributing to the relevance of WHO programs is the organization's experience in matching program design to the burden of disease in partner countries. Another is consultations with key stakeholders at national and local levels during program design.

The WHO appears to be effective in achieving most of its development objectives and expected results with 71% (15/21) of evaluations reporting performance as satisfactory or better. In addition, WHO programs generate benefits for target group members at the individual/household/community level with 64% of 14 evaluations rating performance for this sub-criterion as satisfactory or highly satisfactory. However, evaluations do not consistently report on the number of beneficiaries who benefited from interventions, and no results are

reported for this sub-criterion. Factors contributing to performance in objectives achievement for the WHO include strong technical design of program interventions and high levels of national ownership for key programs.

The benefits of WHO programs appear to be sustainable, but there are challenges in sustaining the capacity of its partners. The benefits of WHO programs are likely to be sustained with 73% of evaluations reporting satisfactory or highly satisfactory results in this area (although only 11 evaluations address this criteria). However, the WHO does face a challenge in the area of building its partners' institutional capacity for sustainability. Only 37% (6/16) of evaluations found WHO programs satisfactory in terms of providing support to local institutional capacity for sustainability. One factor contributing to sustainability has been the use of local networks of service providers to sustain the success of immunization programs.

Efficiency—No Results to Report. Only a few evaluations reported on cost efficiency (9) and on whether implementation of programs and achievement of objectives was timely (5). Evaluation reports that addressed these sub-criteria most often reported factors detracting from efficiency. A common feature of these findings was a link between delays in program implementation and increased costs.

WHO evaluations have not regularly addressed effectiveness in supporting gender equality or environmental sustainability. No evaluations reported on the crosscutting issue of gender equality, and only one reported on environmental sustainability, which prevented the review from identifying any results in this area. The absence of gender equality as an issue in WHO evaluations represents a critical gap in effectiveness information for the organization.

Evaluations reviewed have found WHO systems for evaluation and monitoring to be unsatisfactory. A total of 56% of reported findings on the effectiveness of evaluation systems and processes were classified as unsatisfactory or highly unsatisfactory (9 of 16 evaluations). Similarly, systems for monitoring are unsatisfactory, with 58% (11 of 19) of evaluations' findings classified as unsatisfactory or highly unsatisfactory. Sub-criteria on effective systems and processes for results-based management and evaluation results used to improve development effectiveness were addressed by only 3 and 9 evaluations, respectively. Therefore, no results are presented for these sub-criteria.

In particular, the evaluations reviewed point to a lack of financial resources and trained local staff as important factors contributing to less-than-satisfactory results in the area of evaluation and monitoring. Where evaluation systems are reported as satisfactory, one contributing factor has been the tradition of joint review of program implementation by the WHO and its partners.

Conclusions: Development Effectiveness of WHO

The evaluation function of the WHO needs strengthening: available evaluation reports do not, on balance, provide enough coverage of WHO programs and activities in the period to allow for generalization of the results to the WHO's programming as a whole but provide insights into the development effectiveness of evaluated WHO programs.

Performance: Evaluations carried out between 2007 and 2010 indicate that the WHO's activities are highly relevant to the needs of target group members (16 of 18 evaluations) and are well aligned with national government objectives and priorities (12 of 12 evaluations). In addition, WHO projects in the period under review have achieved their development objectives (15 of 21 evaluations) and resulted in positive benefits for target group members (9 of 14 evaluations). The direct benefits of WHO programming are reported as sustainable in most of the evaluations (8 of 11) that address this issue, although there are persistent challenges regarding the institutional capacity for sustainability of program arrangements (only 6 of 16 evaluations rated well).

Shortcomings: While most WHO programs reviewed have been able to achieve their direct development objectives, the level of expenditure coverage provided by the organization's evaluations is quite low. Additionally, WHO evaluations were often operationally and technically focused and, while well designed within their own parameters, they did not describe resulting changes for the target or beneficiary group. The evaluation function requires significant strengthening in order to cover WHO programs and projects, and to provide more confidence that the findings reported can be generalized to the organization. Similarly, WHO evaluations have not systematically reviewed the effectiveness of its programs in contributing to gender equality.

In an effort to strengthen the evaluation system at the WHO, the Executive Board approved the implementation of a new evaluation policy at its 131th session, held May 28–29, 2012, as part of the organization's management reform.

WHO contributes to Canada's Development Priorities. There is clear evidence that the WHO makes an important direct contribution to the Canadian international development priorities such as increasing food security (especially for pregnant and lactating women, for children and for those affected by crises) and securing the future of children and youth. There is also evidence that WHO activities contribute indirectly to sustainable economic growth through the support of public health systems and by assisting developing countries to reduce the burden of communicable and non-communicable diseases.

Recommendations to CIDA

This section contains the recommendations to CIDA based on the findings and conclusions of this effectiveness review of the WHO. Aimed at improving evaluation and results-based management at the WHO, these recommendations are in line with the objectives of Canada's existing engagements with the WHO. As one of several stakeholders working with the WHO, Canada's individual influence on the organization is limited and it may need to engage with other shareholders to implement these recommendations. (See Annex 8 for CIDA's management response.)

1. Canada should monitor efforts at reforming the evaluation function at the WHO as the new policy on evaluation is implemented. In particular, CIDA should use its influence at the Executive Board and with other donor agencies to advocate for a sufficiently resourced and

capable evaluation function that can provide good coverage of WHO programming over time.

2. CIDA should monitor the implementation of the evaluation policy so that future WHO evaluations sufficiently address gender equality.

3. CIDA should encourage the WHO to implement a system for publishing regular (possibly annual) reports on development effectiveness that builds on the work of the reformed evaluation function. In general, there is a need to strengthen the WHO commitment to reporting on the effectiveness of programs.

4. CIDA should encourage the WHO to systematically manage for results. The ongoing upgrading and further implementation of the Global Management System at the WHO may offer such an opportunity.

1.0 Introduction

1.1 Background

This report presents the results of a review of the development effectiveness of the United Nations' (UN) World Health Organization (WHO). The report utilizes a common approach and methodology developed under the guidance of the Organisation for Economic Co-operation and Development's (OECD) Development Assistance Committee (DAC) Network on Development Evaluation (DAC-EVALNET). Two pilot tests, on the WHO and the Asian Development Bank, were conducted in 2010 during the development phase of the common approach and methodology. The report relies, therefore, on the pilot test analysis of evaluation reports published by the WHO's Office of Internal Oversight Services, supplemented with a review of WHO and CIDA corporate documents, and consultation with the CIDA manager responsible for managing relations with the WHO.

The method uses a common set of assessment criteria derived from the DAC's evaluation criteria (Annex 1). The overall approach and methodology[3] were endorsed by the members of the DAC-EVALNET as an acceptable approach for assessing the development effectiveness of multilateral organizations in June 2011. For simplicity, development effectiveness is hereafter referred to as effectiveness in this report.

From its beginnings, the process of developing and implementing the reviews of development effectiveness has been coordinated with the work of the Multilateral Organization Performance Assessment Network (MOPAN). By focusing on development effectiveness and carefully selecting assessment criteria, the reviews seek to avoid duplication or overlap with the MOPAN process. Normal practice has been to conduct such a review in the same year as a MOPAN survey for any given multilateral organization. A MOPAN survey of the WHO was conducted in 2010 in parallel with this analysis.[4]

1.2 Why Conduct this Review?

The review provides Canada and other stakeholders an independent, evidence-based assessment of the development effectiveness of WHO programs for use by Canada and other stakeholders. In addition, the review satisfies evaluation requirements for all programs established by the Government of Canada's *Policy on Evaluation*.

---

[3] *Assessing the Development Effectiveness of Multilateral Organizations*, DAC Network on Development Evaluation Guidance Document, 2012. (http://www.oecd.org/dac/evaluationofdevelopmentprogrammes/dcdndep/50540172.pdf)

[4] MOPAN defines organisational effectiveness as the extent to which a multilateral organisation is organised to contribute to development results in the countries where it operates. The MOPAN Common Approach examines organisational systems, practices and behaviours that MOPAN believes are important for aid effectiveness and that are likely to contribute to results at the country level. For the WHO in 2010, the Common Approach conducted surveys in ten countries: Afghanistan, Benin, Colombia, Indonesia, Kenya, Nicaragua, Rwanda, Sri Lanka, Vietnam and Zambia.

The objectives of the review are:

- To provide the CIDA with evidence on the development effectiveness of the WHO that can be used to guide Canada's present engagement with WHO;[5] and
- To provide evidence on development effectiveness, which can be used in the ongoing relationship between the Government of Canada and the WHO to ensure that Canada's international development priorities are served by its investments.[6]

Although this report is intended, in part, to support Canada's accountability requirements within the Government of Canada, the results are expected to be useful to other bilateral stakeholders.

1.3 WHO: A Global Organization Committed to Working for Health

1.3.1 Background and Objectives

As the directing and coordinating authority on international health within the UN system, the WHO employs over 8,000 public health experts, including doctors, epidemiologists, scientists, managers, administrators and other professionals. These health experts work in 147 country offices, six regional offices and at the headquarters in Geneva.[7] The WHO's membership includes 194 countries and two associate members (Puerto Rico and Tokelau). They meet annually at the World Health Assembly to set policy for the organization, approve the budget and, every five years, to appoint the Director-General. The World Health Assembly elects a 34-member Executive Board.

The WHO's Eleventh General Programme of Work 2006–2015 defines the following core functions for the organization:

1. providing leadership on matters critical to health and engaging in partnerships where joint action is needed;
2. shaping the research agenda and stimulating the generation, translation and dissemination of valuable knowledge;
3. setting norms and standards and promoting and monitoring their implementation;
4. articulating ethical and evidence-based policy options;
5. providing technical support, catalyzing change, and building sustainable institutional capacity; and
6. monitoring the health situation and assessing health trends.

[5] "Long-term institutional funding can be defined as un-earmarked funding to a Multilateral Organization in support of that organization's mandate." (p.45). *A Review of Evidence of the Effectiveness of CIDA's Grants and Contributions 2005/06-2010/11,* CIDA, 2011.
[6] CIDA's *Review of Evidence* (2011) identifies three main types of CIDA funding to MOs: a) long-term institutional support; b) funding to specific multilateral and global initiatives; and, c) funding to multilateral initiatives delivered by other CIDA branches—including multi-bi funding delivered by Geographic Programs (p.45–46).
[7] Working for Health, An Introduction to the World Health Organization. WHO, 2007.

WHO also serves as the lead agency to coordinate international humanitarian responses in the Health cluster.[8] It hosts a number of independent programs and public private partnerships, including the Global Polio Eradication Initiative, the Stop TB Partnership, and the Partnership for Maternal Newborn and Child Health.[9]

1.3.2 Strategic Plan

WHO's Medium-Term Strategic Plan identifies 11 high-level strategic objectives for improving global health in the 2008 to 2013 period. It also includes two strategic objectives for improving the WHO's performance.

The eleven strategic objectives in global health are:[10]

1. Reduce the burden of communicable diseases;
2. Combat HIV/AIDS, tuberculosis and malaria;
3. Prevent and reduce chronic non-communicable diseases;
4. Improve maternal and child health, sexual and reproductive health, and promote healthy aging;
5. Reduce the health consequences of crises and disasters;
6. Prevent and reduce risk factors for health, including tobacco, alcohol, drugs and obesity;
7. Address social and economic determinants of health;
8. Promote a healthier environment;
9. Improve nutrition, food safety and food security;
10. Improve health services and systems; and
11. Ensure improved access, quality and use of medical products and technologies.

The Medium-Term Strategic Plan also identified two objectives directed toward the WHO's own roles and functions:

12. Provide global health leadership in partnership with others; and,
13. Develop the WHO as a learning organization.

1.3.3 Work and Geographic Coverage

The WHO is funded through both assessed[11] and voluntary contributions from member states. Foreign Affairs and International Trade Canada is responsible for Canada's assessed

[8] As lead for the Health cluster, WHO is accountable to the UN's Office for the Coordination of Humanitarian Affairs. To learn more about the cluster approach, see http://business.un.org/en/documents/6852.
[9] CIDA's Strategy for Engagement with the World Health Organization (WHO). CIDA, 2011.
[10] Medium-Term Strategic Plan, 2008-2013. WHO, 2011, p. 77.

contribution. Similarly to other UN organizations, the WHO prepares a biennium budget covering the two years of operations. The program budget for the 2010–2011 biennium was USD 4.54 billion, of which USD 945 million was assessed contributions.[12]

Since the budget is comprised of both assessed and voluntary contributions, the actual funds available to the WHO for expenditure on a program or priority in any given year may be either more or less than budgeted (depending on the volume of voluntary contributions). Table 1 presents the approved budget amount, the actual funds reported as available over the biennium, and the amount spent.[13]

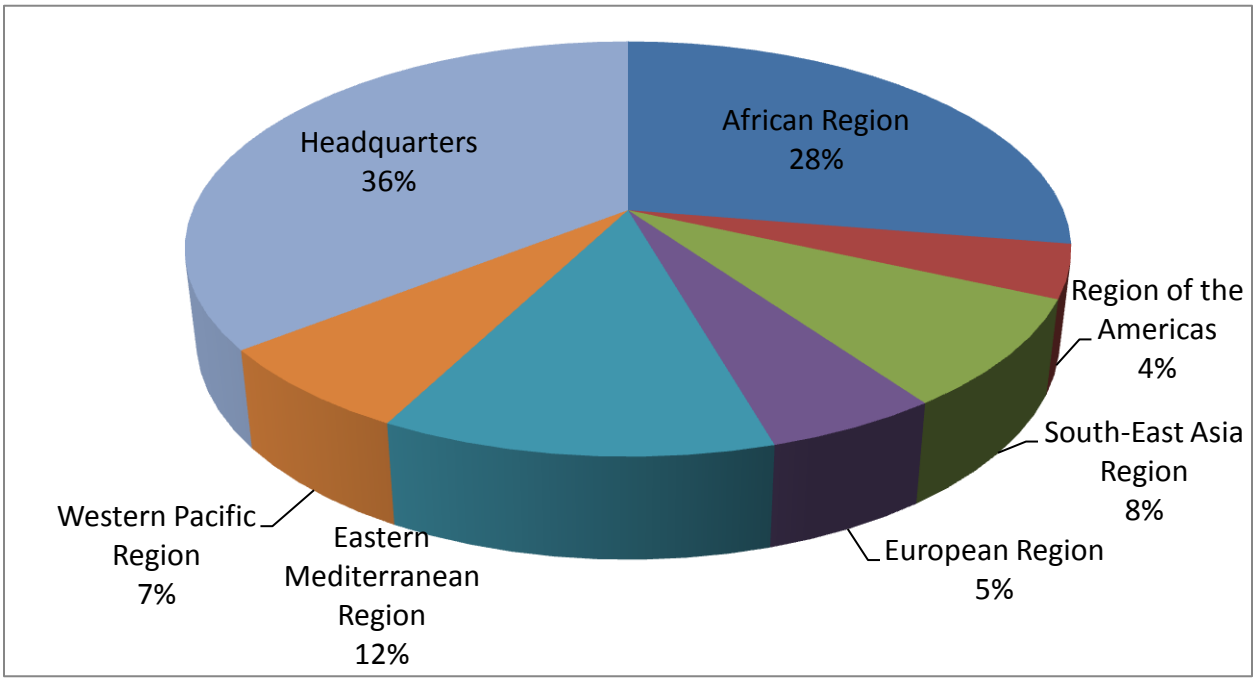
Table 1: WHO Budget (USD) and Expenditures by Strategic Objective (2010–2011)

WHO Strategic Objectives	Approved Budget 2010–2011	Funds Available at Dec. 31, 2011	Expenditures at Dec. 31, 2011	% of Total Expenditures in 2011
1. Communicable Diseases	1,268	1,472	1,290	35%
2. HIV/AIDS, Tuberculosis and Malaria	634	535	446	12%
3. Chronic Non-communicable Diseases	146	112	98	3%
4. Child, Adolescent, Mother Health and Aging	333	222	190	5%
5. Emergencies and Disasters	364	393	312	8%
6. Risk Factors for Health	162	109	94	3%
7. Social and Economic Determinants of Health	114	42	37	1%
8. Healthier Environment	63	94	83	2%
9. Nutrition and Food Safety	120	70	62	2%
10. Health Systems and Services	474	348	298	8%
11. Medical Products and Technologies	115	158	137	4%
12. Global Health Leadership	223	269	264	7%
13. WHO as a Learning Organization[14]	524	420	405	11%
TOTAL	**4,540**	4,244	3,717	100%

Figure 1 presents the share of 2010 expenditures accounted for by each region of operations and by WHO headquarters (see following page).

[11] Assessed contributions are a binding obligation of membership and represent a given nation's share of the costs of operations of an international organization of which the nation is a member.
[12] Assessed Contributions Overview for all Member States as of 31 December 2011, WHO, 2011.
[13] Programme Budget 2010–2011: Performance Assessment Report, WHO, 2012, p. 14.
[14] Strategic Objective 13 covers core administrative functions such as planning, reporting, human resources management, financial management and information technology.

Figure 1: Regional Share of WHO Expenditures in 2010–2011



1.3.4 Evaluation and Results Reporting

Evaluation

Evaluation at the WHO is a decentralized responsibility with most evaluations being commissioned and managed by individual technical programs and regional and country offices. The Office of Internal Oversight Services (OIOS) reports directly to the Director General and conducts internal audits, investigates alleged wrongdoing, and implements the policy on programme evaluation. The OIOS has acted as the custodian of the evaluation function.

When the pilot test was carried out in 2010, evaluations commissioned by WHO were not published either as paper documents or online. Therefore, a request was made to the OIOS to identify and source the evaluations for the pilot test. The OIOS indicated that in most cases, evaluation reports were owned by both WHO and the countries covered in the evaluation. Copies would need to be requested from the WHO Country Offices in question. No central repository of published evaluations was available in either electronic or hard copy form.

