fraserhealth Better health. Best in health care.

A Guide to Planning and Conducting Program Evaluation

May 2009

Research Administration and Development

...fostering a culture of curiosity within a climate of spirited scientific inquiry

Supported by the Michael Smith Foundation for Health Research

Michael Smith Foundation for Health Research

Research Administration and Development

Acknowledgement

This Guide was produced as one of the projects funded by the Michael Smith Foundation for Health Services Research 'Health Authority Capacity Building Grant' which was awarded to FH and overseen by its Research Administration and Development office. Fraser Health acknowledges the valuable contribution that FH personnel and other stakeholders have made to the completion of this document, by way of participation on the Evaluation Guide Advisory Group which included representatives of internal and external stakeholder groups including the Fraser Health Authority, Interior Health Authority, Vancouver Coastal Health Authority, Northern Health Authority, Provincial Health Services Authority, Michael Smith Foundation for Health Research and the University of the Fraser Valley. Heather McNeill, Research Administration and Development, was the primary author of this document.

Fraser Health Authority
#300-10334 152A Street
Surrey BC V3R 7P8

Website: http://www.fraserhealth.ca/Professionals/Research/Pages/default.aspx

TABLE OF CONTENTS

1.0 Introduction

1.1 Project Background

In spring 2008, the Research Administration and Development (RAD) Office of the Fraser Health Authority received additional Health Authority Capacity Building funding from the Michael Smith Foundation for Health Research (MSFHR), through the Health Services & Policy Research Support Network. The objective of this funding was to enable health authorities in British Columbia to increase their ability to use research and engage effectively in local, regional and provincial research, evaluation and planning activities.

While RAD has supported evaluation research through the provision of consultation services by the RAD epidemiologist, there is a growing need and interest to improve the capacity of Fraser Health (FH) personnel to plan and conduct program evaluations. To meet this need, RAD used the Health Authority Capacity Building extension funding to contract an Evaluation Researcher to create this document.

The content of this document is based on a review of needs of FH personnel regarding program evaluation and a synopsis of program evaluation literature, both conducted by RAD. Development of the Guide was informed by the Evaluation Guide Advisory Group that included representatives from FH, other Health Authorities, MSFHR and the University of the Fraser Valley.

1.2 Purpose and Audience

The application of this document supports best practice standards for planning and conducting program evaluation in Fraser Health.

The document will be used by RAD as a means to provide a consistent and professional approach when consulting FH personnel on program evaluation. Evaluation consultation may be requested by individuals responsible for evaluation (or their delegates), including:

- FH employees and managers;
- Privileged physicians.

These consultations may relate to evaluation planning, methodology, data collection and analysis. Ideally, RAD hopes to provide consultation on evaluating planning in tandem with program design, before program implementation.

The document could also serve as a reference (independent of consultation with RAD) for individuals or teams responsible for planning and conducting program evaluation.

1.3 Overview and Benefits

The document has been arranged to provide the user with information to produce an evaluation plan with sufficient detail to subsequently conduct an evaluation that is useful and credible.

Sections 3.0, 4.0 and 5.0 each conclude with a checklist of points to guide evaluation in accordance with standards of effective evaluation practice (utility, feasibility, propriety, accuracy).

Section 2.0 introduction to program evaluation, offers information to:

- Distinguish between evaluation and performance measurement;
- Recognize the importance of evaluation for programs in health care;
- Be aware of the standards that govern the evaluation profession;
- Consider organizational capacity to evaluate goals.

Section 3.0 describes steps to prepare for an evaluation, including:

- Form the Evaluation Project Team;
- Identify the need for an external Evaluator;
- Realize the benefits of forming an Evaluation Steering Committee;
- Undertake administrative tasks (E.g. Evaluation Charter).

Section 4.0 describes steps to develop an Evaluation Plan, which includes:

- Clarify the purpose of the program and expected outcomes;
- Develop a Logic Model to better communicate with stakeholders about the program;
- Identify intended user(s) and intended use(s) of evaluation findings;
- Identify the evaluation questions and a strategy (evaluation methodology) to answer these questions;
- Develop a plan to communicate evaluation findings.

Section 5.0, the stage of conducting evaluation and reporting results, offers guidance to:

- Implement the Evaluation Plan;
- Develop useful recommendations;
- Select a report format that is suited for intended users;
- Ensure the evaluation has been conducted according to program evaluation standards upheld by the Canadian Evaluation Society.

1.4 Summary of Steps to Planning and Conducting Program Evaluation

This document, "A Guide to Planning and Conducting Program Evaluation", describes evaluation of a project or program as a three step process.

Step 1: Preparing for Evaluation

1.1 Getting People Together
- Form an Evaluation Project Team
- Identify Evaluator
- Identify and Engage Key Stakeholders
- Consider forming a Steering Committee

1.2 Administrative Considerations
- Develop an Evaluation Charter
- Review Research Ethics Policy
- Develop a Contract and Budget
- Tasks and Timelines

Step 2: Develop an Evaluation Plan

2.1 Write Program Profile
- Background and Context
- Objectives of the Program
- Target Population

2.2 Develop Program Theory and Logic Model
- Program Theory and Program Components
- Logic Model

2.3 Identify Intended User(s) and Intended Use(s)

2.4 Evaluation Questions
- Develop and Select Evaluation Questions

2.5 Evaluation Methodology
- Data Collection
- Evaluation (Research) Design

2.6 Communications Plan (to Disseminate Evaluation Findings)

Step 3: Conduct Evaluation, Report and Use Findings

- Review and Implement the Evaluation Plan (Conduct Evaluation)
- Develop Recommendations (optional)
- Report Findings
- Use Findings

1.5 Rationale for the Three Step Evaluation Process

This document has conceived evaluation as a three stage process because this division is realistic for the Fraser Health Authority. The goal of the document is to accommodate for scenarios that are feasible for programs within Fraser Health.

Below are five scenarios for how planning and conducting evaluation may take place:

- **Prepare – Plan – Evaluate**
  - The evaluation is planned and conducted in continuum. The evaluation is planned with the clear intent to conduct the evaluation upon completion and acceptance of the plan.