OIOS staff indicated in May 2010 that a major review of the evaluation policy at the WHO was underway and that a new policy and structure would be forthcoming.

The recasting and restructuring of WHO evaluation policy has now become one element in a major, coordinated initiative to reform the management of the organization. In preparation for a

Special Executive Board Meeting in November 2011, the WHO Secretariat produced a five-element proposal for managerial reforms[15] that covered the areas of:

1. Organizational effectiveness, alignment and efficiency;
2. Improved human resources policies and management;
3. Enhanced results-based planning, management and accountability;
4. Strengthened financing of the organization, with a corporate approach to resource mobilization; and,
5. Strategic communications framework.

Under the heading of results-based planning, management and accountability, the proposed managerial reforms aimed to delineate "an approach to independent evaluation."

The new evaluation policy was officially adopted by the Executive Board at its 131st session, held in May 2012.[16] The policy aims to: foster a culture and use of evaluation across the WHO; provide a consolidated institutional framework for evaluation at the three levels of the WHO; and facilitate conformity with best practice and with the norms and standards of the United Nations Evaluation Group.

The new policy[17] opts to strengthen the OIOS rather than create a new evaluation unit reporting directly to the Board. The policy also delineates in considerable detail the roles and responsibilities of the Executive Board, the newly created Global Working Group on Evaluation, and the OIOS with regard to evaluations. It also describes the principles that guide all evaluation work at the WHO.

The most important new duties of the OIOS include: preparing an annual organization-wide work-plan for evaluations; maintaining an online inventory of evaluations performed at the WHO; ensuring that evaluation reports conform to the requirements of the policy; maintaining a system to track the implementation of management responses to evaluations; and submitting an annual report on evaluation activities to the Executive Board through the Director General.

It remains to be seen how these proposals will be implemented and what effect they will have on the strength of the evaluation function at the WHO. The introduction to the proposed new policy describes the challenge facing the WHO:[18]

> "From a broader institutional perspective, it [WHO] has been less successful in fostering an evaluation culture, developing evaluation capacity at all levels of the organization and in promulgating participatory approaches to evaluations. The causes for this include institutional arrangements for the evaluation function (including a lack of a direct mechanism for oversight by

[15] WHO Managerial Reforms. WHO Secretariat. 2011.
[16] Decisions and list of resolutions, World Health Organization Executive Board, 131st session EB131/DIV/2, 2012.
[17] WHO reform: WHO Evaluation Policy. WHO Secretariat. 2011.
[18] WHO Reform, Draft Formal Evaluation Policy, paragraph 6, 2012.

the governing bodies) and the absence of an effective budgetary allocation mechanism to resource the evaluation function."

The WHO has not yet been the subject of a DAC/UNEG Professional Peer Review of the Evaluation Function and the review team did not undertake such a review. However, the review team conducted its own quality review of the evaluations for inclusion in this report.

The results of the review team's quality analysis were mixed, with 52% of reviewed evaluations scoring 30 points or more, and 24% receiving scores of less than 19 out of a possible 48. For quality criteria I, "evaluation findings are relevant and evidence based," only three evaluation reports out of 25 scored less than three from a possible score of five (see Annex 3 for details of the review methodology and Annex 4 for the evaluation quality scoring grid). All evaluations were retained for the review since scores overall were judged reasonable.

WHO evaluation reports were often operationally and technically focused; that is, they were concerned with how well a given service delivery method, surveillance system, or even the introduction of a new vaccine was implemented rather than the resulting changes for the target or beneficiary group. This is a significant problem for assessing evaluation quality because these studies (while often well designed within their own parameters) often lacked key components of a quality evaluation (such as the effects on the target or beneficiary group) when assessed against the quality criteria derived from UNEG standards.

Although no evaluations were screened out due to quality concerns, evaluation reports do not address all the criteria identified essential elements for effective development. As a result, each sub-criteria examined below is addressed by fewer than 25 evaluations. This review examines the data available on each criterion before presenting results, and does not present results for some criteria.

Results Reporting

The WHO does not prepare an annual report on development effectiveness or an annual summary of the results of evaluations. It does provide, however, extensive reporting on the global and regional situation in health to the World Health Assembly each year. It also presents special reports on specific global topics and challenges in public health on an annual basis. Every two years, the WHO publishes a Performance Assessment Report, which describes the extent to which the WHO has achieved its strategic objectives and sub-objectives in the previous biennium.

The Global Management System

For some time, the WHO has been in the process of implementing a system of results monitoring and reporting based on Oracle software. This Global Management System was in development as early as 2008 and is currently being upgraded after a lengthy implementation phase. The Global Management System has as one goal the alignment of program and project planning, implementation and monitoring with agency strategic objectives at a corporate, regional and national level.

Since 2008, the WHO has made an effort to implement the System in each of its regions and by January 2011 was able to report to its Executive Board that it had made "considerable progress" in implementing the system in five regions and at headquarters. The Executive Board (EB128/3) welcomed the reported progress but expressed concern that the Region of the Americas/PAHO had chosen not to implement the system.

In May 2011, the Secretariat at the WHO reported to the Programme, Budget and Administration Committee of the Executive Board on progress in implementing the Global Management System. The Committee in its report to the Executive Board noted that:[19]

> "The Global Management System had been successfully rolled out in the Africa Region. Questions were asked regarding the planned upgrade of the System and its related cost as well as the savings that will result from its implementation. Queries were also raised with regard to harmonization between the Global Management System and the new system in the Region of the Americas/PAHO."

Available documentation on the System suggests its primary focus is still finance, administration, resource allocation planning, and human resources management. It is not yet clear if the System, as implemented, will effectively strengthen the results management and reporting system at the WHO.

At its Special Session on WHO reform in November 2011, the Executive Board welcomed the Director General's proposals on managerial reform and requested that these proposals be taken forward in several areas, including the improvement of monitoring and reporting.[20] As already noted, one consequence of this request was the proposal for a new policy on evaluation, which was officially adopted by the Executive Board in May 2012. It is not yet clear whether this will include an effort to strengthen reporting on the development effectiveness of WHO programs, beyond that expected from the full implementation of the Global Management System.

Finally, it should be noted that the WHO published a performance assessment report in May 2012 to track indicators to measure progress toward the WHO's strategic objectives and sub-objectives over the previous biennium.[21] A similar report was published in 2010. While the reports provide only global (or sometimes regional) information and do not describe the methodology used to track and verify indicators, they represent an excellent step toward reporting on the WHO's performance.

[19] Report of the Programme, Budget and Administration Committee of the Executive Board (EB129/2). WHO, 2011.
[20] Decisions. Executive Board Special Session on WHO Reform (EBSS/2/DIV/2). WHO 2011.
[21] Programme Budget 2010-2011 Performance Assessment Report. WHO. 2012.

2.0 Methodology

This section describes briefly the main elements of the methodology used for the review. A more detailed description of the methodology is presented in Annex 3.

2.1 Rationale

As an important United Nations (UN) Organization, the WHO was chosen for the pilot test of the common approach, together with the Asian Development Bank (a Multilateral Development Bank). The selection of the WHO allowed for testing the approach on a specialized agency of the UN with a strong social mandate. DAC-EVALNET members also expressed considerable interest in an effectiveness review of the WHO as an organization critical to efforts to achieving the health-related Millennium Development Goals (MDGs).

The term "common approach" describes the use of a standard methodology, as implemented in this review, to assess consistently the (development) effectiveness of multilateral organizations. It offers a rapid and cost effective way to assess effectiveness relative to a more time-consuming and costly joint evaluation.[22] The approach was developed to fill an information gap regarding the effectiveness of multilateral organizations. Although these organizations produce annual reports to their management and/or boards, bilateral shareholders were not receiving a comprehensive overview of the organizations' performance on the ground. The Multilateral Organization Performance Assessment Network (MOPAN) seeks to address this issue through organizational effectiveness assessments. This approach complements MOPAN's assessments.

The approach suggests conducting a review based on the organization's own evaluation reports when two specific conditions exist:[23]

1. There is a need for field-tested and evidence-based information on the effectiveness of the multilateral organization.
2. The multilateral organization under review has an evaluation function that produces an adequate body of reliable and credible evaluation information that supports the use of a meta-evaluation methodology to synthesize an assessment of the organization's effectiveness.

The WHO met one of the two requirements for successfully carrying out an effectiveness review at the time of the pilot test. There was a clear need for more field tested and evidence-based information on the effectiveness of WHO programming. Results for the second test were more marginal. The supply of reasonable quality evaluation reports available at the time of the pilot test was limited, with only 25 such evaluations provided by the WHO over the 2007–2010 period

[22] "Joint evaluation" refers to a jointly funded and managed comprehensive institutional evaluation of an MO. It does not refer to DAC/UNEG Peer Reviews of the Evaluation Function.
[23] *Assessing the Development Effectiveness of Multilateral Organizations: Approach, Methodology and Guidelines*, Management Group of the Task Team on Multilateral Effectiveness, DAC EVALNET, 2011.

to the pilot test team. The review was completed because these 25 evaluations were able to address moderately four of the six main criteria used to assess effectiveness. However, this narrow supply of reasonable evaluations limits the extent to which the results can be generalized across the organization.

2.2 Scope

The sample of 25 evaluations available for this review of the WHO provides limited coverage of the over 4.5 billion USD in programming budget available over the 2010–2011 biennium. It is difficult to estimate the level of coverage provided because the evaluation reports often do not include data on the overall value of the programs under evaluation. Nonetheless, the evaluations provide coverage at the country, regional and global/thematic level, and there are some interesting points of congruence between the sample and the profile of the WHO budget.

- Communicable Diseases (strategic objective 1): 8 of the 25 evaluations deal with the implementation of Extended Programs of Immunization in a range of countries (Central African Republic, the Democratic Republic of Congo, Cameroon, Vietnam, Sierra Leone, Zambia and the Philippines). These programs directly contribute to the most significant WHO strategic objective in dollar terms.
- Emergencies and Disasters (strategic objective 5): 3 of the 25 evaluations deal with Health Action in Crisis at the regional or country level: 1 for Africa, 1 for Myanmar, and 1 for Palestine. (In addition, a program evaluation of Health Action in Crisis is included in the global category below.) These programs contribute to the third-largest strategic objective in terms of funding.
- A significant number of the evaluations reviewed are global or organizational in scope. They include:

    1. Evaluation of the Making Pregnancy Safer Department (2010);
    2. Independent evaluation of major barriers to interrupting Poliovirus transmission (2009);
    3. Independent Evaluation of the Stop TB Partnership (2008);
    4. Review of the Nutrition Programmes of the WHO in the context of current global challenges and the international nutrition architecture (2008);
    5. Assessment of the Implementation, Impact and Process of WHO Medicines Strategy (2007);
    6. Health Actions in Crisis Institutional Building Program Evaluation (2007);
    7. Programmatic Evaluation of Selected Aspects of the Public Health and Environment (PHE) Department (2007); and
    8. Thematic Evaluation of the WHO's Work with Collaborating Centres (2007).

The evaluations covered in this review were all produced by the WHO in the period from early 2007 to mid-2010 when the review was carried out (Annex 3). While some covered programming periods before 2007, most of the WHO program activities covered in the reviewed evaluations will have occurred between 2007 and 2010. The review team also analyzed

selected WHO documents published in 2011 and early 2012 to provide an update to some of the findings of the reviewed evaluations.

In summary, while the list of suitable evaluations for review obtained from the organization by the pilot test team cannot be easily compared to the geographic and programmatic distribution of activities, it does provide at least a partial body of field-tested evaluation material on effectiveness. For that reason (and to learn what lessons could be drawn from the experience of conducting the study) the team proceeded with the pilot test effectiveness review of the WHO.

In addition to the 25 evaluation reports, the review examined relevant WHO policy and reporting documents, such as the reports of the Programme, Budget and Administration Committee to the Executive Board, Reports on WHO Reform by the Director-General, Evaluation Policy Documents, Annual Reports and the Interim Assessment of the Medium-Term Strategic Plan (see Annex 6). These documents allowed the review team to assess the ongoing evolution of evaluation and results reporting at the WHO and to put in context the findings reported in the evaluation reports.

The review team also carried out an interview with staff of the Office of Internal Oversight Services (OIOS) at the WHO to understand better the universe of available WHO evaluation reports and to put in context the changing situation of the evaluation function. Finally, the review team interviewed the CIDA manager most directly responsible for the ongoing relationship between CIDA and the WHO in order to better assess the organization's contribution to Canada's international development priorities.

2.3 Criteria

The methodology does not rely on a particular definition of (development) effectiveness. The Management Group and the Task Team created by the DAC-EVALNET to develop the methodology had previously considered whether an explicit definition of effectiveness was needed. In the absence of an agreed upon definition of effectiveness, the methodology focuses on some of the essential characteristics of developmentally effective multilateral organization programming, as described below:

1. **Relevance of interventions:** The programming is relevant to the needs of target group members;
2. **Achievement of Development Objectives and Expected Results:** Programming contributes to the achievement of development objectives and expected development results at the national and local levels in developing countries;
3. **Sustainability of Results/Benefits:** The benefits experienced by target group members and the development results achieved are sustainable in the future;
4. **Efficiency:** Programming is delivered in a cost-efficient manner;
5. **Crosscutting Themes (Environmental Sustainability and Gender Equality):** Programming is inclusive in that it would support gender equality and would be environmentally sustainable (thereby not compromising the development prospects in the future); and

6. **Using Evaluation and Monitoring to Improve Effectiveness:** Programming enables effective development by allowing participating and supporting organizations to learn from experience and uses performance management and accountability tools, such as evaluation and monitoring, to improve effectiveness over time.

The review methodology therefore involves a systematic and structured meta-synthesis of the findings of WHO evaluations, as they relate to these six main criteria and 18 sub-criteria that are considered essential elements of effective development (Annex 5). The main criteria and sub-criteria are derived from the DAC evaluation criteria.

2.4 Limitations

As with any meta-evaluation, there are methodological challenges that limit the findings. For this review, the limitations include: sampling bias; the challenge of ensuring adequate coverage of the criteria used; and problems with the classification of evaluation findings.

The major limitation to this review of the WHO has been the number of evaluation reports available at the central OIOS and made available to the review team in 2010 (covering the period 2007 to 2010). The set of available and valid evaluation reports does not provide, on balance, enough coverage of WHO programs and activities in the period to allow for generalization of the results to WHO programming as a whole. The 25 available evaluation reports do, however, provide insights into the development effectiveness of WHO programs evaluated during the period.

A further limitation arises from the fact that many of the 25 evaluations did not address some of the sub-criteria used to assess effectiveness. Because of the limitations arising from the small number of evaluations available and the lack of coverage of some sub-criteria, findings are reported below for only those criteria where coverage was rated either strong or moderate.

3.0 Findings on the Development Effectiveness of WHO

Insufficient evidence available to make conclusions about WHO

The major limitation to this review was that only 25 evaluation reports were available at the central OIOS and made available to the review team. This small sample does not provide enough coverage of WHO programs and activities to allow for generalization of results to the WHO as a whole.

The limited number of evaluation reports also did not allow reviewers to control for selection bias in the evaluation sample. This challenge is compounded by the fact that evaluation reports did not always report the programme budget that was evaluated.

Finally, many of the available evaluations did not address the sub-criteria used in this review to assess effectiveness, limiting the amount of information this review is able to report.

Taken together, these limitations mean that there is insufficient information available to make conclusions about the WHO's development effectiveness. However, in the interest of providing useful, synthesized information, some findings are presented below.

WHO's 2012 evaluation policy

An analysis of the 2012 WHO evaluation policy (Section 3.6.4) indicates that while the approval of an evaluation policy represents a positive step, gaps remain in the policy regarding the planning, prioritizing, budgeting and disclosure of WHO evaluations.

A 2012 United Nations Joint Inspection Unit review also raises concerns about independence and credibility of WHO evaluations, suggests that the WHO should have a stronger central evaluation capacity, and recommends that a peer review on the evaluation function be conducted by the United Nations Evaluation Group and be presented to the WHO Executive Board by 2014.

Observations on Development Effectiveness of the WHO

This section presents the results of the development effectiveness review as they relate to the six main criteria and their associated sub-criteria (Table 2 and Annex 5). In particular, Table 2 below describes the ratings assigned by the review team of "satisfactory" or "unsatisfactory" for each of the six major criteria and their associated sub-criteria. The table also presents the numbers of evaluations that addressed each sub-criterion (represented by the letter *a*).[24]

No results are provided for sub-criteria addressed in less than 10 evaluations. Where coverage for a given sub-criterion was strong (that is, addressed by 18–25 evaluation reports), or moderate (addressed by 10–17 evaluation reports), results on effectiveness are presented.

Each of the following sections begins with a summary of the coverage and key findings, and follows with the main factors contributing to these results. A quantification of how many evaluations identified a particular factor describes the importance of positive and negative factors contributing to results under each assessed criteria.