- **Prepare – Plan….then Evaluate**
  - Time elapses between completion of the evaluation plan and conducting the evaluation.

- **Prepare – Plan…but no Evaluation**
  - The program prepares the evaluation plan, but the evaluation does not take place.

- **Prepare – Plan**
  - An evaluation plan was completed (e.g. to meet funding requirements), but there was no firm intent to evaluate.

- **Plan**
  - The preparatory steps may identify administrative or resource issues that hold-up the evaluation project. It is not feasible to neither create an evaluation plan nor conduct an evaluation.

2.0 Background on Program Evaluation

2.1 What is Program Evaluation?

Program evaluation is *"...a rich and varied combination of theory and practice..."* that creates "*information for planning, designing, implementing, and assessing results of our efforts to address and solve problems...."* [1]

A defining characteristic of program evaluation is that the results are used for making decisions. Program evaluation has been defined as:

> *"The systematic collection, analysis and reporting of information about a program to assist in decision-making."* [2]

Program evaluation is often used to determine program effectiveness, where effectiveness describes the extent to which a program or initiative is meeting its expected outcomes, and to inform action to change.

Program evaluation has been defined as:

- *"... The use of social research methods to systematically investigate the effectiveness of social intervention programs in ways that are adapted to their political and organizational environments and are designed to inform social action to improve social conditions."* [3]

To determine effectiveness, program evaluation often seeks to minimize confounding factors that attribute observed outcomes of the intervention or program.[4]

2.2 What is a Program?

The term **program** generally refers to *"any group of related, complementary activities intended to achieve specific outcomes or results."* [5] The scale of activity can vary among programs.

In Fraser Health, a program may include deliver of a specific service (e.g. Medical Imaging Services) or a clinical intervention with multiple services (e.g. Hospice Palliative Care Program). The program may be delivered at a particular location (e.g. Burnaby Hip and Knee Centre) or throughout the region (e.g. Nursing Support Services).

In addition to programs, the concepts in this document can be applied to projects (e.g. CapPlan Demonstration Project) or initiatives (e.g. iCare).

---

[1] McDavid, J. C., Hawthorn, L. R. L. (2006).
[2] Van Marris & King. (2006)
[3] Rossi, P. H., Lipsey, M. W, & Freeman, H. E. (2004).
[4] Mayne, J. (2001).
[5] Van Marris & King. (2006)

2.3 When is Program Evaluation Useful?

Use evaluation to inform decisions to improve a program. Evaluation may be used to generate information to improve a program; this type is described as formative evaluation. Evaluation also supports "evidence-informed" decisions about how to improve program management.

In addition, using evaluation for program improvement may promote a spirit of inquiry among program personnel, those most intimate with program function and the effects of policy. The opportunity to contribute to further success of a program may improve employee morale and their commitment to an organization.

Use evaluation to demonstrate accountability to stakeholders or funders. Evaluation may also be conducted to show accountability and to provide information about program effectiveness to decision makers; this type is described as summative evaluation.[6] Program evaluation serves an important role to interpret trends that are revealed through statistical analysis and to examine program implementation.

Accountability for public spending is an important theme among governmental organizations and may be achieved through performance measurement and evaluation.[7] For example, within the BC Provincial Government and the Fraser Health Authority, performance measurement is used routinely for accountability reporting.[8]

2.4 How does Performance Measurement Compare to Evaluation?

Performance measurement describes the process and systems of selection, and development of performance indicators. A performance indicator is composed of a number (how much) and a unit (of what). The indicator measures an important component of the program (an input, activity, output or outcome).

Performance measurement is increasingly relied upon for information about the success of public spending. In Canada, and most industrialized countries, government agencies are re-organizing the delivery of social services to demonstrate and measure the changes that occur as a result of the investment and intervention.[9]

In health care, performance measurement serves many purposes, including:

- To report on quality of care;
- To make comparisons (benchmarking) over time between places (e.g. hospitals);
- To make judgments and set priorities (e.g. choosing a hospital or surgery);
- To support accountability, regulation, and accreditation;

---

[6] McDavid, J. C., Hawthorn, L. R. L. (2006).

[7] Policy on Transfer Payments: http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=13525

[8] Currently, the accountability mechanism between the Health Authorities and the Ministry of Health is the Annual Report, a document that responds to requirements outlined in the Government Letter of Expectations (elements of which are reiterated in the Health Authorities' Service Plans). Since 2006, Fraser Health has reported to the FH Executive Board by way of a Balanced Scorecard, which includes performance indicators presented in the Annual Report.

[9] Coyne, K. & P. Cox. (2004).

- To support quality improvement; and to support patient choice of providers.[10]

Performance measurement can inform how a program is performing relative to respective targets and can detect the presence (or absence) of change in an intended outcome. In some instances, analysis of performance indicators reveals sufficient information to establish a causal relationship between a specific outcome and the activities and outputs of a program.

Developing a plan to monitor and track performance indicators enables managers to use trend analysis to inform decision-making. Performance measurement systems can be designed to provide information on an on-going basis.

However, performance measurement cannot explain changes in nor attribute changes to a particular program. Nor can performance measurement, alone, address the strengths and weaknesses of program design.

Evaluation, on the other hand, seeks multiple sources of information as a means to improve program implementation and attribute observed outcomes to the program. Evaluation requires additional data sources and methodologies to answer more extensive questions. As a result of the amount of resources and time required to conduct program evaluation, it is conducted periodically.

❖ **For More Information**
Appendix A describes the benefits of a Performance Measurement Plan and defines **performance indicator**, **target**, **benchmark**, and **baseline**.

2.5 What is Performance Management?

In fact, information derived from on-going performance measurement activities and periodic evaluations can be coordinated as a **cycle of performance management** to provide more insightful information to decision makers. [11]

❖ **For More Information**
Together, "*performance measurement and evaluation present valuable opportunities to learn and adjust so that the desired results may be achieved*."

See the Treasury Board of Canada Secretariat (2007) "Performance Reporting Good Practices Handbook". http://www.tbs-sct.gc.ca/rma/rbm-gar_e.asp

---

[10] Mainz, J. (2003)
[11] McDavid, J. C., & Hawthorn, L. R. L. (2006).