[24] *a* = the number of evaluations that addressed the sub-criteria, *n* = the number in the sample

Table 2: Summary of Findings by Criteria for Assessing Development Effectiveness

Relevance of interventions

Sub-criteria	*a**	Coverage Level**	Evaluations Rated Satisfactory (%)***	Evaluation Rated Unsatisfactory (%)***
1.1 Programs are suited to the needs of target group members	18	Strong	89%	11%
1.2 Programs are aligned with national development goals	12	Moderate	100%	0%
1.3 Effective partnerships with governments	18	Strong	61%	39%
1.4 Program objectives remain valid	21	Strong	100%	0%
1.5 Program activities are consistent with program goals	20	Strong	60%	40%

Achieving Development Objectives and Expected Results

Sub-criteria	*a**	Coverage Level**	Evaluations Rated Satisfactory (%)***	Evaluation Rated Unsatisfactory (%)***
2.1 Programs and projects achieve stated objectives	21	Strong	71%	29%
2.2 Positive benefits for target group members	14	Moderate	64%	36%
2.3 Substantial numbers of beneficiaries	8	Weak	N/A	N/A

Sustainability of Results/Benefits

Sub-criteria	*a**	Coverage Level**	Evaluations Rated Satisfactory (%)***	Evaluation Rated Unsatisfactory (%)***
3.1 Program benefits are likely to continue	11	Moderate	73%	27%
3.2 Programs support institutional capacity for sustainability	16	Moderate	37%	63%

Efficiency

Sub-criteria	*a**	Coverage Level**	Evaluations Rated Satisfactory (%)***	Evaluation Rated Unsatisfactory (%)***
4.1 Programs evaluated as cost efficient	9	Weak	N/A	N/A
4.2 Program implementation and objectives achieved on time	5	Weak	N/A	N/A

Crosscutting Themes: Inclusive Development Which can be Sustained (Gender Equality and Environmental Sustainability)

Sub-criteria	*a**	Coverage Level**	Evaluations Rated Satisfactory (%)***	Evaluation Rated Unsatisfactory (%)***
5.1 Programs effectively address gender equality	0	Weak	N/A	N/A
5.2 Changes are environmentally sustainable	1	Weak	N/A	N/A

Using Evaluation and Monitoring to Improve Development Effectiveness

Sub-criteria	*a**	Coverage Level**	Evaluations Rated Satisfactory (%)***	Evaluation Rated Unsatisfactory (%)***
6.1 Systems and processes for evaluation are effective	16	Moderate	44%	56%
6.2 Systems and processes for monitoring are effective	19	Strong	42%	58%
6.3 Systems and processes for RBM are effective	3	Weak	N/A	N/A
6.4 Evaluation results used to improve development effectiveness	9	Weak	N/A	N/A

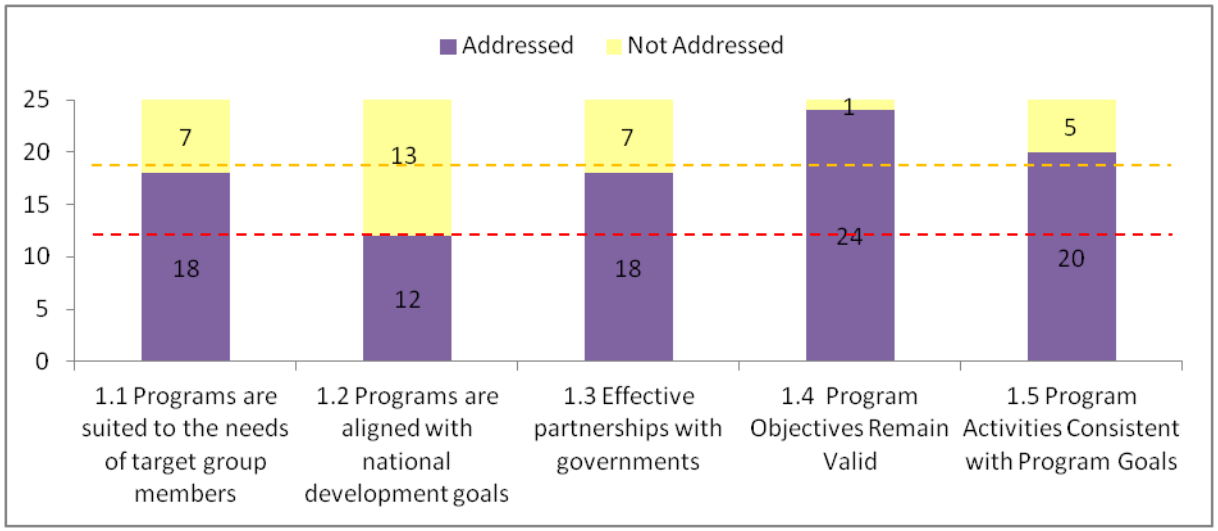
- **a* = number of evaluations addressing the given sub-criterion
- ** Coverage Level defined as: Strong: *a* = 18–25, Moderate: *a* = 10–17, Weak: *a* = under 10
- *** Satisfactory ratings include “satisfactory” and “highly satisfactory”; unsatisfactory ratings include “unsatisfactory” and “highly unsatisfactory”

3.1 WHO programs appear relevant to stakeholder needs and national priorities

3.1.1 Coverage of Sub-criteria

As demonstrated in Figure 2, the evaluations reviewed generally addressed the topic of relevance, with four of five sub-criteria (1.1, 1.3, 1.4 and 1.5) rated strong in coverage. Coverage in one sub-criterion (1.2) was rated moderate, as it was addressed in 12 evaluations.

Figure 2: Number of Evaluations Addressing Sub-criteria for Relevance



3.1.2 Key Findings

In summary, the evaluations reviewed rate WHO supported projects and programs high on scales of relevance. In particular, the programs evaluated are well suited to the needs of target group members and aligned with national priorities, and their objectives remain valid over time, as described in Figure 3 below.

On the question of whether or not WHO-supported programs and projects are suited to the needs of target group members (sub-criterion 1.1), the review found that 16 of 18 evaluations reports that addressed the criterion (89%) rated as reporting satisfactory or better findings and half of those rated as highly satisfactory. All 12 evaluations addressing the question of alignment of WHO-supported programs with national development goals and priorities (sub-criterion 1.2) were rated satisfactory or better.

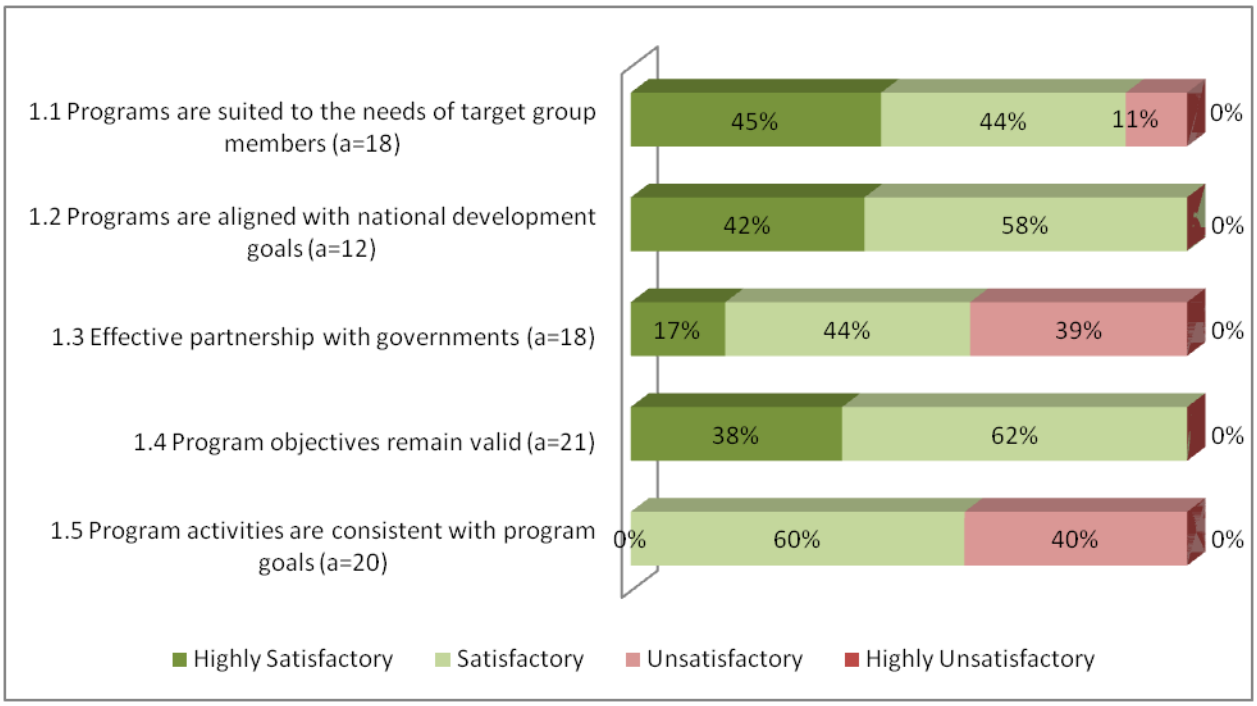
More evaluations considered sub-criterion 1.3, the effectiveness of partnerships with all levels of government, with 11 of the 18 evaluations (61%) rated satisfactory or better. On the other hand, 7 evaluations (39%) were rated as unsatisfactory.

All 20 evaluations that addressed sub-criterion 1.4 on the continued validity of program objectives reported findings of satisfactory or better. The question of the fit between program objectives and program activities (sub-criterion 1.5) is not quite so clear-cut with only 12 of 20 evaluation reports (60%) reporting findings classified as satisfactory. This also reflects the

technically focused nature of some WHO evaluations, which did not allow the review team to verify that the design of projects includes a systematic assessment of causal linkages between program activities and outputs and objectives achievement.

Findings from this review and from the 2010 MOPAN survey converge on the subject of relevance. [25] The WHO ranked at the top end of 'adequate' on the MOPAN indicator for 'results developed in consultation with beneficiaries' and 'strong' for the indicators 'expected results consistent with national development strategies' and 'supporting national plans.'

Figure 3: Relevance of Interventions (Findings as percentage of number of evaluations addressing sub-criterion (= a), n = 25)



Highlight Box 1 below provides an illustration of successful results for criterion 1.2, "Programs are aligned with national development goals," as remarked on in the evaluation of child health in Guyana.

Highlight Box 1

Aligning with national priorities in Guyana

A national strategic plan for the reduction of maternal and neonatal mortality 2006-2012 has been developed, which focuses on achieving the MDG mortality targets set in the UN General Assembly Special Session in 2000. Improvement of the health status of mothers and children is also given priority in the National Health Plan 2003-07, and the Poverty Reduction Strategy Paper (2002).

Review of Child Health in Guyana

[25] MOPAN Common Approach Institutional Report for the World Health Organization (WHO) 2010 (Vol. I), Multilateral Organisation Performance Assessment Network, 2011, pp. 21 and 35.

3.4.3 Contributing Factors

Two important factors contributed to the positive evaluation findings in the area of relevance:

- The WHO's experience in matching program design to the burden (morbidity and mortality) of disease in programming countries (11 evaluations);[26] and
- The use of consultations with key stakeholders at national and local levels to ensure program design matched user needs and national priorities.

Highlight Box 2 provides an example of how global consultations were used to help define the framework for the WHO's intervention under Health Action in Crisis programming in crisis-affected countries.

A number of factors contributed to some of the unsatisfactory evaluation findings in the area of relevance:

- Unclear relations and responsibilities among participating government and non-government organizations (2 evaluations).
- Lack of coordination among supporting organizations (the WHO and the UN Office for Coordination of Humanitarian Affairs, for example), which made it difficult to coordinate with regional and local government partners (1 evaluation).
- Capacity weaknesses among both government and non-government partners (1 evaluation).
- Misunderstandings within the programs over the roles of different agencies and different units of government (1 evaluation).

Highlight Box 2

Consultations Used to Define Institutional Support for Health Action in Crisis (HAC)

In 2005, a consultative process involving over 300 stakeholders globally defined four core functions for WHO's work in countries affected by crises. This framework was endorsed by the 2005 World Health Assembly resolution WHA58.1. The first core function was to promptly assess health needs of populations affected by crises. This was considered to be particularly well understood and implemented. The evaluation noted increased satisfaction with the improvement of WHO's capacity for needs assessments and that it improved in all countries visited, although needs always exceeded resources.

Evaluation of HAC Institutions Building Program

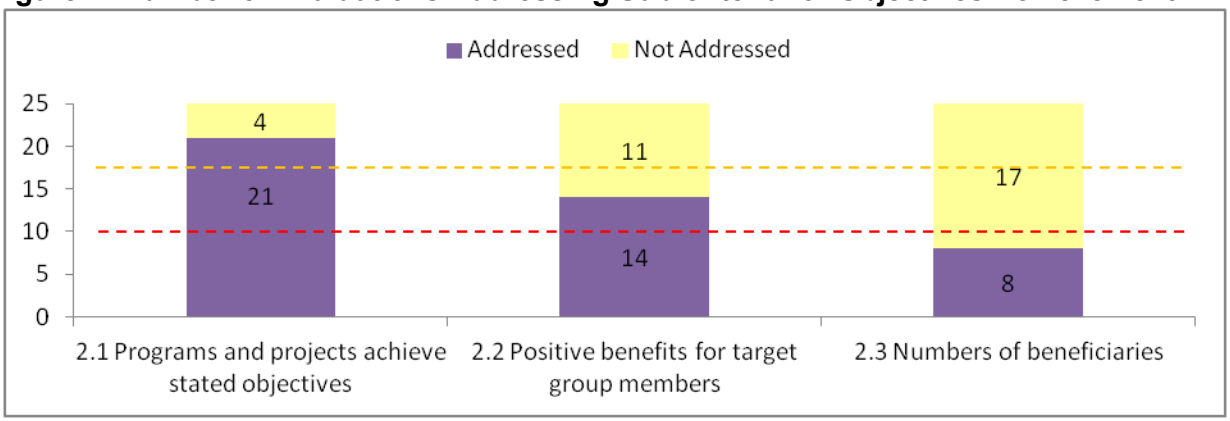
[26] The numbers in brackets track the frequency with which analysts in the review team highlighted comments when reviewing evidence to support the findings ratings for a given evaluation. The figures cited do not provide an exhaustive census of how many citations were made but, rather, an overall portrait of the emphasis given in different evaluation reports. They are meant to be illustrative, not a definitive count of occurrence of the factors.

3.2 The WHO appears to be effective in achieving its development objectives and expected results

3.2.1 Coverage

Two of the three sub-criteria for objectives achievement and expected results have a strong (sub-criteria 2.1) or moderate level (sub-criteria 2.2) of coverage. As illustrated in Figure 4, coverage of sub-criteria 2.3 (programs and projects made differences for a substantial number of beneficiaries) was weak with only 8 evaluations addressing the number of program beneficiaries.

Figure 4: Number of Evaluations Addressing Sub-criteria for Objectives Achievement

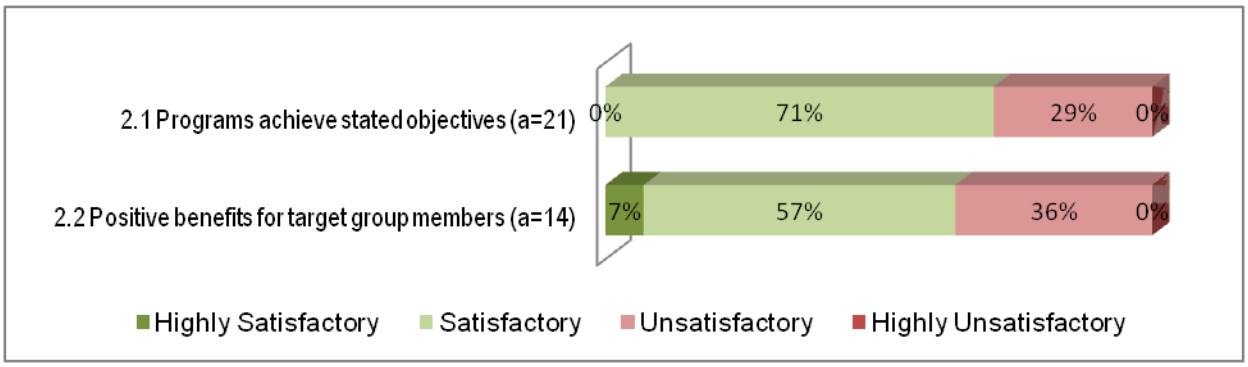


3.1.2 Key Findings

In summary, the evaluations reviewed indicate that WHO programs achieve their developmental objectives and that they result in benefits for the designated target group members.

Of 21 evaluation reports that addressed sub-criterion 2.1, “Programs and projects achieve stated objectives,” 15 (71.4%) reported findings rated as satisfactory while only 6 (28.6%) were scored unsatisfactory. WHO programs also resulted in benefits for target group members, as noted in the findings of 9 (64%) of the 14 evaluations that addressed sub-criterion 2.2.

Figure 5: Results for Objectives Achievement (Findings as percentage of number of evaluations addressing sub-criterion (= a), n = 25)



Highlight boxes 3 and 4 provide an illustration of how WHO programs achieve their development objectives. Highlight Box 3 reports that the WHO was able to play a neutral brokering role in order to provide leadership in the coordination of the UN Health Cluster during emergency operations in Africa. Highlight Box 4 provides an example of WHO programming contributing to positive outcomes in newborn and child health in Cambodia.

Highlight Box 3

WHO Health Action in Crisis (HAC) in Africa

The evaluation found evidence that WHO is able to put the neutral brokering role in practice without undermining its organizational mandate. The evaluation confirmed that WHO can implement the leadership role for the coordination of the Health Cluster. Country Offices provided good support to partners with regard to needs assessments, health outcome and health services surveys, and providing regular disease surveillance data.

Evaluation of HAC's Work in Africa

Highlight Box 4

Contributing to Newborn and Child Health in Cambodia

Overall neonatal and child mortality rates fell between 1996–2000 and 2001–2005. Improvements have been noted in a number of areas, including: neonates protected against tetanus at birth; neonates and mothers receiving early postnatal care contacts; initiation of early breastfeeding; exclusive breastfeeding to six months; living in households using iodized salt; and vaccination coverage. Improvements are needed in other areas, including: antenatal care coverage and skilled attendance at birth.

Review of Newborn and Child Health Program in Cambodia

3.2.3 Contributing Factors

Two common factors were noted in the evaluations as contributing to the achievement of development objectives in WHO programs:

- Strong technical elements in program design which matched the program intervention to the burden of disease (11 evaluations); and

- High levels of national and local ownership resulting from consultative processes of program development (4 evaluations).