2.6 Why is Program Evaluation Important for Healthcare?

Decision-makers tend to focus on trend analysis of indicators to inform new policies and programs. Historically, and in present times, it is less often the case that decision-makers seek an explanation about why observed results occur.[12]

In Canada, the healthcare system has experienced "dramatic and largely invisible changes" over the last decade, including changes in service delivery models, organizational structure, costs and funding models, and technology.[13]

In this complex, quickly changing environment, decision-makers need to understand why observed results occur. An evaluation process can seek to understand the attribution or contribution of observed results to a program. Evaluation can also inform new policies and programs that respond to these challenges.

2.7 Are There Evaluation Standards?

In March 2008, the Canadian Evaluation Society National Council adopted standards for effective Canadian evaluation practice. [14]

There are 30 Evaluation Standards and these can be described by four categories:

1) **Utility standards** are intended to ensure that an evaluation will serve the information needs of intended users.

2) **Feasibility standards** are intended to ensure that an evaluation will be realistic, prudent, diplomatic, and frugal.

3) **Propriety standards** are intended to ensure that an evaluation will be conducted legally, ethically, and with due regard for the welfare of those involved in the evaluation, as well as those affected by its results.

4) **Accuracy standards** are intended to ensure that an evaluation will reveal and convey technically adequate information about the features that determine worth or merit of the program being evaluated.

Sections 3.0, 4.0 and 5.0 conclude with relevant checkpoints related to these standards.

❖ **For More Information**
The Evaluation Standards, definitions and criteria are available at: http://evaluationcanada.ca/site.cgi?s=6&ss=10&_lang=en.

---

[12] Rutman, L. (1977).
[13] Centre for Health Services and Policy Research. (2003).
[14] Developed by the Joint Committee, a coalition of major professional associations in the US.

2.8 When Not to Evaluate?

The Evaluation Standards also provide guidance regarding when to proceed with an evaluation. If it is anticipated that the standards will not be met, then the evaluation should not be conducted less it risk not meeting professional standards.

In addition, there are four circumstances when it is not advisable to evaluate:[15]

- There are no questions about the program;
- The program has no clear direction;
- Stakeholders cannot agree on the program objectives;
- When there is not enough money to conduct a sound evaluation.

2.9 Why and When to Develop an Evaluation Plan?

Development of an evaluation plan establishes:

- Program description and background, including key stakeholders;
- A logical program design, defining how activities lead to expected results;
- Evaluation questions and appropriate methodology to answer these questions.

Ideally, evaluation planning occurs in tandem with program design, before the program is implemented. The advantages of planning an evaluation at this stage include:

- Opportunity to clarify the purpose of the program;
- Time to garner stakeholder involvement and support for the evaluation;
- Time to gather baseline data;
- Time to establish data systems to collect meaningful performance indicators.

In reality, evaluation planning often takes place after program implementation. In this case, evaluation planning may face a number of challenges:

- Baseline data has not been collected;
- Need to seek consent from program recipients to collect data;
- If data is collected retrospectively, it may not be accurate.

Still, there are advantages to evaluation planning for an established program:

- Individuals have more program knowledge and experience;
- Model of program delivery has evolved;
- Program staff and organizational structure is more stable.

[15] Issel, L.M. (2004).

2.10 Is Your Organization Ready to Evaluate?

Organizational capacity for evaluation describes the ability to plan and conduct an evaluation. **Organizational readiness** is a related concept that describes organizational understanding, capacity and willingness to monitor and evaluate its goals.

❖ **Additional References**

"A Checklist for Building Organizational Evaluation" is available through The Evaluation Center at the Western Michigan University at: http://www.wmich.edu/evalctr/checklists/ecb.pdf

Read more about readiness assessment in "Ten Steps To a Results-based Monitoring and Evaluation System" sponsored by the World Bank at: http://www.preval.org/documentos/00804.pdf

---

3.0 Preparing for Evaluation

3.1 Getting People Together

3.1.1 Form an Evaluation Project Team

Typically, there is an individual accountable to ensure an evaluation is planned and conducted, such as a Program Manager or Director. This individual will strike an Evaluation Project Team to complete the task.

The Evaluation Project Team is comprised of individuals who can offer expert program knowledge and other skills as necessary to complete the evaluation process.

In addition to the Program Manager or Director, the Evaluation Project Team will likely include individuals who can fulfill the roles of:

- Evaluator (see below);
- Program Expert and Data Systems Expert (if necessary);
- Stakeholders, including Intended Users;
- Evaluation Coordinator (to prepare and circulate documents).

3.1.2 Who is an Evaluator?

An Evaluator is an individual with the necessary expertise to plan and conduct the evaluation. Essential competencies for program evaluators have been categorized as:[16]

- **Professional Practice:** knowing and observing professional norms and values, including evaluation standards.[17]

- **Systematic Inquiry:** expertise in the technical aspects of evaluations, such as design, measurement, data analysis, interpretation, and sharing results.

- **Situational Analysis:** understanding and attending to the contextual and political issues of an evaluation, including determining evaluability, addressing conflicts, and attending to issues of evaluation use.

- **Project Management:** the nuts and bolts of managing an evaluation from beginning to end, including negotiating contracts, budgeting, identifying and coordinating needed resources, and conducting the evaluation in a timely manner.

- **Reflective Practice:** an awareness of one's program evaluation expertise as well as the needs for professional growth.

---

[16] Reproduced from M.Q. Patton (2008) (p.200)

[17] See the Program Evaluation Standards, upheld by the Canadian Evaluation Society, at: http://evaluationcanada.ca/site.cgi?s=6&ss=10&_lang=en

- **Interpersonal Competence:** the people skills needed to work with diverse groups of stakeholders to conduct program evaluations, including written and oral communication, negotiation, and cross-cultural skills.

3.1.3 Is the Evaluator Internal or External to the Organization?

The Evaluator may be an individual who is internal (e.g. an employee, FH personnel) or external (e.g. a consultant) to the organization.