Where evaluations reported that benefits for target group members were missing or limited in scope they noted:

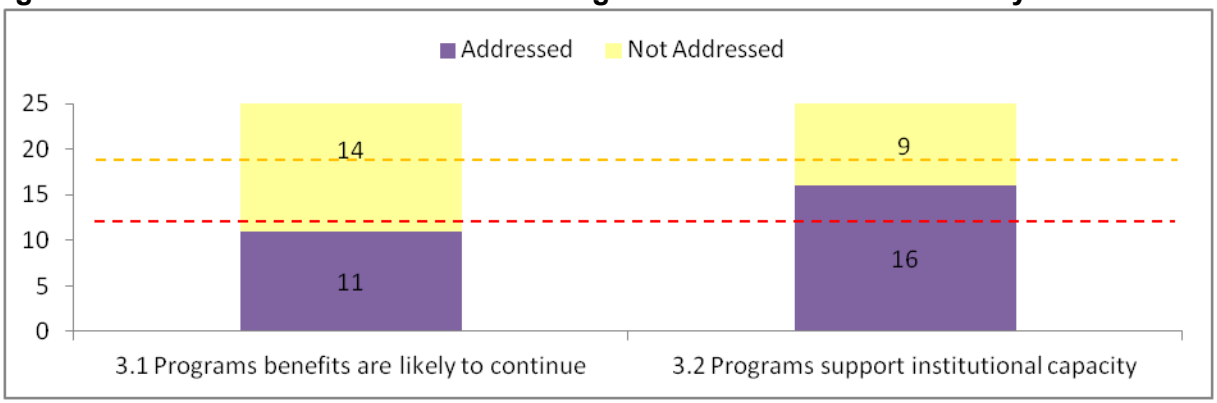
- Weak or delayed implementation (2 evaluations);
- Lack of adequate financing and human resources invested in the program (1 evaluation); and
- Delays in the expected increase in donor funding (1 evaluation).

3.3 Benefits of WHO programs appear to be sustainable but there are challenges in sustaining the capacity of partners

3.3.1 Coverage of Sub-criteria

Evaluations provided a moderate level of coverage for both the sub-criteria for assessing sustainability. Sub-criterion 3.1, "Program benefits are likely to continue," was addressed by 11 evaluation reports, while sub-criterion 3.2, "Programs support institutional capacity for sustainability," was addressed by 16 of 25 evaluation reports.

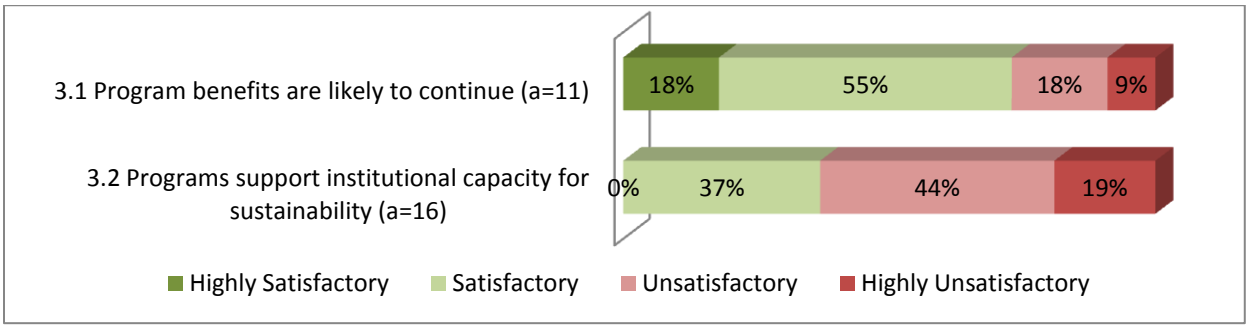
Figure 6: Number of Evaluations Addressing Sub-criteria for Sustainability



3.3.2 Key Findings

The findings regarding sustainability reflect a mixed level of performance (**Figure 7**). Evaluations reviewed indicate that the benefits of WHO programs are sustainable but that there are important challenges to ensuring that the institutional arrangements for ongoing program delivery are sustainable. On sub-criterion 3.1, "Program benefits are likely to continue," 8 of 11 evaluation reports (73%) reported findings of satisfactory or better. In contrast, for sub-criterion 3.2, "Programs support institutional capacity for sustainability," only 37% of evaluations reported positive findings, with 10 (63%) of 16 evaluations classified as unsatisfactory or worse.

Figure 7: Sustainability of Results/Benefits (Findings as percentage of number of evaluations addressing sub-criterion (= a), n = 25)



3.3.3 Contributing Factors

Three factors were cited in evaluations as contributing to the sustainability of the results of WHO programming:

- Strong national and local ownership (4 evaluations);
- Consultative processes for identifying key health issues and agreeing on implementation arrangements for solutions (4 evaluations); and
- Use of local networks for sustaining the success of immunization program arrangements.

Two factors were identified as contributing to less than satisfactory results for sustainability:

- The absence of adequate and sustained financial resources from both government and donors to sustain program services at current levels (1 evaluation); and
- Problems in the disruption of WHO services to countries and areas in crisis (1 evaluation).

Highlight Box 5 provides an illustration of strong local institutional capacity and the use of networking to improve program effectiveness and sustainability.

Highlight Box 5

Contribution to Capacity Development

Vietnam's EPI program and health system is well functioning and well positioned to meet these coming challenges...strong networks established between commune health centers and village health workers have been identified by this review to be a critical factor in immunization program success. The implementation of the school-based measles second dose program also testifies to the strength of local area institutional and social networks in facilitating access of the population to health care services.

Vietnam EPI Evaluation

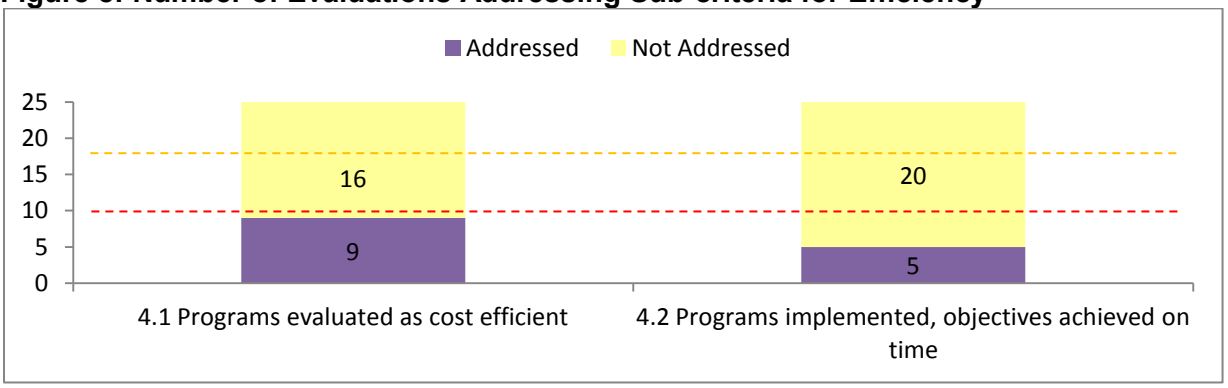
3.4 WHO evaluations did not address efficiency

3.4.1 Coverage

WHO evaluations generally did not address efficiency in the 2007 to 2010 time frame. For the two sub-criteria grouped under the overall heading of efficiency, the combination of a small sample size and few valid cases severely undercuts the validity of any general observation that can be made from the evaluations reviewed. Sub-criterion 4.1, "Programs evaluated as cost efficient," was addressed in only nine evaluation reports (36% of the sample). Only five evaluations addressed sub-criterion 4.2, "Program implementation and objectives achieved on time." Therefore, no results are presented for these sub-criteria.

MOPAN survey results and document review are not directly comparable with this review's criteria of efficiency. MOPAN does measure timeliness of implementation, but not the timeliness of the achievement of objectives.

Figure 8: Number of Evaluations Addressing Sub-criteria for Efficiency



3.4.2 Contributing Factors

Those evaluation reports that did address these sub-criteria most often reported factors detracting from efficiency. A common feature of these findings was a link between delays in program implementation and increased costs. Factors contributing to unsatisfactory results in program efficiency include:

- Inefficient and time-consuming procurement practices (1 evaluation);
- Poor ongoing monitoring of expenditures (1 evaluation);
- Delays in disbursement of funds by partner governments (1 evaluation);
- Lack of adequate training for logistical officers (1 evaluation);
- Delays in the approval process for transfer of funds to other UN partners (1 evaluation);
- Poor financial information, especially on the costs of operations of local partner organizations (1 evaluation); and
- Delays in mobilizing resources, including contracted personnel (1 evaluation).

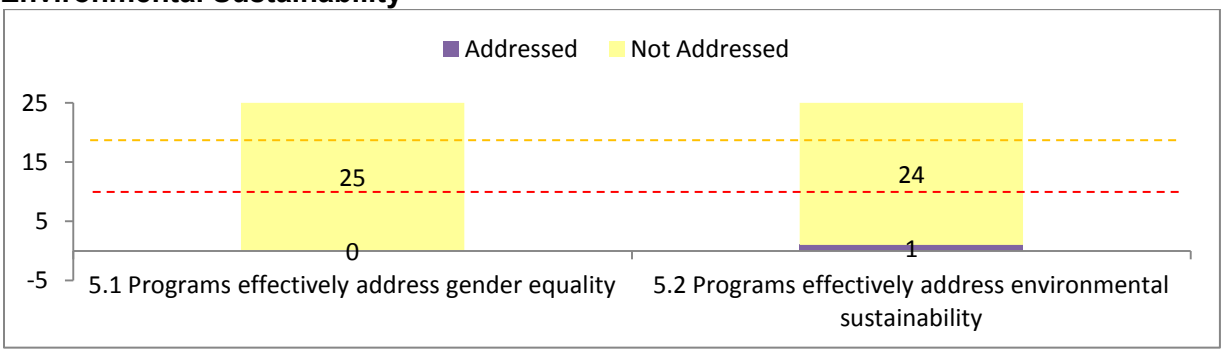
3.5 WHO evaluations did not address gender equality and environmental sustainability

3.5.1 Coverage

WHO evaluations did not regularly address effectiveness in supporting gender equality or environmental sustainability (Figure 9). Sub-criterion 5.1, "Programs effectively address gender equality," was not addressed in any of the 25 evaluation reports reviewed. Sub-criterion 5.2, "Changes are environmentally sustainable," was addressed in one evaluation report. Therefore, no results are presented for these sub-criteria.

The absence of gender equality as an issue in WHO evaluations represents a critical gap in effectiveness information for the organization.

Figure 9: Number of Evaluations Addressing Sub-criteria for Gender Equality and Environmental Sustainability



The absence of gender equality considerations in evaluations supplements the 2010 MOPAN study of the WHO, which rated integration of gender as strong in its document review but only adequate in survey responses. MOPAN noted "On WHO's integration of gender equality and human rights-based approaches, divergent ratings between the document review and survey suggest that while WHO has the policy frameworks and guidance required in its documents, it may not yet be applying these consistently in its programming work at all levels of the organization."[27]

This review's findings are in line with the WHO's own 2011 report of the baseline assessment of the WHO Gender Strategy, which found that less than 5% of planning officers "strongly" integrated gender into the monitoring and evaluation phases of WHO programming.[28]

On the crosscutting theme of the environment, MOPAN was more positive, noting that: "WHO's attempts to mainstream environment in its programmatic work were seen as adequate by

[27] MOPAN Common Approach Institutional Report for the World Health Organization (WHO) 2010 (Vol. I), Multilateral Organisation Performance Assessment Network, 2011, p. viii.
[28] *Gender mainstreaming in WHO: where are we now? Report of the baseline assessment of the WHO Gender Strategy*, WHO, 2011, p. 13.

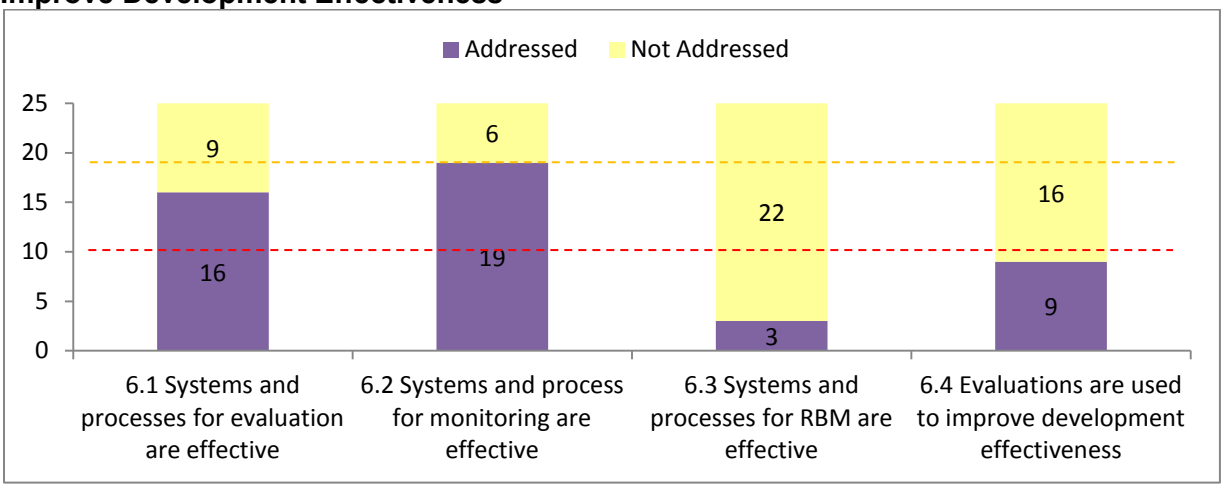
survey respondents and strong by the document review." This review is unable to provide results on environmental integration.

3.6 Evaluations report weaknesses in systems for monitoring and evaluation

3.6.1 Coverage of Sub-criteria

Some care is required in interpreting the results reporting regarding the use of monitoring and evaluation to improve effectiveness since two of the four sub-criteria were rated weak in coverage (Figure 10). Sub-criterion 6.1, "systems and processes for evaluations are effective," was addressed in 16 evaluation reports and rated moderate in coverage. Sub-criterion 6.2, "systems and processes for monitoring are effective," was addressed in 19 evaluations and rated strong in coverage. The last two sub-criteria, "systems and processes for results-based management are effective" and "evaluation results used to improve development effectiveness," were addressed in less than 10 evaluation reports and were both rated weak in coverage.

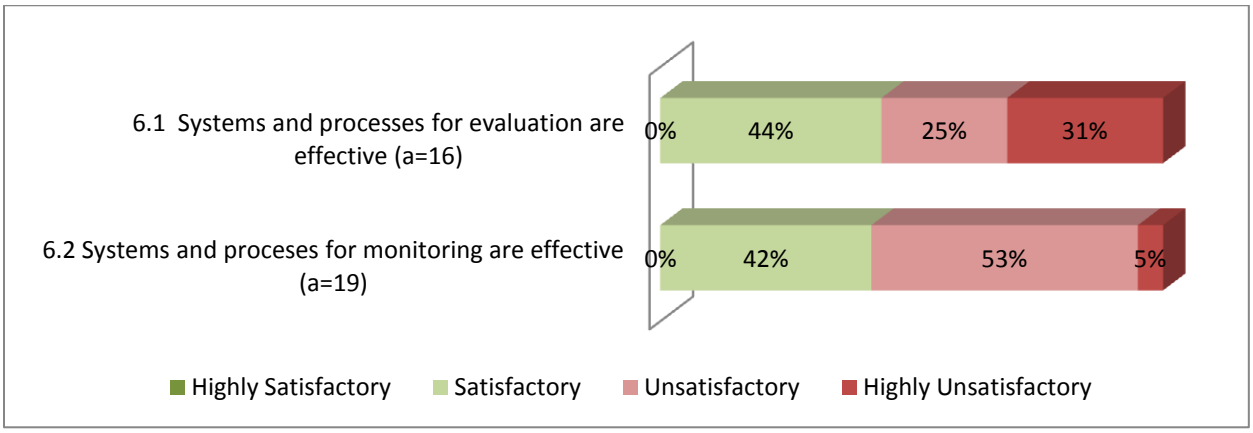
Figure 10: Number of Evaluations Addressing the Sub-criteria for Use of Evaluation to Improve Development Effectiveness



3.6.2 Key Findings

The WHO's systems and processes for using monitoring and evaluation to improve effectiveness was assessed as unsatisfactory. Evaluation reports that addressed sub-criterion 6.1, "systems and processes for evaluations are effective," often reported unsatisfactory findings, with only 7 of 16 evaluations (44%) producing findings classified as satisfactory or better. Similarly, only 8 of 19 evaluations (42%) reporting findings coded as satisfactory or better for sub-criterion 6.2, "systems and processes for monitoring are effective."

Figure 11: Using Evaluation and Monitoring to Strengthen Development Effectiveness
(Findings as percentage of number of evaluations addressing sub-criterion (= a), n = 25)



Findings from the MOPAN survey converge with findings of this review. The MOPAN report concluded, "the independence of the Office of Internal Oversight Services (OIOS) was considered adequate by survey respondents and the review of documents." However, other assessment findings suggest that the WHO's evaluation function should be strengthened: evaluation coverage is limited and difficult to ascertain because of the decentralised nature of evaluation; there is no repository of evaluations (although an inventory does exist) and evaluations are difficult to access through the WHO website." [29]

MOPAN survey results indicated the likelihood that programs would be subject to independent evaluation was near the bottom of the 'adequate' range. However, 40% of respondents answered 'don't know' to this question, and the document review rated the WHO as 'inadequate'.

MOPAN's document review rated as 'adequate' adjustments to strategies and policies as well as to programming on the basis of performance information, but noted "Although there are periodic evaluations of WHO programs (which assess the outcomes of the WHO's work along the lines of thematic, programmatic or country evaluations), the reports to the Executive Board do not seem to draw on the evaluation findings or recommendations."[30] Similarly, MOPAN found only one concrete example of performance information leading to adjustment to programming.