Internal evaluation is the "*process of using staff members who have the responsibility for evaluating programs or problems of direct relevance to managers*." About three quarters of evaluation studies in North America use internal evaluation. [18]

If the Evaluator is an employee, she/he can offer:

- Organizational knowledge to ensure evaluation methodology is relevant;
- Potentially, a responsibility to use the evaluation to achieve on-going organizational improvement.

External evaluation refers to contracting with an external consultant to complete the evaluation. If the Evaluator is an external consultant, she/he can offer:

- Necessary (and often specialized) expertise;
- Objectivity.

In fact, there are many possible combinations of internal and external evaluation. Depending on the needs of the evaluation, an external or internal Evaluator could assume different roles, including those to:

- Direct the evaluation planning and conducting process;
- Guide evaluation by involving key decision-makers, acting as a facilitator and resource;
- Act as an "empowerment" facilitator and enable a team that has total authority and resources to evaluate;[19]
- Act as a counsel, with ownership shared by team and evaluator.

To decide whether internal or external evaluation is most appropriate, consider:

- Do personnel/staff have necessary methodological and technical expertise to plan and/or conduct evaluation?

- Do the intended users require the evaluation to be conducted with objectivity and independence from the organization?

- Is the evaluation intended for internal management purposes or to show accountability to an external funder?

- Are there sufficient resources to contract for evaluation expertise?

---

[18] Love, A.J. (1991)

[19] Fetterman, D. M., Kaftarian, S. J., and Wandersman, A. (1996).

3.1.4 Identify and Engage Stakeholders, including Intended Users

Evaluation stakeholders are individuals and groups (both internal and external) who have an interest in the evaluation, that is, they are involved in or affected by the evaluation.[20]

Evaluation stakeholders include:

- Program management (managers, team leaders, executive sponsors);
- Funding agencies;
- Program personnel (first line leaders, support staff);
- Volunteers and community representatives;
- Program participants.

Evaluation stakeholders include anyone who makes decisions or desires information about a program.[21]

Identify stakeholders

Ask persons in leadership roles to identify stakeholders. Do not exclude any potential stakeholders because of gender, ethnicity, or language background.

To identify stakeholders, consider:

- Who is funding the program?
- Who delivers the program (e.g., third party delivery agencies)?
- Who has requested the evaluation (e.g. funding agency, decision makers)?
- Who will use the results of the evaluation and how?
- How will the organization, stakeholders and personnel respond to findings?
- Who will the evaluation results be disseminated to?

Learn about stakeholders

The values and interests of key stakeholders affect what is evaluated, how information is collected and interpreted, and how the findings are used.

To learn about stakeholders, consider:

- What are the interests of each stakeholder?
- Whose interests/views will be given priority?
- What factors and pressures are encouraging the evaluation?
- Is the purpose of the evaluation to prove results and accountability to the program funder?
- Is the purpose of the evaluation to improve program management and provision of service delivery?

Tips for working with stakeholders

- Learn how key stakeholders view the evaluation's importance, how they would like to use its results, and what particular information would be useful.

[20] Van Marris & King. (2006).
[21] Patton, M.Q. (2008).

- Reach an understanding about the relative importance of the potential stakeholders and their information requirements.

- Involve key stakeholders directly in designing and conducting the evaluation.

- Where necessary, help them to develop realistic expectations that account for the methodological, financial, and political constraints on the evaluation.

- Maintain flexibility in including additional key stakeholders throughout the process.

Focus on Primary Intended Users

Typically, the information needs (potential evaluation questions) will exceed the resources (e.g. time, money, expertise) to complete the evaluation.

One approach to prioritize evaluation questions is to focus on the needs of **primary intended users** of the evaluation, a subset of evaluation stakeholders. Primary intended users are expected to use the evaluation findings to improve the program or to make judgements about the program.

The practice of designing the evaluation according to the information needs of the primary intended users of the evaluation is known as the practice of **utilization-focused evaluation**.[22] If the primary intended users work with the evaluator then it is more likely that the evaluation will be useful, meaningful, relevant and credible.

❖ **For Consideration**

Consider forming an Evaluation Steering Committee as a means to:

- Provide strategic and critical feedback on the work of the Project Team;
- Create a formal mechanism for key stakeholders to provide feedback to the Evaluation Project Team.

Considerations for forming a Steering Committee include:

- Budget for meetings;
- Time constraints and availability of prospective members;
- Responsibilities of the Evaluation Steering Committee.

Develop a Terms of Reference for the Steering Committee that describes the scope, time commitment and specific responsibilities (E.g. Feedback on the content of documents; Final approval of the evaluation plan and report).

Membership and terms of the Evaluation Steering Committee may change if significant time elapses between completion of the planning stage and the start of the evaluation.

---

[22] Patton, M.Q. (2008).

3.2 Administrative Considerations

3.2.1 Develop an Evaluation Charter

An Evaluation Charter is a document (like a Project Charter) that is developed to seek formal approval from internal management to proceed with an evaluation project. The evaluation project may be to develop an evaluation plan and/or conduct an evaluation.

The Evaluation Charter describes:

- Goals of the evaluation project;
- Objectives of the evaluation project (concrete steps to completing the project);
- Evaluation stakeholders and primary intended users;
- Assumptions about how the evaluation project will proceed;
- Known risks to completing the evaluation project;
- Roles and responsibilities of the evaluation project team members.

❖ **For More Information**
See Appendix B for a sample Evaluation Charter (to create an Evaluation Plan).

3.2.2 Review Research Ethics Policy

FH personnel should consult the addendum to the FH "Studies Not Requiring Ethical Review", which gives insight regarding how the policy applies to program evaluation.

❖ **For Consideration**
FH Research Ethics Policy is available at:
http://www.fraserhealth.ca/Professionals/Research/Pages/default.aspx

---

3.2.3 Develop a Contract and Budget

The program may choose to contract with an Evaluator (internal or external) to plan and conduct the evaluation. A more plausible scenario within Fraser Health, however, is that the program completes the evaluation plan and then contract with an external Evaluator to conduct the evaluation.