3.6.3 Contributing Factors

Three factors were cited as contributing to positive results in relation to the strength of evaluation and monitoring systems.

- A tradition of joint review of WHO programs involving the WHO, host governments and other stakeholders (3 evaluations);

[29] MOPAN Common Approach Institutional Report for the World Health Organization (WHO) 2010 (Vol. I), Multilateral Organisation Performance Assessment Network, 2011, p. ix.
[30] *Ibid,* p. 44.

- The practice of conducting regular or mid-term evaluations of new programs such as the introduction of a new vaccine (3 evaluations); and
- The practice of independent external evaluations of large WHO-supported programs (2 evaluations).

The most frequent critique of evaluation systems and procedures was that the evaluation reports did not refer to similar evaluations of the same program in the past or plans for the future (4 evaluations). In general, the evaluation reports did not include information that would allow the reviewer to place this particular evaluation in the context of a wider system or process calling for systematic evaluation of the programs under review.

Other factors that contributed to less-than-satisfactory results for the strength of evaluation and monitoring systems include:

- Institutional weakness among partners and, more specifically, failure to staff designated monitoring and evaluation positions which are a feature of program design requirements (4 evaluations);
- Missing data or a failure to collect agreed-upon data on a regular and reliable basis (6 evaluations); and
- A sense among some WHO partners that data requirements are overly bureaucratic and that the data is not being used, so less effort is put into data collection (1 evaluation).

Highlight Box 6 provides an illustration of how lack of resources and weak commitment to the requirement for results reporting (seen as overly bureaucratic) have undermined the effectiveness of monitoring and evaluation.

Highlight Box 6

Lack of resources and commitment to evaluation and monitoring by collaborating centres

With few exceptions, the lack of systematic monitoring and evaluation is obvious. The reasons mentioned are lack of manpower and interest, ambiguity of responsible technical officers about their role, and uncertainty regarding the role of regional Collaborating Centre focal points. The annual report submitted by Collaborating Centres is often perceived as a bureaucratic formality, rather than a useful instrument to assess progress and improve collaboration, especially in the case of active networks that have active ongoing monitoring and reporting mechanisms.

Thematic Evaluation of WHO's Work with Collaborating Centres

3.6.4 A look at WHO's 2012 evaluation policy

A brief comparison of WHO's 2012 evaluation policy to standards from the United Nations Evaluation Group highlights improvements and areas for continued attention at the WHO. The United Nations Evaluation Group is a network to bring together evaluation units in the UN system. Its standards for evaluation in the United Nations system describe, among other things, the expectations for evaluation policies in UN organizations. Table 3 compares the UNEG standard with the WHO's new evaluation policy.

Table 3: Comparing United Nations Evaluation Group Standards to WHO Evaluation Policy

UNEG requirement:	Addressed by WHO evaluation policy?
Clear explanation of the concept and role of evaluation within the organization	Yes
Clear definition of the roles and responsibilities of the evaluation professionals, senior management and programme managers	Yes, but not clearly—role of program managers described under utilization and follow-up section, with no clear accountability and oversight.
An emphasis on the need for adherence to the organization's evaluation guidelines	Partially—principles and norms of evaluation clearly identified, which contains a 'quality' sub-section with a reference "applicable guidelines." Although the new policy replaces its previous evaluation guidelines, WHO does not appear to have new evaluations guidelines yet.
Explanation of how evaluations are prioritized and planned	Partially—list of factors that will be considered in planning and prioritizing evaluations, but vague overall.
Description of how evaluations are organized, managed and budgeted	Partially—clear description of evaluation organization and management, but budgeting does not describe sources of funding.
An emphasis on the requirements for the follow-up of evaluations	Yes.
Clear statement on disclosure and dissemination	Partially—Statement provided, but vague, in stating "WHO shall make evaluation reports available in accordance with the Organization's disclosure policy." The reference to the disclosure policy confuses readers as to which evaluation reports will be disclosed.

As demonstrated in Table 3, gaps remain in the evaluation policy regarding the planning, prioritizing, budgeting and disclosure of WHO evaluations. Additionally, the WHO can further clarify the roles and responsibilities of programme managers regarding evaluations and specify the guidelines it will use to judge the quality of evaluations.

In 2012, the United Nations Joint Inspection Unit conducted a review of management, administration and decentralization in the World Health Organization.[31] The review raises concerns about independence and credibility of WHO evaluations, suggests that the WHO should have a stronger central evaluation capacity, and recommends that a peer review on the evaluation function be conducted by the United Nations Evaluation Group and be presented to the WHO Executive Board by 2014.

4.0 WHO and Canada's Priorities in International Development

In May 2009, the Minister of International Cooperation announced Canada's intention to focus its development assistance on three thematic priorities: increasing food security; stimulating sustainable economic growth; and securing the future of children and youth. The section first reviews Canada's relationship with the WHO, including management responsibility within CIDA, and then assesses the extent to which the WHO contributes toward Canada's priorities in international development, and to the implementation of CIDA's strategic objectives for engagement with the WHO.

Figure 12: Canada's Thematic Priorities

- Increase food security for the poor in those partner countries and regions where food security is identified as a key priority.
- Create sustainable economic growth that will increase revenue generation, create employment and lead to poverty reduction in developing countries.
- Support girls, boys, young women and young men to become healthy, educated, and productive citizens of tomorrow.

4.1 CIDA Support to the WHO

In 2010–2011, Canada, through CIDA, was the third-ranked country in terms of the amount of voluntary funding provided to the WHO, only the United States and the United Kingdom ranked higher[32].

With 284 million Canadian dollars of support provided by CIDA in the four fiscal years from 2007–2008 to 2010–2011, the WHO ranks eighth among multilateral organizations supported by CIDA in dollar terms. In the area of health, only the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) ranks higher with 450 million Canadian dollars of support from CIDA in the same period.

[31] Review of Management, Administration and Decentralization in the World Health Organization, JIU/REP/2012/6, United Nations Joint Inspection Unit, Geneva, 2012.
[32] CIDA Strategy for Engagement with the World Health Organization (WHO). p.2. CIDA, 2011.

The volume of CIDA's financial support to the WHO illustrates the organization's importance as a channel for Canadian development assistance. It also demonstrates Canada's importance to the WHO as a source of funding. Table 4 provides an overview of all forms of CIDA support to the WHO over the past four years.

Table 4: CIDA Support to WHO 2007/08 to 2010/11 ($ Canadian millions)

Branch and Type of Funding	2007–2008	2008–2009	2009–2010	2010–2011	Total
Multilateral and Global Programs Branch Initiative-specific funding	55.20	45.75	47.34	48.50	**196.79**
Other CIDA branches	14.53	34.85	27.28	10.45	**87.11**
Total	**69.72**	**80.60**	**74.62**	**58.95**	**283.89**

Source: Statistical Analysis and Reporting Section—Chief Financial Officer Branch (CIDA) (2011).

Notes: (1) Other CIDA branches includes: Partnerships with Canadians Branch, Strategic Policy and Performance Branch, and the Office for Democratic Governance.
(2) Long-Term Institutional Funding is not reported since CIDA did not provide this form of support to WHO during this period.

4.2 Managing CIDA's relationship with the WHO

Health Canada has the overall lead for the Government of Canada's engagement with the WHO, and is head of the Canadian delegation to the World Health Assembly. In 2009, Canada was elected to a three-year term on the Executive Board (until 2012).

Health Canada's strategic framework guides Canada's tenure on the Executive Board and its overall engagement with the WHO. It includes two broad objectives: fostering a stronger, more capable organization which responds more effectively to the evolving needs of member states; and, protecting the health and well being of Canadians, while advancing Canada's domestic, foreign and international development policy objectives related to global health.

CIDA's Strategy for Engagement with the WHO describes its role in managing this relationship in some detail[33]:

> "CIDA engages with WHO in two significant ways, policy dialogue and development assistance programming in such area as infectious diseases, child health, and humanitarian assistance. CIDA's longstanding and productive relationship with WHO includes partnering on many key initiatives such as the "three by five" Initiative, the Global Polio Eradication Initiative and the Global Drug Facility. WHO is also playing a key role in developing health indicators and data collection in support of the G8 in Muskoka Initiative on Maternal, Newborn and Child Health championed by Canada."

Multilateral and Global Programs Branch staff at CIDA's headquarters coordinate with staff at the Permanent Mission of Canada to the Offices of the United Nations in Geneva to ensure

[33] CIDA Strategy, p. 3.

consistent and continuous dialogue on policies and programmes. Most recently, Multilateral and Global Programs Branch staff have been active in promoting the ongoing process of reform at the WHO, with particular emphasis on improving evaluation and monitoring policies and practices.

4.3 Alignment with Canada's Thematic Priorities

4.3.1 Increasing Food Security

In line with WHO strategic objective 9, the organization engages in critical technical and normative work in support of improved nutrition, especially for pregnant and lactating women, for children and for those affected by crises. In addition, as noted in its Review of Newborn and Child Health programs in Cambodia, the WHO promotes early breastfeeding, continuation of breastfeeding for at least six months, and micronutrient fortification, including the use of iodized salt. Furthermore, as the head of the Global Health Cluster for humanitarian relief, the WHO provides guidelines on effective rations to be used in emergency food aid and in longer-term relief and rehabilitation operations. WHO expenditures under strategic objective 9 (nutrition and food safety) contribute directly to the goal of increasing food security in developing countries.

4.3.2 Stimulating Sustainable Economic Growth

The WHO contributes indirectly to the goal of stimulating sustainable economic growth. Through strategic objective 10 (health systems and services), it works to improve public health services in developing countries through better governance, financing, staffing and management. It also engages directly in the support of health information systems to provide reliable and accessible evidence as a sound basis for public health policy. In 2010, WHO spent 136 million USD in support of health systems and services (7.3% of overall expenditures).

Of course, expenditures in other areas of the WHO's mandate will also support sustainable economic growth by reducing working hours lost to illness and by reducing the burden of disease, which acts as a drag on productivity and drives people into poverty. These other areas of the WHO's work include: communicable disease (strategic objective 1); HIV/AIDS, tuberculosis, and malaria (strategic objective 2); chronic and non-communicable diseases (strategic objective 3); risk factors for health (strategic objective 6); social and economic determinants of health (strategic objective 7); and medical products and technologies (strategic objective 11). Together, these objectives, which accounted for over 1.1 billion USD in expenditures by the WHO in 2010 (61% of the total), make an important contribution to creating the necessary conditions for sustainable economic growth in many developing countries.

4.3.3 Securing the Future of Children and Youth

To the extent that WHO programs have been found to be effective in achieving development objectives, there can be no doubt of the organization's contribution to the Canadian priority of securing the future of children and youth. WHO expenditures under strategic objective 4 (child, adolescent, mother health and aging) are often focused on improving the health of children, adolescents and their mothers and thereby helping to secure their future. Similarly, expenditures

under strategic objective 9 (nutrition and food security) are often focused on the nutritional needs of children and their mothers.

Programming under strategic objective 1 (communicable diseases) includes extended programs of immunization that directly benefit children by providing protection from immuno-preventable infections. Strategy objective 2 (HIV/AIDS, tuberculosis and malaria) includes support to the prevention of mother to child transmission of HIV/AIDS as well as, of course, support to national programs of malaria and tuberculosis control.

Taken together, WHO expenditures under strategic objectives 1, 2, 4 and 9 accounted for 1.1 billion USD in 2010 (58% of the total).

Finally, the WHO engages regularly in important normative work on the health of children and adolescents, including, for example, the development and promotion of guidelines in the Integrated Management of Childhood Illness (IMCI). One such example is illustrated in Highlight Box 7.

Highlight Box 7

Addressing Child Mortality in Guyana

Under the national strategy supported by the WHO, the trend for under-five child mortality in the last decade in Guyana has been downwards. One result has been an improvement in access delivery facilities so that a high proportion of women deliver in facilities and are attended by a skilled attendant.

Review of Child Health in Guyana

4.4 How is the WHO Fulfilling the Strategic Objectives that CIDA Identified?

CIDA, in collaboration with other Canadian government departments, currently focuses on two strategic objectives for its engagement with the WHO.[34]

1. Strengthen CIDA's partnership with the WHO in support of development priorities in the area of health, including maternal and child health and nutrition; and

2. Support the WHO as a key institutional partner by fostering a stronger, more effective and efficient organization.

4.4.1 Encouraging WHO Action in Development Priorities in Health

Under the first of its two priorities for engaging with the WHO, CIDA has committed to supporting its mandate as the lead global agency in public health. CIDA will also encourage the organization to capitalize on its strengths in supporting Canada's development priorities, particularly in maternal and child health and nutrition. The evaluation reports reviewed do not directly address the question of whether the WHO is making the most of its strengths in the area

[34] CIDA Strategy, p. 6.

of maternal and child health and nutrition, although the organization does dedicate significant resources on programs in this area. Similarly, the profile of WHO budgets and expenditures indicates that the organization is active in CIDA's priority areas of human resources for health and strengthening health information and management systems.

4.4.2 Fostering a Stronger, More Effective and Efficient Organization

Interviews with CIDA staff and a review of the documents on reform of the WHO (including reports to and decisions of the Board in 2010, 2011 and 2012) indicate that CIDA has been very active in pursuing important organizational reforms at the level of the Executive Board and World Health Assembly. The proposals by the Director General and the subsequent decisions of the Executive Board at its special session on UN Reform (EBSS/2/DIV/2) in November 2011 indicate that important changes are being set in motion in relation to: programs and priority setting; governance; and managerial reforms. Under the heading of managerial reforms, detailed proposals were called for by the November special session in relation to:

- organizational effectiveness, alignment and efficiency;
- financing of the WHO;
- human resources policies and management;
- results-based planning, management and accountability (including independent evaluation); and
- strategic communications.

CIDA has been vocal in support of these reforms and will play an important role in following up on their effectiveness. It is too early to say if the proposals approved at the May 2012 meeting of the WHO's Executive Board will produce the expected improvements in overall organizational effectiveness. Nonetheless, they represent an important first step in realizing this important CIDA priority in its engagement with the organization.

5.0 Conclusions

5.1 Insufficient evidence available to make conclusions about the WHO

The major finding of this review is that the limited set of available and valid evaluation reports means that there is not enough information to draw conclusions about the WHO's development effectiveness.

The limited number of evaluation reports that are available provide insights into the effectiveness of those WHO programs. Results from the review of these evaluations are presented below, but cannot be generalized to the organization as a whole.

An analysis of the 2012 WHO evaluation policy indicates that while the approval of an evaluation policy represents a positive step, gaps remain in the policy regarding the planning,

prioritizing, budgeting and disclosure of WHO evaluations. In addition, the WHO could further clarify the roles and responsibilities of program managers regarding evaluations and provide guidance to judge the quality of evaluations.

A 2012 United Nations Joint Inspection Unit review also raises concerns about independence and credibility of WHO evaluations, suggests that the WHO should have a stronger central evaluation capacity, and recommends that a peer review on the evaluation function be conducted by the United Nations Evaluation Group and be presented to the WHO Executive Board by 2014.

Based on a systematic review of available evaluation reports, and the key findings and related contributing factors, this review concludes that:

1. ***Based on the limited sample available, WHO programs appear to be highly relevant to stakeholder needs and national priorities.*** Evaluations reported that WHO programs are well suited to the needs of stakeholders (16 of 18 or 89% rated satisfactory or better) and well aligned with national development goals (100% of 12 evaluations rated satisfactory or highly satisfactory). Further, the objectives of WHO-supported projects and programs remain valid over time (100% of 21 evaluations rated satisfactory or better). There is room, however, for better linking WHO program activities to their objectives during program design (60% of 20 evaluations rated satisfactory) and for more effective partnerships with governments (61% of 18 evaluations rated satisfactory or highly satisfactory). These high levels of positive relevance reported in evaluations indicate that the WHO has been able to ensure that its programs address important and enduring needs while remaining well aligned with national priorities in health.
2. ***The WHO appears to be effective in pursuing development objectives*** with 71% (15 of 21) of evaluations reporting performance as satisfactory or better. In addition, WHO programs generate positive benefits for target group members with 64% of 14 evaluations rating performance for this sub-criterion as satisfactory or highly satisfactory.
3. ***The benefits of assessed WHO programs appear to be sustainable but there are important challenges to institutional sustainability of program arrangements.*** The benefits of WHO programs are likely to be sustained with 73% of evaluations reporting satisfactory or highly satisfactory results in this area (although only 11 evaluations cover this criteria). The WHO does, however, face a challenge in the area of institutional capacity for sustainability. Only 37% (6 of 16) of evaluations found WHO programs were satisfactory in terms of providing support to local institutional capacity for sustainability.
4. ***No results to report on efficiency*** in the 2007 to 2010 time frame. Only a few evaluations reported on cost efficiency (9) and on whether implementation of programs and achievement of objectives was timely (5). The combination of a small sample size and few evaluations covering the efficiency sub-criteria precludes any general findings about efficiency. Evaluation reports that did address these sub-criteria most often reported factors detracting from efficiency. A common feature of these findings was a link between delays in program implementation and increased costs.

5. ***WHO's evaluation function has not adequately addressed effectiveness in supporting gender equality or environmental sustainability.*** This prevented the review from identifying any results in this area.

    i. The crosscutting theme of gender equality was not addressed in the WHO's evaluation reports, and represents a critical gap in effectiveness information for the organization.

    ii. Possibly because few WHO-supported programs directly impact environmental sustainability, only one evaluation addressed this crosscutting theme.

6. ***Evaluations reviewed have found WHO systems for evaluation and monitoring at country level to be unsatisfactory.*** A total of 56% of reported findings on the effectiveness of evaluation systems and processes were classified as unsatisfactory or highly unsatisfactory (9 of 16 evaluations). Similarly, systems for monitoring are unsatisfactory, with 58% (11 of 19) of evaluations' findings classified as unsatisfactory or highly unsatisfactory. In particular, the evaluations reviewed point to a lack of financial resources and trained local staff as important factors contributing to less than satisfactory results in the area of evaluation and monitoring. Where evaluation systems are reported as satisfactory, one contributing factor has been the tradition of joint review of program implementation by the WHO and its partners. Sub-criteria on effective systems and processes for RBM and evaluation results used to improve development effectiveness were addressed by only 3 and 9 evaluations, respectively. Therefore, no results are presented for these sub-criteria.

5.2 The WHO and Canada's international development priorities

The WHO contributes to Canada's Development Priorities. There is clear evidence that WHO programs make an important direct contribution to the Canadian international development priorities of food security (especially for pregnant and lactating women, for children and for those affected by crises) and securing the future of children and youth. There is also good evidence that WHO programs contribute indirectly to sustainable economic growth through the support of public health systems and by assisting developing countries to reduce the burden of communicable and non-communicable diseases.

5.3 Evaluation and reporting on Effectiveness

Evaluation and reporting on effectiveness should be strengthened. The review of the evaluation function carried out by the pilot test team and the gaps identified in this review demonstrate the need to strengthen the evaluation function within the WHO. In particular, there is a need to ensure adequate coverage of WHO programs and projects through a systematic and sufficiently-resourced evaluation function. In addition, there is a need to make evaluation reports produced by the WHO readily available to staff and to external stakeholders. An effort to improve the evaluation system resulted in the adoption of a new policy on evaluation by the Executive Board at its meeting in May 2012.

6.0 Recommendations for CIDA

This section contains the recommendations to CIDA based on the findings and conclusions of this effectiveness review of the WHO. Aimed at improving evaluation and results-based management at the WHO, these recommendations are in line with the objectives of Canada's existing engagements with the WHO. As one of several stakeholders working with the WHO, Canada's individual influence on the organization is limited, and it may need to engage with other shareholders to implement these recommendations (See Annex 8 for CIDA's management response.)

1. Canada should monitor efforts at reforming the evaluation function at the WHO as the new policy on evaluation is implemented. In particular, CIDA should use its influence at the Executive Board and with other donor agencies to advocate for a sufficiently resourced and capable evaluation function that can provide good coverage of WHO programming over time.
2. CIDA should monitor the implementation of the evaluation policy so that future WHO evaluations sufficiently address gender equality.
3. CIDA should encourage the WHO to implement a system for publishing regular (possibly annual) reports on development effectiveness that builds on the work of the reformed evaluation function. In general terms, there is a need to strengthen WHO commitment to reporting on the effectiveness of programs.
4. CIDA should encourage the WHO to systematically manage for results. The ongoing upgrading and further implementation of the Global Management System at the WHO may offer such an opportunity.

Annex 1: Criteria Used to Assess Development Effectiveness

Relevance of interventions

Sub-criteria
1.1 Programs are suited to the needs of target group members
1.2 Programs are aligned with national development goals
1.3 Effective partnerships with governments
1.4 Program objectives remain valid
1.5 Program activities are consistent with program goals

Achieving Development Objectives and Expected Results

Sub-criteria
2.1 Programs and projects achieve stated objectives
2.2 Positive benefits for target group members
2.3 Substantial numbers of beneficiaries

Sustainability of Results/Benefits

Sub-criteria
3.1 Program benefits are likely to continue
3.2 Programs support institutional capacity for sustainability

Efficiency

Sub-criteria
4.1 Programs evaluated as cost-efficient
4.2 Program implementation and objectives achieved on time

Cross Cutting Themes: Inclusive Development Which can be Sustained (Gender Equality and Environmental Sustainability)

Sub-criteria
5.1 Programs effectively address gender equality
5.2 Changes are environmentally sustainable

Using Evaluation and Monitoring to Improve Development Effectiveness

Sub-criteria
6.1 Systems and processes for evaluation are effective
6.2 Systems and processes for monitoring are effective
6.3 Systems and processes for RBM are effective
6.4 Evaluation results used to improve development effectiveness

Annex 2: Evaluation Sample

#	Year	Title	Type
1	2010	Making Pregnancy Safer Department	HQ
2	2009	Central African Republic—The Surveillance assessment	EPI
3	2009	Democratic Republic of Congo—The Surveillance assessment	EPI
4	2009	Joint field mission to study WHO Disaster Preparedness and Response in the occupied Palestinian territory	HAC
5	2009	Democratic Republic of Congo—Post introduction Evaluation (PIE) of new vaccines—Pentavalent	EPI
6	2009	Cameroon—The Surveillance assessment	EPI
7	2009	Vietnam—EPI assessment	EPI
8	2009	Sierra Leone—Yellow Fever and Measles Vaccination Campaign, Rapid Evaluation of Vaccine Coverage Using the Lot Quality Assurance Sampling (LQAS) Methodology	EPI
9	2009	Independent evaluation of major barriers to interrupting Poliovirus transmission	HQ
10	2009	Zambia—Post-introduction Evaluation (PIE) of new vaccines switch from lyophilized to liquid Pentavalent	EPI
11	2009	Philippines—Vaccine procurement assessment	EPI
12	2008	Independent Evaluation of the Stop TB Partnership—McKinsey & Co.	Stop TB Partnership
13	2008	Review of Acute Flaccid Paralysis (AFP) Surveillance in Afghanistan	Polio Eradication Team, EMRO
14	2008	Review of the Nutrition Programmes of the WHO in the context of current global challenges and the international nutrition architecture	NHD

#	Year	Title	Type
15	2008	Short review of Newborn and Child Health Programme, Nepal	SEARO & CAH/HQ
16	2008	Joint field mission to study WHO Disaster Preparedness and Response in the context of the Health Cluster response to cyclone Nargis in Myanmar	HAC
17	2008	External Disease Surveillance Review, Ethiopia	VPD/AFRO
18	2008	Short Program Review for Child Health in Guyana	AMRO, CAH/HQ
19	2008	Review of Newborn and Child Health Program, Cambodia	WPRO & CAH/HQ
20	2007	Assessment of implementation, impact and process of WHO Medicines Strategy, 2004–2007	Org-wide
21	2007	Health Action in Crisis Institutional building program (The 3-year program)	Org-wide
22	2007	Programmatic Evaluation—Selected Aspects of the PHE [Public Health and Environment] Department (HQ)	Global thematic
23	2007	Thematic Evaluation of the WHO's Work with Collaborating Centres	Global thematic
24	2007	Health Action in Crisis' work in Africa	AFR
25	2007	Evaluation of Second Generation HIV Surveillance, Nepal	Country thematic

Annex 3: Approach and Methodology

This Annex provides a more thorough explanation of the key elements of the methodology used for the review of the development effectiveness of the WHO. It is structured around the sequence of tasks undertaken during the review: determining the rationale for the review; drawing the sample of evaluations; undertaking the process of review and controlling for quality during the analysis phase; and assessing the level of coverage provided by the development effectiveness review. The review team also carried out an interview with staff of the Office of Internal Oversight Services (OIOS) at the WHO to understand better the universe of available WHO evaluation reports and to put in context the changing situation of the evaluation function.

This review of evaluation reports was supplemented by a review of WHO corporate documents related to evaluation and reporting on development effectiveness and a consultation with CIDA manager responsible for managing relations with the WHO. [35] These were done to contextualize the results of the review and to take account of advances since the pilot test was carried out in 2010. A list of the documents consulted is provided in Annex 5.

Relevant WHO corporate documents include the reports of the Programme, Budget and Administration Committee to the Executive Board, Reports on WHO Reform by the Director-General, Evaluation Policy Documents, Annual Reports and the Interim Assessment of the Medium-Term Strategic Plan. These documents allowed the review team to assess the ongoing evolution of evaluation and results reporting at the WHO and to put in context the findings reported in the evaluation reports.

Rationale for a Development Effectiveness Review

The common approach and methodology offer a rapid and cost-effective way to assess the development effectiveness of the multilateral organization relative to a more time-consuming and costly joint evaluation. The approach was developed to fill an information gap regarding the development effectiveness (development effectiveness) of multilateral organizations. Although these multilateral organizations produce annual reports to their management and/or boards, bilateral shareholders were not receiving a comprehensive overview of the performance on the ground of these organizations. The Multilateral Organization Performance Assessment Network (MOPAN) seeks to address this issue through organizational effectiveness assessments. This approach complements MOPAN's assessments.

The approach suggests conducting a review based on the organization's own evaluation reports when two specific conditions exist:[36]

[35] The reviewers note that future reviews could benefit from interviews to provide context and additional information.

[36] *Assessing the Development Effectiveness of Multilateral Organizations: Approach, Methodology and Guidelines*, Management Group of the Task Team on Multilateral Effectiveness, DAC EVALNET, 2011.

1. There is a need for field-tested and evidence-based information on the effectiveness of the multilateral organization.
2. The multilateral organization under review has an evaluation function that produces an adequate body of reliable and credible evaluation information that supports the use of a meta-evaluation methodology to synthesize an assessment of the organization's effectiveness.

In reference to condition number one, the Medium-Term Strategic Plan Interim Assessment Report does provide some insight into how member states perceive elements of the WHO's performance. However, a survey of perceptions does not replace a regular, evidence-based report on the development effectiveness of WHO activities and programs. As a result, condition 1, the need for more evidence-based, field-tested information on development effectiveness is met in the case of the WHO.

Results for the second test, discussed below, were more marginal. The supply of reasonable quality evaluation reports available at the time of the pilot test was limited, with only 25 such evaluations provided by the WHO to the pilot test team. The development effectiveness review was completed because these 25 evaluations were able to provide moderate coverage of four of the six main criteria used to assess development effectiveness. However, this narrow supply of reasonable evaluations limits the extent to which the results are generalizable across the organization.

The WHO's Evaluation Function (Quantity and Quality)

Quantity of Evaluations

When the pilot test was carried out in 2010, evaluations commissioned by the WHO were not published either as paper documents or electronically. Therefore, a request was made to the OIOS to identify and source the evaluations for the pilot test. The OIOS indicated that in most cases, evaluation reports were owned by both the WHO and the countries covered in the evaluation. Copies would need to be requested from the WHO Country Offices in question. There was no central repository of published evaluations available in either electronic or hard copy form.

For the 2007 to 2010 period, just 58 studies were identified as the universe of evaluations, and the WHO was able to provide 34 of these studies to the evaluation team in the available time frame. Only 25 of the 34 studies provided were evaluation studies in the sense that they presented evaluation findings relevant to the criterion under assessment.

Quality of Evaluations

WHO evaluation reports were often operationally and technically focused; that is, they were often concerned with how well a given service delivery method, surveillance system, or even the introduction of a new vaccine, was implemented rather than the resulting changes for the target or beneficiary group. This is a significant problem for assessing evaluation quality because these studies (while often well designed within their own parameters) often lacked key

components of a quality evaluation when assessed against the quality criteria derived from UNEG standards.

The WHO has not yet been the subject of a DAC/UNEG Professional Peer Review of the Evaluation Function, and the review team did not undertake such a review. However, the review team conducted its own quality review of the evaluations for inclusion in this report.

The results of the review team's own quality analysis were mixed, with 52% of reviewed evaluations scoring 30 points or more, and 24% receiving scores of less than 19 out of a possible 48. For quality criteria I, "evaluation findings are relevant and evidence based," only three evaluation reports scored less than three from a possible score of five. (See Annex 4 for the evaluation quality scoring grid). All evaluations were retained for the review since scores overall were reasonable.

On balance, however, the set of available evaluation reports does not provide enough coverage of WHO programs and activities in the period to allow for generalization of the results to WHO programming as a whole. The 25 available evaluation reports do provide insights into the development effectiveness of WHO programs evaluated during the period.

WHO Reporting on Development Effectiveness

The WHO does not prepare an annual report on development effectiveness or an annual summary of the results of evaluations. It does provide, however, extensive reporting on the global and regional situation in health to the World Health Assembly each year. It also presents special reports on specific global topics and challenges in public health on an annual basis.

For some time now, the WHO has been in the process of implementing a system of results monitoring and reporting based on Oracle software. This Global Management System was in development as early as 2008 and is currently being upgraded after a lengthy implementation phase. One goal of the Global Management System is to align program and project planning, implementation and monitoring with agency strategic objectives at a corporate, regional and national level.

Since 2008, the WHO has made an effort to implement the System in each of its regions and by January 2011 was able to report to its Executive Board that it had made "considerable progress" in implementing the system in five regions and at headquarters. The Executive Board (EB128/3) welcomed the reported progress but expressed concern that the Region of the Americas/PAHO had chosen not to implement the system.

In May 2011, the Secretariat at the WHO reported to the Programme, Budget and Administration Committee of the Executive Board on progress in implementing the Global Management System. The Committee in its report to the Executive Board noted that:[37]

[37] Report of the Programme, Budget and Administration Committee of the Executive Board (EB129/2). WHO, 2011.

> "The Global Management System had been successfully rolled out in the Africa Region. Questions were asked regarding the planned upgrade of the System and its related cost as well as the savings that will result from its implementation. Queries were also raised with regard to harmonization between the Global Management System and the new system in the Region of the Americas/PAHO."

Available documentation on the System suggests its primary focus is still finance, administration, resource allocation planning, and human resources management. It is not yet clear if the System, as implemented, will effectively strengthen the results management and reporting system at the WHO.

At its Special Session on WHO reform in November 2011, the Executive Board welcomed the Director General's proposals on managerial reform and requested that these proposals be taken forward in several areas, including the improvement of monitoring and reporting.[38] As already noted, one consequence of this request was the proposal for a new policy on evaluation, which was officially adopted by the Executive Board at its meeting in May 2012. It is not yet clear whether this will include an effort to strengthen reporting on the development effectiveness of WHO programs, beyond that expected from the full implementation of the Global Management System.

Finally, it should be noted that the WHO published an interim assessment of progress toward meeting the goals of the Medium Term Strategic Plan in 2011.[39] This report was based on a survey of all member states and associate members carried out between November 2010 and March 2011. Responses were received from 104 member states and one associate member.

The assessment included questions for members and associate states in five topic areas:

1. The overall health situation and trends;
2. National policies and health systems;
3. Cooperation and collaboration with partners;
4. Mobilization and management of resources; and
5. Adequacy of cooperation with the WHO.

For most questions, the surveyed countries reported on progress they had made (with WHO support) toward achieving the objectives set out in the Medium-Term Strategic Plan. They also reported (under topic five) on the adequacy of WHO performance in each of six areas of work set out in its eleventh program of work:[40]

1. Providing technical support;
2. Providing leadership and engaging partnership;
3. Setting norms and standards;

---

[38] Decisions. Executive Board Special Session on WHO Reform (EBSS/2/DIV/2). WHO 2011.
[39] Medium Term Strategic Plan 2008 – 2013 Interim Assessment. WHO. 2011.
[40] Engaging for Health: Eleventh General Programme of Work 2006 – 2015: A Global Health Agenda. WHO, 2006.

4. Monitoring health situation and assessing trends;
5. Shaping research agenda and disseminating knowledge; and
6. Articulating ethical, evidence-based policy options.

The overall response to the question of WHO performance was positive, with 92% of respondents assessing the WHO's contribution as either meeting or above expectations and 8% as below expectations.[41] On the other hand, performance was rated as below expectations by both lower- and upper-middle-income countries, particularly in addressing needs related to four strategic objectives:

- Chronic non-communicable diseases (strategic objective 3);
- Social and economic determinants of health (strategic objective 7);
- Health systems and services (strategic objective 10); and
- Medical products and technologies (strategic objective 11).

The Medium-Term Strategic Plan Interim Assessment provides an insight into how member states perceive elements of the WHO's performance. However, a survey of this type cannot replace a regular, evidence-based report on the development effectiveness of WHO activities and programs.

Selecting the Evaluation Sample

Identifying and Obtaining Evaluations

The WHO's evaluation reports are not published and are not distributed electronically. Therefore, a request was made to the OIOS, which is responsible for the evaluation function, to identify and source the evaluations for the pilot test in 2010. OIOS staff indicated that there was no central repository of published evaluations available in either electronic or hard copy form. The OIOS also indicated that in most cases, evaluation reports were owned by both the WHO and the countries covered in the evaluation. Copies would need to be requested from WHO Country Offices in question. During a conference call, OIOS staff explained some of the difficulties of obtaining host country approval for the release of the evaluations in question.

For the 2007 to 2010 period, just 58 studies were identified as the universe of evaluations, and WHO was able to provide 34 of these studies to the evaluation team in the available time frame. Only 25 of the 34 studies provided were evaluation studies in the sense that they presented evaluation findings relevant to the criterion under assessment.

The OIOS noted that the practice of conducting country program evaluations was stopped at the WHO prior to 2006 and replaced by a series of internal audits. The pilot test team reviewed the 2009 Report of the Internal Auditor and noted that the audits of country offices concentrated on assessing risk management measures, including administrative, financial and programmatic risks, and, as such, did not represent country program evaluations suitable for this review.