Cost-effectiveness and **fiscal responsibility** are two pertinent concerns when contracting with an external Evaluator. Here are some tips:

- The terms of the contract should detail how the Evaluator will produce information of sufficient value, so that the resources expended can be justified.

- The allocation and expenditure of resources within the contract must be prudent and ethically responsible. Expected expenses must be described (obtain quotes if applicable).

❖ **For More Information**

When writing a contract for an external Evaluator, refer to the "Evaluation Contract Checklist":
http://www.wmich.edu/evalctr/checklists/contracts.pdf

Also consider the "Checklist for Developing and Evaluating Evaluation Budgets"
http://www.wmich.edu/evalctr/checklists/evaluationbudgets.pdf

3.2.4 Tasks and Timelines

Example 1 (Plan and Conduct Evaluation): The following brief work plan suggests how the Evaluator and Project Team could complete an evaluation over four months. This is the minimum suggested timeframe for completing an evaluation. Data collection and analysis will vary depending on the complexity of the methodology.

The schedule for completion dates assumes:

- Evaluator (internal or external) works full-time on the evaluation;
- Project Team and (if applicable Steering Committee) provides feedback and (if applicable) approves deliverables within one week of presentation;
- Data collection is possible to complete in two weeks;
- A final written evaluation report is necessary.

Deliverable	**Responsibility**	**Completion Date**
Gather and review documentation	Evaluator	Week 1
Prepare Project Description	Evaluator	Week 2
Review and Provide Feedback	Project Team	Week 3
Deliverable 1:	**Project Description**	**Week 3**
Prepare Evaluation Plan	Evaluator	Week 4
Approve Evaluation Plan	Project Team	Week 5
Deliverable 2:	**Evaluation Plan**	**Week 5**
Data Collection Phase	Evaluator	Week 6-8
Data Analysis	Evaluator	Week 8-9
Present Preliminary Findings	Evaluator	Week 10
Deliverable 3:	**Preliminary Findings**	**Week 10**
Present Draft Report	Evaluator	Week 11
Deliverable 4:	**Draft Report**	**Week 11**

Develop Recommendations	Evaluator /Team	Week 12
Prepare Final Report	Evaluator	Week 13-14
Deliverable 5:	**Final Report**	**Week 14**
Present Evaluation Report	Evaluator or Team	Week 15

Example 2 (Plan Evaluation): The following work plan suggests how the tasks required for developing and writing an evaluation plan could be completed over a period of six months.

The schedule for completion dates assumes:

- Evaluator (internal or external) works one day a week on the planning process;
- Project team meets every two weeks, alternating between half day face-to-face meetings (working sessions) and one hour follow-up planning teleconferences;
- Project Team and (if applicable Steering Committee) provides feedback and (if applicable) approves deliverables within one week of presentation.
- A final written evaluation plan is necessary.

Meetings and Goals	**Date**
Meeting #1: Logic Model Workshop and discussion	Month 1
Teleconference #1: allocate work on Program Description	mid - Month 1
Meeting #2: Review Program Description and Logic Model	Month 2
Teleconference #2: Follow-up to Meeting #2	mid - Month 2
Meeting #3: Discuss Evaluation Questions, Audience, Use	Month 3
Teleconference #3: Follow-up to Meeting #3	mid - Month 3
Meeting #4: Prioritize Evaluation Questions	Month 4
Teleconference #4: Follow-up to Meeting #4	mid - Month 4
Meeting #5: Data Collection Tools, Evaluation Design	Month 5
Teleconference #5: Follow-up to Meeting #5	mid - Month 5
Meeting #6: Final Review of Evaluation Plan	Month 6
Teleconference #6: Follow-up to Meeting #6	mid - Month 6
Deliverable: Evaluation Plan	**Month 6**

3.3 Checklist of Evaluation Standards – Preparing for Evaluation

The following Evaluation Standards and checkpoints apply to evaluation planning:[23]

- **Stakeholder Identification:** Are persons involved in or affected by the evaluation identified, so that their needs can be addressed?

  - Clearly identify the stakeholders, including primary intended user(s);
  - Engage leadership figures to identify stakeholders;
  - Consult stakeholders to identify their information needs;
  - Ask stakeholders to identify other stakeholders;
  - Arrange to involve stakeholders throughout the evaluation;
  - Keep the evaluation open to serve newly identified stakeholders.

- **Evaluator Credibility:** Are the persons planning and conducting the evaluation will be both trustworthy and competent to perform the evaluation, so that the evaluation findings achieve maximum credibility and acceptance?

  - Engage a competent Evaluator(s);
  - Engage an Evaluator whom the stakeholders trust;
  - Engage Evaluator(s) who can address stakeholders' concerns;
  - Help stakeholders understand the evaluation process;
  - Attend appropriately to stakeholders' criticisms and suggestions.

- **Formal Agreements:** Have obligations of the formal parties to an evaluation (e.g. Work plan for Evaluation Project Team, contract with an external Evaluator, Terms of Reference for Steering Committee) been agreed to in writing?

  Reach written agreements on:
  - Evaluation purpose and questions;
  - Audiences;
  - Editing;
  - Release of reports;
  - Evaluation procedures and schedule;
  - Evaluation resources.

- **Human Interactions:** Have individuals involved in the evaluation shown respect for human dignity, so that no person is threatened or harmed?

  - Consistently relate to all stakeholders in a professional manner;
  - Honour time commitments;

[23] There are 30 standards in total (relating to utility, feasibility, propriety, accuracy). The standards indicated above were selected because they are relevant to this section of the document. The checkpoints are reproduced from Stufflebeam, D. L. (1999).

- Be sensitive to stakeholders' and participants' diversity of values and cultural differences;
- Be evenly respectful in addressing different stakeholders;
- Do not ignore or help cover up any individual's incompetence, unethical behaviour, fraud, waste, or abuse.

- **Conflict of Interest:** Has any conflict of interest been dealt with openly and honestly, so that it does not compromise the evaluation processes and results?
  - Identify potential conflicts of interest early in the evaluation;
  - As appropriate and feasible, engage multiple evaluators;
  - Maintain evaluation records for independent review (e.g. records of discussion or decision);
  - If feasible, have the funding authority (e.g. FH portfolio) establish the contract for an external Evaluator, rather than the funded program;
  - If feasible, have the Evaluator report directly to the Director or Executive Director of the program;
  - Engage uniquely qualified persons to participate in the evaluation;
  - Take steps to counteract conflict of interest.