[41] MTSP Interim Assessment. p.x

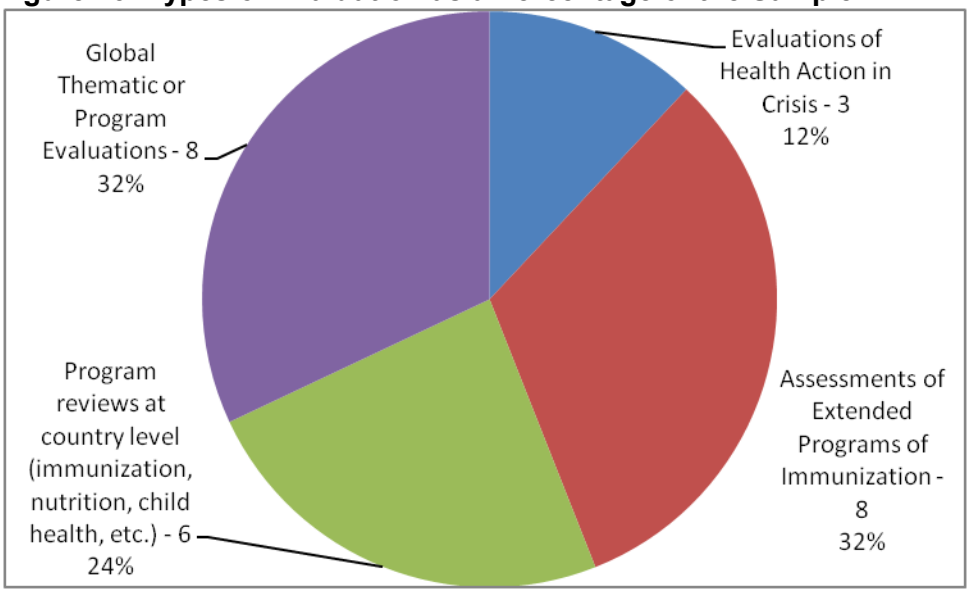
As a result, the set of available evaluations for review was in fact very small. The set of WHO evaluations reviewed should be seen not as a sample but as a census of the evaluation reports available and suitable for review at the time of the test. The list of evaluations is provided in Annex 2.

Evaluation Coverage of WHO Programming

The 25 evaluations available for the pilot test of the WHO do not provide adequate coverage of the over 4.5 billion USD in programming available to the WHO for expenditures over a two-year period. Nonetheless, there are several points of congruence between the sample and the profile of WHO budgeting. Figure 13 describes the types of WHO evaluations, which include:

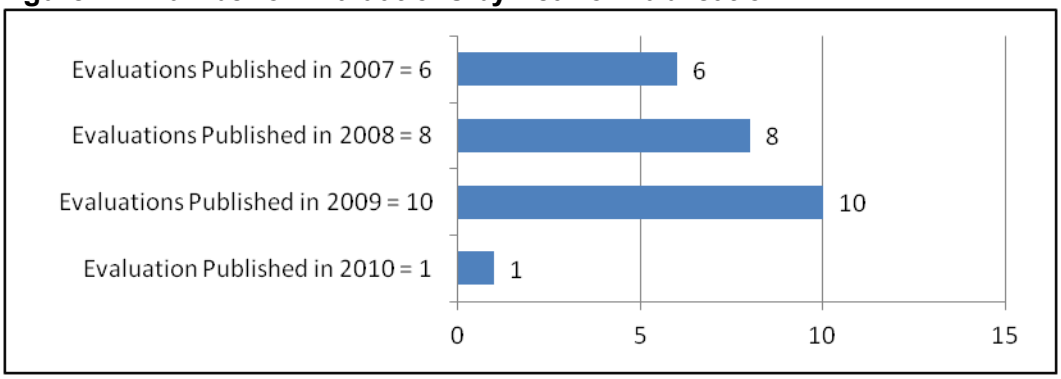
- 8 of the 25 evaluations deal with the implementation of Extended Programs of Immunization in a range of countries (Central African Republic, the Democratic Republic of Congo, Cameroon, Vietnam, Sierra Leone, Zambia and the Philippines). These programs directly contribute to strategic objective 1, communicable diseases, which is the most significant in dollar terms;
- 3 of the 25 evaluations deal with Health Action in Crisis at the regional or country level: 1 for Africa, 1 for Myanmar, and 1 for Palestine. (In addition, a program evaluation of Health Action in Crisis is included in the global category below.) These programs contribute to the third-largest strategic objective in terms of funding: emergencies and disasters; and
- A significant number of the evaluations reviewed are global or organizational in scope. They include:
    1. Evaluation of the Making Pregnancy Safer Department (2010);
    2. Independent evaluation of major barriers to interrupting Poliovirus transmission (2009);
    3. Independent Evaluation of the Stop TB Partnership (2008);
    4. Review of the Nutrition Programs of the WHO in the context of current global challenges and the international nutrition architecture (2008);
    5. Assessment of the Implementation, Impact and Process of WHO Medicines Strategy (2007);
    6. Health Actions in Crisis Institutional Building Program Evaluation (2007);
    7. Programmatic Evaluation of Selected Aspects of the Public Health and Environment (PHE) Department (2007); and
    8. Thematic Evaluation of the WHO's Work with Collaborating Centres (2007).

Figure 13: Types of Evaluation as a Percentage of the Sample



The evaluations covered in this review were all published by the WHO in the period from early 2007 to mid-2010 when the review was carried out. The number of evaluations published in each of these years is demonstrated by Figure 14.

Figure 14: Number of Evaluations by Year of Publication



In summary, while the list of suitable evaluations for review obtained from the WHO by the pilot test team cannot be easily compared to the geographic and programmatic distribution of activities, it does have some interest as a body of evaluation material on development effectiveness. For that reason, and to learn what lessons could be drawn from the experience of conducting the evaluation review, the team proceeded with the pilot test on the WHO.

The problem of a small sample of evaluations is further compounded by the large number of evaluation reports that do not address important criteria of development effectiveness. The net result is that for many of the criteria chosen to assess development effectiveness, the number of observations is too small to allow for general findings or lessons. In these instances, no findings are reported.

The Review Process and Quality Assurance

The review itself was conducted by a team of four analysts and a team leader. A two-day training session was held for analysts to build a common understanding of the review criteria. Following, the analysts and team leader conducted a pre-test to independently review two evaluations. The team compared their ratings from these two evaluations and developed a common agreement on the classification of results for all sub-criteria. This process helped to standardize classification decisions made by the analysts. During the review of evaluations, analysts conferred regularly over any classification issues that arose.

Once the reviews were completed, the team leader reviewed the coded findings and carefully examined the cited evidence and contributing factors. Based on this examination, the team leader made a small number of adjustments to the coded findings. The process of training, testing and monitoring the evaluation review process minimized any inter-analyst reliability issues and controlled for bias on the part of any one reviewer.

Review Coverage of Development Effectiveness criteria

In order to assess the level of coverage of a given sub-criterion, the review team developed ranges that defined coverage as **strong** when the number of evaluations addressing the criteria (*a*) was in the range from 18 to 25. Criteria where *a* was between 10 and 17 were rated as **moderate** in coverage. Finally, criteria addressed by less than 10 evaluations were rated as **weak** in coverage.

Table 5: Levels of Coverage for Each Assessment Criteria and Sub-Criteria

Relevance of interventions

Sub-criteria	*a**	Coverage Level**
1.1 Programs are suited to the needs of target group members	18	Strong
1.2 Programs are aligned with national development goals	12	Moderate
1.3 Effective partnerships with governments	18	Strong
1.4 Program objectives remain valid	21	Strong
1.5 Program activities are consistent with program goals	20	Strong

Achieving Development Objectives and Expected Results

Sub-criteria	*a**	Coverage Level**
2.1 Programs and projects achieve stated objectives	21	Strong
2.2 Positive benefits for target group members	14	Moderate
2.3 Substantial numbers of beneficiaries	8	Weak

Sustainability of Results/Benefits

Sub-criteria	*a**	Coverage Level**
3.1 Program benefits are likely to continue	11	Moderate
3.2 Programs support institutional capacity for sustainability	16	Moderate

Efficiency

Sub-criteria	*a**	Coverage Level**
4.1 Programs evaluated as cost-efficient	9	Weak
4.2 Program implementation and objectives achieved on time	5	Weak

Cross Cutting Themes: Inclusive Development Which can be Sustained (Gender Equality and Environmental Sustainability)

Sub-criteria	*a**	Coverage Level**
5.1 Programs effectively address gender equality	0	Weak
5.2 Changes are environmentally sustainable	1	Weak

Using Evaluation and Monitoring to Improve Development Effectiveness

Sub-criteria	*a**	Coverage Level**
6.1 Systems and processes for evaluation are effective	16	Moderate
6.2 Systems and processes for monitoring are effective	19	Strong
6.3 Systems and processes for RBM are effective	3	Weak
6.4 Evaluation results used to improve development effectiveness	9	Weak

- **a* = number of evaluations addressing the given sub-criteria
- ** Coverage Level defined as: Strong: *a* = 18 – 25, Moderate: *a* = 10 – 17, Weak: *a* = under 10

Of the 18 sub-criteria, only 6 received valid findings in 18 or more of the evaluation reports and are rated strong in coverage. Another 5 received valid findings in the moderate range (10 to 17 evaluation reports). A total of 7 sub-criteria, including all those relating to gender equality and efficiency, were adequately addressed in less than 10 evaluations and received weak ratings.

Annex 4: Evaluation Quality—Scoring Guide and Results

	Component of UNEG Standards on Evaluation to be Scored	Points Available
A	**Subject to be evaluated is clearly described.** Evaluation report describes the activity/program being evaluated, expected achievements, how the development problem would be addressed by the activity and the implementation modalities used.	**4**
B	**Purpose and context of the evaluation is clearly stated.** Evaluation report describes why the evaluation is being done, what triggered it (including timing in the project/program cycle) and how it will be used.	**3**
C	**Evaluation objectives are realistic and achievable.** Evaluation objectives follow directly from the stated purpose of the evaluation. They are clear and report notes agreement from key stakeholders.	**3**
D	**Scope of the evaluation is clearly defined.** The report defines the boundaries of the evaluation in terms of coverage of time period, phase of implementation, geographic area and dimensions of stakeholder involvement being examined. Limitations of evaluation scope are also noted.	**5**
E	**Evaluation criteria used to assess the subject to be evaluated are clearly spelled out in the evaluation report.** Normally these would include most of the following: • Relevance • Objectives Achievement • Efficiency • Impacts • Sustainability	**5**
F	**Evaluation methodologies chosen are sufficiently rigorous to assess the subject and to ensure a complete, fair and unbiased assessment.** The evaluation report clearly describes the methods chosen and the data sources used. Different sources of information are used to ensure accuracy, validity and reliability. All affected stakeholders are considered. Methodology addresses issues of gender and participation of under-represented groups.	**5**
G	**Evaluation methodologies are appropriate to the criteria being addressed.** The evaluation report describes the suitability of evaluation methods used to address the main evaluation criteria. Where samples are relied on, the sample chosen is described and its reliability and validity assessed. The use of qualitative and/or quantitative methods is identified and strengths and weaknesses discussed. Where impacts are assessed (using either quantitative or qualitative methods), a theory of how impacts are expected to occur (theory-based approach) is described and/or a counterfactual is presented.	**5**
H	**The evaluation acknowledges the limitations of the methodologies chosen.** The evaluation report includes an assessment of the limits of the methodologies chosen from a design perspective (prior to implementation). The report also describes any limitation arising from the experience of the evaluation team in implementing the chosen methodologies. The report provides an overall assessment of the appropriateness of the methodologies chosen based on the limitations noted.	**5**
I	**Evaluation findings and conclusions are relevant and evidence-based.** The report includes evaluation findings relevant to the assessment criteria (including issues and questions) specified. Evaluation findings are supported by clearly presented evidence resulting from an analysis of data	**5**

	Component of UNEG Standards on Evaluation to be Scored	Points Available
	gathered from the chosen methodologies. Conclusions are linked to the evaluation findings as reported.	
J	**Evaluation recommendations follow clearly from stated conclusions.** Evaluation recommendations are clearly related to the conclusions stated in the report.	**3**
K	**There is an explicit response from governing authorities and management.** The evaluation report, in either the report or an annex, includes a response from management. This response and the report have been circulated at the governance (Board) level (if indicated). Management response indicates a commitment to implement some or all recommendations. A time frame for implementation is stated. This response may be outside the evaluation report for most evaluations but is sometimes included.	**5**
	Total	**48**

Evaluation Quality Scoring Results

During the pilot test, the Management Group of participating development agencies guiding the work on behalf of DAC-EVALNET suggested grouping quality score results for each evaluation into groups of five (in total score). This was seen as presenting the best level of "granularity" and transparency. It allows independent observers to reach their own conclusions on the distribution of quality scores.

Table 10: Evaluation Quality Scoring Results

Evaluation Quality Scores in Groups of 5 (Max = 45)	Evaluations in Each Bracket (#)	Evaluations in Each Bracket (%)
43–48	1	4%
37–42	2	8%
31–36	10	40%
25–30	4	16%
19–24	2	8%
13–18	6	24%
7–12	0	0%
0–6	0	0%
Total	25	100%

Annex 5: Guide for Review Team to Classify Evaluation Findings

1. Relevance

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
1.1 Multilateral Organization (MO)-supported programs and projects are suited to the needs and/or priorities of the target group	The evaluation report indicates that substantial elements of program or project activities and outputs were unsuited to the needs and priorities of the target group.	The evaluation report indicates that no systematic analysis of target group needs and priorities took place during the design phase, or the evaluation reports some evident mismatch between program and project activities and outputs and the needs and priorities of the target group.	Evaluation report finds that the MO-supported activity, program or project is designed taking into account the needs of the target group as identified through a process of situation or problem analysis and that the resulting activities are designed to meet the needs of the target group.	Evaluation report identifies methods used in project development to identify target group needs and priorities (including consultations with target group members), and finds that the program and project takes those needs into account and is designed to meet those needs and priorities (whether or not it does so successfully.
1.2 MO-supported projects and programs align with national development goals:	The evaluation reports that significant elements of MO-supported program and project activity run counter to national development priorities with a resulting loss of effectiveness, overlap or duplication of effort.	The evaluation reports a significant portion (1/4 or more) of the MO-supported programs and projects subject to the evaluation are not aligned with national plans and priorities, but there is no evidence that they run counter to those priorities or result in overlap and duplication.	Most MO-supported programs and projects are reported in the evaluation to be fully aligned with national plans and priorities as expressed in national poverty eradication and sector plans and priorities. Wherever MO-supported programs and projects are reported in the evaluation as not directly supportive of national plans and priorities, they do not run counter to those priorities or result in overlap and duplication.	All MO-supported projects and programs subject to the evaluation are reported in the evaluation to be fully aligned to national development goals as described in national and sector plans and priorities, especially including the national poverty eradication strategy and sector strategic priorities.
1.3 MO has developed an effective partnership with governments, bilateral and multilateral development organizations and NGOs for	The evaluation report indicates that the MO experiences significant divergence in priorities from those of its (government, NGO or donor)	The evaluation reports that the MO has experienced significant difficulties in developing an effective relationship with partners and	The evaluation reports that the MO has improved the effectiveness of its partnership relationship with partners over time during the evaluation	The evaluation reports that the MO has consistently achieved a high level of partnership during the evaluation period.

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
planning, coordination and implementation of support to development	partners and lacks a strategy or plan which will credibly address the divergence and which should result in strengthened partnership over time.	that there has been significant divergence in the priorities of the MO and its partners.	period and that this partnership was effective at the time of the evaluation or was demonstrably improved.	
1.4 Objectives of MO-supported programs remain valid	The evaluation reports that either a significant number of sub-objectives or some of the most important objectives of MO-supported programs and projects are no longer valid to the needs and priorities of the target group at the time of the evaluation and that this raises important concerns regarding effectiveness.	The evaluation reports that, while the majority of the objectives of MO-supported programs and projects remain valid in terms of addressing target group needs and priorities, some objectives and/or sub-objectives are no longer valid. Nonetheless, the evaluation reports that the most important objectives remain valid.	The evaluation reports that, while no systematic effort has been made by MO-supported programs and projects to assess and adjust program objectives in order to confirm their validity, the objectives do remain valid in terms of addressing target group needs and priorities.	The evaluation reports that the MO-supported programs and projects subject to evaluation have carried out a systematic review of the continued validity of program objectives, and have either confirmed validity or made appropriate adjustments to the objectives.
1.5 Activities and outputs are consistent with program goal and with objectives achievement	The evaluation report finds that there are serious deficiencies in the causal link between the activities and outputs of MO-supported projects and programs and their objectives. This can occur either because the linkages are weak or non-existent or because the scale of activities and outputs is not matched to the scale of the objectives to be achieved. Note: the evaluation should recognize that not all project and program inputs will be provided by the MO in joint and country-led projects and programs.	The evaluation report is not able to verify that the design of MO-supported programs and projects includes a systematic assessment of causal linkages between program activities and outputs and objectives achievement. Nonetheless, there is no indication that these links do not exist in the program as implemented.	The evaluation report notes that the activities and outputs of MO-supported programs and projects are clearly linked to a causal process that should logically contribute significantly to the achievement of stated objectives. However, the scale of the activities and outputs is either not described or is inconsistent with the contribution to achieving the stated objectives.	The evaluation report notes that the activities and outputs of MO-supported programs and projects are clearly linked to a causal process that should logically contribute to the achievement of stated objectives. Further, the scale of the activities and outputs is consistent with the expected contribution to achieving the objectives as stated or the MO makes a significant contribution to overall strategy in the sector.

2. Achievement of Development Objectives and Expected Results

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
2.1 MO-supported programs and projects achieve their stated objectives and attain expected results.	The evaluation reports that one or more very important output and outcome level objective(s) has not been achieved.	The evaluation reports that half or less than half of stated output and outcome level objectives are achieved.	The evaluation reports that MO-supported programs and projects either achieve at least a majority of stated output and outcome objectives (more than 50% if stated), or that the most important of stated output and outcome objectives are achieved.	The evaluation reports that MO-supported programs and projects achieve all or almost all significant development objectives at the output and outcome level.
2.2 MO-supported programs and projects have resulted in positive changes for target group members.	The evaluation reports that problems in the design or delivery of MO-supported activities mean that expected positive impacts have not occurred or are unlikely to occur.	The evaluation report finds that it is not possible to make a credible assessment of program impacts because the program design did not specify intended impacts. If credible data is available and the design specifies impacts but sufficient time has not passed for expected impacts to emerge, this should be coded not addressed.	The evaluation report finds that MO-supported projects and programs have resulted in positive changes experienced by target group members (at the individual, household or community level).	The evaluation report finds that MO-supported projects and programs have resulted in widespread and significant positive changes experienced by target group members, as measured using either quantitative or qualitative methods (possibly including comparison of impacts with non-program participants).
2.3 MO programs and projects made differences for a substantial number of beneficiaries.	Evaluation finds that MO-supported projects and programs have not contributed to positive changes in the lives of beneficiaries, as measured quantitatively or qualitatively.	Evaluation finds that MO-supported projects and programs have contributed to positive changes in the lives of only a small number of beneficiaries (when compared to project or program targets and goals if established).	Evaluation finds that MO-supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries, as measured quantitatively or qualitatively.	Evaluation finds that MO-supported projects and programs have contributed to positive changes in the lives of substantial numbers of beneficiaries and accounting for most members of the target group, as measured quantitatively or qualitatively.