- **Fiscal Responsibility**: Does the allocation and expenditure of resources reflect sound accountability procedures so that expenditures are accounted for and appropriate?
  - Specify and budget for expense items in advance;
  - Keep the budget sufficiently flexible to permit appropriate reallocations to strengthen the evaluation;
  - Plan to maintain accurate records of sources of funding and expenditures and resulting evaluation services and products;
  - Plan to maintain adequate personnel record concerning job allocations and time spent on the evaluation project;
  - Plan to be frugal in expending evaluation resources.

4.0 Develop an Evaluation Plan

This section describes components of an Evaluation Plan and could be used as a template to create an Evaluation Plan. The main sections, described below, are:

- Program Profile
- Program Theory and Logic Model
- Primary Intended User and Intended Use
- Evaluation Questions
- Evaluation Methodology
- Communication Plan to Disseminate Evaluation Plan and/or Evaluation Findings

4.1 Program Profile

The Evaluation Plan begins with a Program Profile that describes the rationale for the program, the context, main goals, the evaluation stakeholders and primary intended users and program recipients (target population). The components of a program profile are described below.

4.1.1 Background and Context

The issue that the program seeks to address exists within a particular context or situation of social, political, environmental, and economic conditions.

- Why was the program established?
- What do existing research and experience say about this issue?
- Are there related studies that provide rationale for this program?
- Are there programs that offer similar or complementary services?
- What factors and trends in the larger environment may influence program success or failure?
- Who are the evaluation stakeholders and primary intended users and their associated organizations?

The stage of development of the program offers important contextual information – it influences the type of evaluation questions posed and the way that evaluation results may be used.

- Is the program just getting started?
- How long is the implementation stage expected to last?
- Has the program been underway for a significant period of time?

❖ **For Consideration**

A search of published and grey literature can help identify similar program evaluations and this information could save time in completing the Evaluation Plan (identifying outcomes, evaluation questions and methodologies).

Contact **Fraser Health Library Services**!

4.1.2 Program Goals

Describe the changes that the program aims to achieve.

- What is the program striving to achieve in the long-term?
- How do these program goals link with strategic goals of the organization?

4.1.3 Target Population

The **target population** is the set of individuals and /or organizations that an activity is intended to influence.

Describe the target population (individual, household, group, community) with detailed demographic information.

- Who is the program intended to serve?
- How will the program benefit the target population?

4.2 Program Theory and Logic Model

4.2.1 Program Theory and Program Components

Program theory describes how the program works. It describes the relationships and assumptions about "planned work" (inputs and activities) and "intended results" (outputs and outcomes).

The program theory is summarized into a simplistic graphical description in the Logic Model (below).

The program theory should provide answers to these broad questions:

- How are the planned activities and outputs going to lead to the intended changes for the target population(s)?
- What assumptions have been made about the program components and how do they link together?

A program can be described by components: inputs, activities, outputs and outcomes.

Inputs (Resources)

Program inputs are resources required to implement activities necessary to accomplish intended outcomes. Inputs are the financial and non-financial resources used by organizations, policies, programs and initiatives to produce outputs and accomplish outcomes.

Inputs may include:

- Funds
- Personnel
- Equipment and/or facilities

- Supplies
- Partnerships
- Research
- Best practices (e.g. clinical guidelines, policies, procedures)

Activities

Activities are the tasks, operations or work processes internal to an organization or program that use inputs to produce outputs. Activities may include:

- Clinical activities: assessment, consultation, counselling, treatment, education
- Research
- Construction
- Management
- Planning activities and policy creation
- On-going data collection and performance measurement
- On-going gap analysis

Outputs

Outputs are services or products generated by program activities; they are usually within the control of the organization and can be controlled by modifying program activities. Outputs may include:

- Procedures completed
- Care/treatment plan completed
- Creation of documents (e.g. manuals)
- Health promotion pamphlets delivered
- Immunizations delivered

Outcomes

Outcomes are the intended results or goals of a program. Outcomes are not within the control of a single program, but can be influenced by a program. Outcomes are usually further qualified as immediate, intermediate, or final (ultimate). Immediate outcomes are strongly influenced by a program, and therefore can be strongly attributed to the program. Final outcomes are influenced by more factors and therefore are less attributable to the program.

To identify outcomes, consider:

- How and in what way does the target population need to change?
- What specific action does the target population need to take?

Tips for articulating outcomes:

- Outcomes are written as change statements (e.g. increase, decrease, or stay the same);
- Outcomes can be qualified depending on the degree to which a program can reasonably assume control and responsibility;[24]
- Formulate outcomes as positive statements because it will be easier to gather consensus amongst program representatives by speaking positively;

[24] These categories of outcomes are consistent with those used by the Treasury Board of Canada Secretariat, which were also adopted by both cited CHSPR documents.

- Each outcome should capture only one improvement area.

Immediate outcomes:

- Outcomes for which a program can reasonably assume direct control and responsibility;
- Expected to arise in the short-term (e.g. 1-2 years);
- Are achievable within the funding and reporting periods specified;
- For example, an increase in awareness among a target population.

Intermediate outcomes:

- Outcomes for which a program can assume less direct control and responsibility;
- Consequence of one or more immediate outcomes;
- Expected to occur in the medium-term after one or more immediate outcomes have been achieved (e.g. 3-4 years);
- For example, a behavioural change among a target population.

Final outcomes:[25]

- Consequence of one or more intermediate outcomes;
- Highest-level outcomes that can be reasonably attributed to outputs;
- Expected to occur in the long-term (e.g. 5+ years);
- Represents the main goal of a program;
- Contribute to the strategic priorities of the organization;
- For example, a change of state/condition among a target population.

4.2.2 Logic Model

A logic model is a graphical description of how a program is intended to work (program theory). The program components (inputs, activities, outputs, outcomes) comprise the logic model. The logic model is supported by the Program Profile that explains the purpose, context and assumptions about the relationship between activities, outputs, and outcome.

The process of creating a logic model may begin with activities (for established programs) or with final outcomes (for new programs), depending on the program representative perspectives. The logic model diagram can present activities or final outcomes first, in a horizontal or vertical arrangement. There is "no one way to do it"!

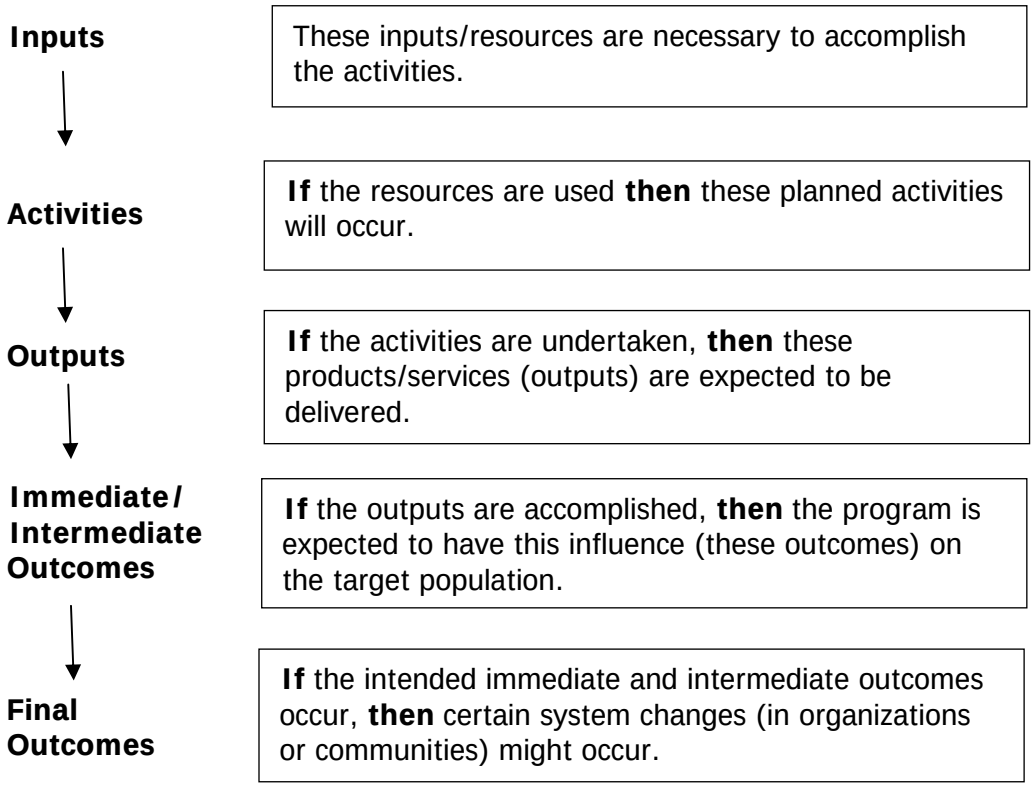
Common types of program inputs/outputs/outcomes can be grouped together for easier graphical representation in the logic model. For instance, outcomes could be classified as client-focused, system-focused or employee-focused.

The figure below describes how the logic model connects program components by a series of "If…Then" statements.[26]

[25] The W.K. Kellogg "Logic Model Development Guide" uses the term "impact" to describe the fundamental change occurring within 7-10 years.

[26] W.K.K. Foundation (2004).

Figure: Logic Model as a Series of Causal Links



4.2.3 Why develop a Logic Model?

The logic model is an important tool that can be used for:

- **Achieving consensu**s: Logic model development is a process by which program staff, program funders, or other stakeholders can achieve consensus about the intended outcomes and target population of the program.

- **Planning**: By clearly defining the intended outcomes, a logic model can enable program planners to plan activities for targeted people to achieve these outcomes.

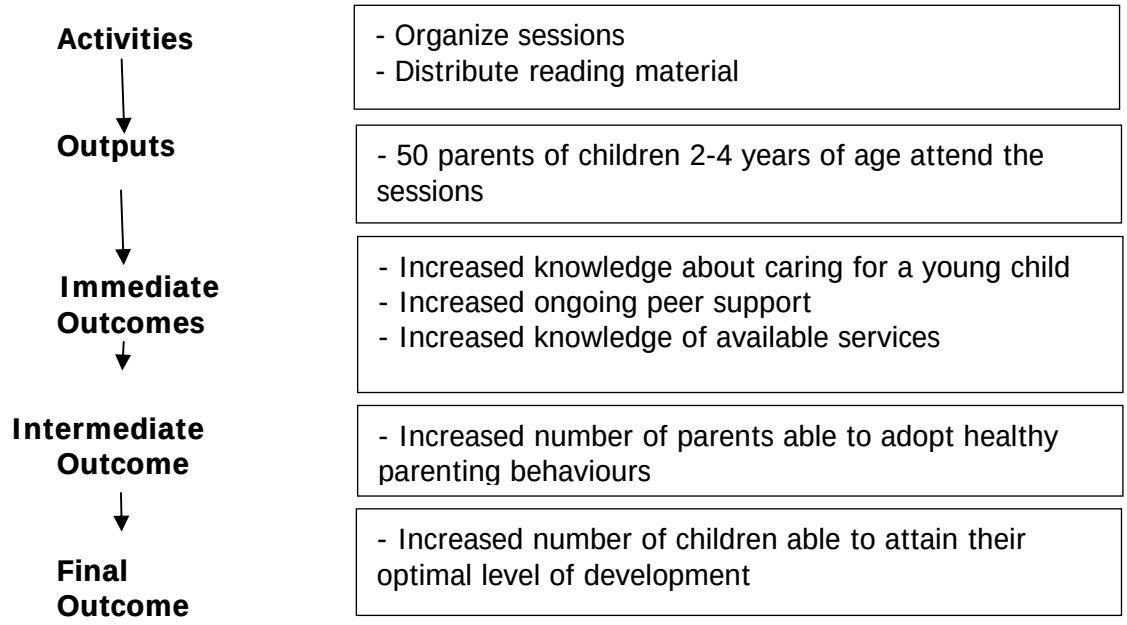
- **Program management**: A logic model displays the connections between resources, activities and outcomes and thereby is a basis for developing a more detailed management plan. During program implementation, a logic model can be used to explain, track and monitor operations, processes and functions.