3. Sustainability

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
3.1 Benefits continuing or likely to continue after project or program completion	The evaluation finds that there is a very low probability that the program/project will result in continued intended benefits for the target group after project completion.	The evaluation finds that there is a low probability that the program/project will result in continued benefits for the target group after completion.	The evaluation finds it is likely that the program or project will result in continued benefits for the target group after completion.	Evaluation finds that it is highly likely that the program or project will result in continued benefits for the target group after completion.
3.2 Extent MO-supported projects and programs are reported as sustainable in terms of institutional and/or community capacity	The design of MO-supported programs and projects failed to address the need to strengthen institutional and/or community capacity as required.	MO programs and projects may have failed to contribute to strengthening institutional and/or community capacity.	MO programs and projects may have contributed to strengthening institutional and/or community capacity, but with limited success	Either MO programs or projects have contributed to significantly strengthen institutional and/or community capacity as required, or institutional partners and communities already had the required capacity to sustain program outcomes.

4. Efficiency

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
4.1 Program activities are evaluated as cost-efficient:	The evaluation reports that there is credible information indicating that MO-supported programs and projects are not cost-efficient.	The evaluation indicates that the MO-supported programs and projects under evaluation do not have credible, reliable information on the costs of activities and inputs and, therefore, the evaluation is not able to report on cost-efficiency.	The evaluation reports that the level of program outputs achieved when compared to the cost of program activities and inputs is appropriate even when the program design process did not directly consider alternative program delivery methods and their associated costs.	The evaluation reports that MO supported programs and projects are designed to include activities and inputs that produce outputs in the most cost-efficient manner available at the time.
4.2 Evaluation indicates implementation and objectives achieved on time	The evaluation reports that less than half of stated output and outcome level objectives of MO-supported programs and projects are achieved on	The evaluation reports that less than half of stated output and outcome level objectives of MO supported programs and projects are achieved on	The evaluation reports that more than half of stated output and outcome level objectives of MO supported programs and projects are achieved on	The evaluation reports that nearly all stated output and outcome level objectives of MO supported programs and projects are achieved on time.

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
	time, and there is no credible plan found by the evaluation that would suggest significant improvement in on-time objectives achievement in the future.	time, but the program or project design has been adjusted to take account of difficulties encountered and can be expected to improve the pace of objectives achievement in the future.	time and that this level is appropriate to the context faced by the program during implementation.	

5. Cross Cutting Themes: Gender Equality and Environmental Sustainability

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
5.1 Extent MO-supported activities effectively address the cross-cutting issue of gender equality.	The evaluation finds MO-supported activities are unlikely to contribute to gender equity or may in fact lead to increases in gender inequities.	The evaluation report finds that MO-supported activities either lack gender equality objectives or achieve less than half of their stated gender equality objectives at the outcome level. (Note: where a program or activity is clearly gender-focused (maternal health programming, for example) achievement of more than half its stated objectives warrants a satisfactory rating.)	MO-supported programs and projects achieve a majority (more than 50%) of their stated gender equality objectives at the outcome level.	MO-supported programs and projects achieve all or nearly all of their stated gender equality objectives at the outcome level.
5.2 Extent changes are environmentally sustainable.	MO-supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. In addition, the evaluation reports that changes resulting from MO-supported programs and projects are not environmentally sustainable.	MO-supported programs and projects do not include planned activities or project design criteria intended to promote environmental sustainability. There is, however, no direct indication that project or program results are not environmentally sustainable.	MO-supported programs and projects include some planned activities and project design criteria to ensure environmental sustainability. These activities are implemented successfully, and the evaluation reports that the results are environmentally sustainable	MO-supported programs and projects are specifically designed to be environmentally sustainable and include substantial planned activities and project design criteria to ensure environmental sustainability. These plans are implemented successfully, and the evaluation reports that the

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
				results are environmentally sustainable.

6. Using Evaluation and Monitoring to Improve Development Effectiveness

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
6.1 Systems and process for evaluation effective	Evaluation report specifically notes that evaluation practices in use for programs and projects of this type are seriously deficient.	No indication in the evaluation report that the program is subject to systematic and regular evaluations.	Evaluation report notes that program being evaluated is subject to systematic and regular evaluations or describes significant elements of such practice. No mention of policy and practice regarding similar programs and projects.	Evaluation report notes that program being evaluated (along with similar programs and projects) is subject to systematic regular evaluations or describes significant elements of such practice.
6.2 Systems and processes for monitoring and reporting on program results are effective	The evaluation notes an absence of monitoring and reporting systems for the program.	While monitoring and reporting systems for the program exist, either they do not report on a regular basis or they are inadequate in frequency, coverage or reliability.	Evaluation notes that monitoring and reporting systems for the program are well established and report regularly.	Evaluation notes that monitoring and reporting systems for the program are well established and report regularly. The quality of regular reports is rated highly by the evaluation, and results are reportedly used in the management of the program.
6.3 Results-Based Management (RBM) systems effective	Report notes that there is no evidence of the existence of an RBM system for the program, and no system is being developed.	While an RBM system is in place, or being developed, it is unreliable and does not produce regular reports on program performance.	Evaluation notes that RBM system is in place and produces regular reports on program performance.	Evaluation notes that RBM system is in place for the program, and there is evidence noted in the evaluation that the system is used to make changes in the program to improve effectiveness.

Sub-criteria	(1) Highly Unsatisfactory	(2) Unsatisfactory	(3) Satisfactory	(4) Highly Satisfactory
6.4 MO makes use of evaluation to improve development effectiveness	Evaluation report does not include a management response and does not have one appended to it or associated with it. There is no indication of how the evaluation results will be used. There is no indication that similar evaluations have been used to improve effectiveness in the past.	Evaluation report includes a management response (or has one attached or associated with it), but it does not indicate which recommendations have been accepted, or there is some non-specific indication that similar evaluations have been used to improve program effectiveness in the past.	Evaluation report includes a management response (or has one attached or associated with it) that indicates which recommendations have been accepted. On the other hand, there is a clear indication that similar evaluations in the past have been used to make clearly identified improvements in program effectiveness.	Evaluation report includes a management response (or has one attached or associated with it) that describes a response to each major recommendation which is appropriate and likely to result in the organizational and programmatic changes needed to achieve their intent.

Annex 6: Corporate Documents Reviewed

CIDA Corporate Documents

CIDA Strategy for Engagement with the World Health Organization. 2011.

Review of the Effectiveness of CIDA's Grants and Contributions, 2005/06 to 2010/11. 2011.

WHO Corporate Documents

Decisions. Executive Board Special Session on WHO Reform. WHO. 2011.

Decisions and list of resolutions, Executive Board 131st session (EB131/DIV/2). WHO. 2012.

Engagement for Health: Eleventh General Programme of Work 2006-2015: A Global Health Agenda. 2006.

Gender mainstreaming in WHO: where are we now? Report of the baseline assessment of the WHO Gender Strategy. 2011.

Medium-Term Strategic Plan 2008-2013 Interim Assessment. 2011.

Report of the Programme, Budget and Administration Committee of the Executive Board. 2011.

Review of Management, Administration and Decentralization in the World Health Organization. JIU/REP/2012/6. United Nations Joint Inspection Unit, Geneva, 2012.

WHO Managerial Reforms. 2011.

WHO Reform, Draft Formal Evaluation Policy. 2012.

Working for Health. An Introduction to the World Health Organization. 2007.

Annex 7: CIDA Funding to Multilateral Humanitarian and Development Organizations

Long-term Institutional Funding[42]

Long-term institutional funding can be defined as un-earmarked funding to an MO in support of that organization's mandate. According to the OECD, there are currently 170 MOs active in development and eligible to receive aid funding. As of 2010–2011, CIDA provided long-term institutional funding to 30 of these MOs. CIDA's funding was highly concentrated with 9 MOs receiving 80% of its total long-term institutional funding from 2007–2008 to 2010–2011.

Funding to Specific Multilateral and Global Initiatives

Specific multilateral and global funding can be defined as funding to MOs in support of a key program or activity usually in a specific thematic area and often global in scope. Within this category, there are two sub-types: 1) humanitarian assistance, and 2) other global initiatives programming.

Humanitarian assistance is provided based on need and usually in response to specific appeals issued by MOs with expertise in providing humanitarian assistance. The main MOs involved in providing humanitarian assistance are the World Food Programme (WFP), the United Nations High Commissioner for Refugees (UNHCR), the International Committee of the Red Cross (ICRC) and the UN Office for Coordination of Humanitarian Affairs. The United Nations Children's Fund (UNICEF), although not primarily a humanitarian organization, also delivers humanitarian assistance with a specific emphasis on the needs of children.

The second sub-type of specific multilateral and global funding involves global initiatives in other sectors. These initiatives are in sectors that deal with issues that transcend borders and thus lend themselves to a multilateral approach. The main sectors CIDA supports with this type of funding are health, environment and economic growth. The health sector is the most important of these, especially in light of the challenges of infectious diseases like AIDS and tuberculosis that do not respect international borders. Bilateral programming in a single country is unlikely to succeed in meeting the challenges of infectious diseases in the absence of regional and global programs.

[42] All the information in this section has been extracted from *A Review of Evidence of the Effectiveness of CIDA Grants and Contributions.* (pp. 45–46). CIDA. 2011

Funding to Multilateral Initiatives Delivered by other CIDA Branches

Multilateral initiatives can also receive funding from other CIDA branches, mostly through multi-bi funding from Geographic programs. Multi-bi funding refers to earmarked funding to a specific MO initiative by a CIDA geographic program to support a specific activity in a specific country or group of countries. It is considered “bilateral” assistance because it is funded through CIDA’s geographic programs in the context of the program’s country strategies or programming frameworks.

Multi-bi funding accounts for a large and growing share of CIDA resources. It more than tripled in the five years from 2002–2003 to 2007–2008, mainly because of substantial funding to programs in fragile states. By 2007–2008, CIDA multi-bi funding had reached $691 million, with 53% spent in fragile states, including 37% of all multi-bi funding spent in Afghanistan.

In fragile states, where United Nations (UN) organizations and the World Bank are often assigned specific roles by member governments, use of multi-bi funding by CIDA can sometimes help the Agency to limit fiduciary risk and result in a reduced administrative burden on the very weak national institutions. The use of this type of funding is also consistent with Canada’s commitment to the Paris Declaration principles of aid effectiveness that includes a call for donors to harmonize their aid and use program-based approaches where they can be effective.

It is important to note that CIDA’s geographic programs manage multi-bi funding according to the same basic processes that govern all of the Agency’s geographic programming. For example, CIDA’s geographic programs are responsible for monitoring and reporting on the effectiveness of funds used in this way. Country Program Evaluations that examine CIDA’s bilateral programs in a given country include in their remit programming delivered by MOs and supported by multi-bi funding.

Annex 8: Management Response

Overall management response to the evaluation report:

The Review of the World Health Organization's (WHO) Development Effectiveness from 2007–2010 by CIDA's Evaluation Division has provided a valuable opportunity to comprehensively review the WHO's capacity to effectively deliver on programming areas of interest to CIDA. The Report is particularly timely considering the current process of reforming the WHO, at administrative, managerial and technical levels, that was set in motion in 2011, aimed at improving the efficiency, transparency and accountability of the Organization. In this regard, the Review will be a useful tool for informing CIDA's efforts to support the organization as a key institutional partner by fostering a stronger, more effective organization. This is in line with CIDA's strategic objectives for engagement with the WHO as set out in our Institutional Strategy.

The major finding of the Review is that there is not enough information to make generalized conclusions about the WHO's programming as a whole due to limited set of available evaluation reports. Nevertheless, the Review acknowledges that the available evaluation reports do provide insights into the effectiveness of WHO programs that were evaluated.

The Review highlights key non-generalized institutional strengths and challenges that are highly relevant to CIDA's programming with the Organization in the years ahead. The Review concludes that the WHO appears to have achieved positive results in terms of the relevance of its programs to stakeholder needs and national priorities, its ability to achieve most of its development objectives and expected results, and the sustainability of WHO programs. Areas where the WHO requires strengthening largely centre on its evaluation function, including insufficient program coverage provided by the organization's evaluations, limited reporting on cost-efficiency, limited integration of crosscutting issues such as gender equality and environmental sustainability in WHO evaluations, and limited accessibility of evaluation reports to Member States and external stakeholders.

CIDA accepts all of the recommendations of the Review aimed at improving evaluation and results-based management at the WHO. These recommendations are in line with the objectives of CIDA's existing engagement with the WHO. As the report noted (p.35), as one of several stakeholders working with the WHO, Canada is limited to the extent which it can unilaterally influence improvements on the Organization. However, CIDA benefits from direct access to senior management at the WHO and has already engaged with them about these issues. As noted in the report (p. 32), CIDA has been vocal in support of WHO reforms and will play an important role in following up on its effectiveness.

CIDA's interventions to address the recommendations will focus on the following: 1) CIDA will engage with WHO staff and senior management to address the issues and recommendations outlined in the Review, including monitoring the WHO's commitment to implement its new evaluation policy, exploring ways to support WHO plans to mainstream gender across the organization, and monitoring progress in implementing the Global Management System; 2) CIDA will liaise with like-minded donor agencies to build support around these recommendations which can be highlighted during the meetings of the WHO governing bodies.

Recommendations	Commitments / Actions	Responsibility Centre	Target Completion Date	Status
1. Canada should monitor efforts at reforming the evaluation function at the WHO as the new policy on evaluation is implemented. In particular, CIDA should use its influence at the Executive Board and with other donor agencies to advocate for a sufficiently resourced and capable evaluation function that can provide good coverage of WHO programming over time.	Agreed. 1.1 Through Canada's membership on the Executive Board, CIDA has been actively involved in the evolution of the WHO's new evaluation policy, which was recently adopted by the World Health Assembly in May 2012. As part of the adoption of the new evaluation policy, the Executive Board will provide input into the biennial Organization-wide evaluation workplan and revise the evaluation policy as needed based on annual reports from the Office of Internal Oversight Services. 1.2 CIDA will actively engage with WHO staff and senior management to advocate for the strengthening of the WHO's evaluation function, as appropriate. 1.3 CIDA will liaise with like-minded donor agencies to identify areas of mutual interest for the strengthening of the WHO's evaluation function.	CIDA/ Multilateral and Global Programs Branch (Global Initiatives Directorate)	1.1 2012 1.2 2013 1.3 2013	1.1 Completed – Please reference EDRMS#:6068730

Recommendations	Commitments / Actions	Responsibility Centre	Target Completion Date	Status
2. CIDA should monitor the implementation of the evaluation policy so that future WHO evaluations sufficiently address gender equality.	Agreed. 2.1 CIDA is working with the WHO/Gender Equity and Rights (GER) Unit to encourage them to mainstream crosscutting priorities, including gender, equity and rights, across all levels of the WHO. CIDA will communicate the results of this review relating to the issue of gender equality in evaluations to the WHO GER unit. CIDA will continue to engage in dialogue with the WHO's GER Unit and like-minded donors to explore ways to support WHO plans to mainstream gender across the organization, including the evaluation function. 2.2 CIDA will convene a "Friends of WHO" meeting with colleagues from across the Agency to outline the results of this review and to advocate for the inclusion of gender equality in future WHO evaluations for CIDA-funded initiatives.	CIDA/ Multilateral and Global Programs Branch (Global Initiatives Directorate)	2.1 2013 2.2 2013	

Recommendations	Commitments / Actions	Responsibility Centre	Target Completion Date	Status
3. CIDA should encourage the WHO to implement a system for publishing regular (possibly annual) reports on development effectiveness that builds on the work of the reformed evaluation function. In general terms, there is a need to strengthen the WHO's commitment to reporting on the effectiveness of programs.	Agreed. 3.1 Through our membership on the Executive Board and the Programme, Budget, Administration Committee (2010–2012), Canada has been advocating to have access to evaluation and audit reports publically available. A commitment was made by the WHO Secretariat at the May 2012 Executive Board meeting to improve access to audit and evaluation reports through a limited access website later this year, but it is not yet in place. CIDA feels strongly that access to this information is critical to ensuring due diligence. 3.2 CIDA will continue to advocate for the WHO to increase accessibility to evaluation and audit reports and deliver on the promised website in a timely manner.	Director General, Global Initiatives Directorate	3.1 Completed 3.2 2013	3.1 Please reference EDRMS#: 6054296

Recommendations	Commitments / Actions	Responsibility Centre	Target Completion Date	Status
4. CIDA should encourage WHO to systematically manage for results. The ongoing upgrading and further implementation of the Global Management System at the WHO may offer such an opportunity.	Agreed. The WHO has made progress in implementing new management and administrative mechanisms to improve its effectiveness, including the Global Management System. However, despite the availability of new tools, aspects of the organizational culture limit their effectiveness, including financial planning and allocation of resources, human resources management, and results-based management. 4.1 CIDA will liaise with the WHO on progress in implementing the Global Management System and advocate for a more effective results-based management system. 4.2 CIDA will actively engage at the Executive Board and other forums, to monitor the implementation of Global Management System and encourage its use as the basis for managing results.	Director-General, Global Initiatives Directorate	4.1 2013 4.2 2013	