- **Evaluation:** The logic model can guide the selection of performance indicators and help articulate evaluation questions that test and verify assumptions about how the program works.

- **Communication:** A simple graphic representation helps to communicate about the program with program staff, program funders, or other stakeholders (and to reach consensus about the intended outcomes).

❖ **An Example in Healthcare**

Program: Educational Parenting Program
Target Population: Parents of children 2 to 4 years of age



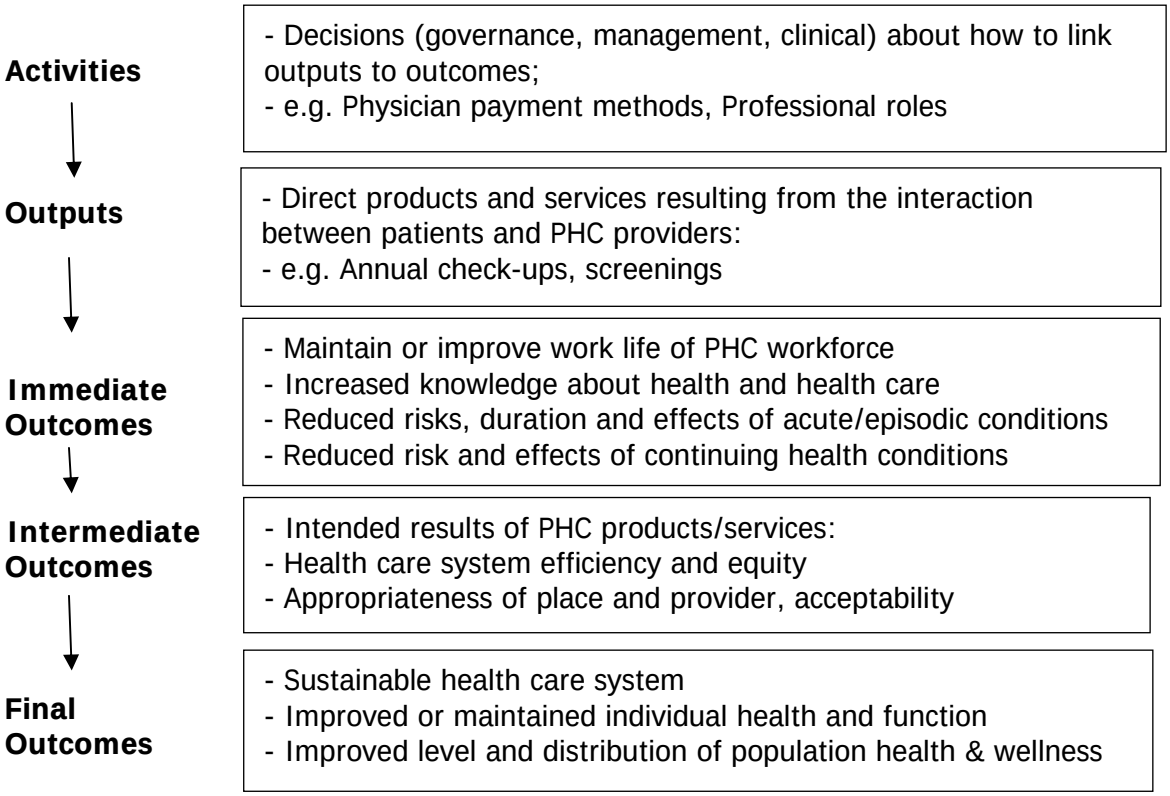
❖ **An Example in Healthcare**

In 2000, $800 million from the federal government was designated for the Primary Health Care Transition Fund. To support the National Evaluation Strategy for this fund, the University of British Columbia Centre for Health Services and Policy Research (CHSPR) developed "A Results-Based Logic Model for Primary Health Care". [27]

This logic model describes Canada's Primary Health Care (PHC) system and serves to establish a common understanding among stakeholders regarding: [28]

- Inputs, activities, outputs and outcomes of the PHC system;
- Outcomes that stakeholders are accountable for (e.g. immediate outcomes).

Primary Health Care System in Canada – Logic Model



[27] Watson, D., Broemeling, A., Reid, R., Black, C. (2004).
[28] Summarized from Broemeling, A. M., Watson, D. E., Black, C., & Reid, R. J. (2006).

❖ **An Example in Healthcare**

There are many different ways to design a logic model. They may vary in:

- Scope of logic model (how much they cover);
- Number of levels included;
- Description of levels included;
- Direction of information flow;
- Amount of text; and
- Visual layout.

The Health Communication Unit at the Centre for Health Promotion, University of Toronto discusses these factors and provides examples. The "Logic Models Workbook" is available at:
http://www.thcu.ca/infoandresources/resource_display.cfm?res_topicID=4

❖ **For Consideration**

While logic models present a quick overview, they have been criticised for presenting a closed system image (which rarely occurs).

The International Development Research Centre (a Canadian public corporation) has researched and developed an alternative logic model, "Temporal Logic Model", which includes program responsiveness to environmental changes and the organizational learning process. This report is available at:
http://www.idrc.ca/uploads/user-S/10553603900tlmconceptpaper.pdf

❖ **Additional References**

There are numerous other references to guide logic model development.

- University of Wisconsin offers an on-line course "Enhancing Program Performance with Logic Models", available at:
  http://www.uwex.edu/ces/lmcourse/

- The W.K. Kellogg Foundation "Logic Model Development Guide" is available at http://www.wkkf.org/. Enter "Logic Model Development Guide" in the search box. Also see the W.K. Kellogg Foundation "Evaluation Handbook".

- The Centre of Excellence for Child and Youth Mental Health at Children's Hospital of Eastern Ontario "Program Logic Model" described in "Doing More with Program Evaluation" Toolkit, available at:
  http://www.onthepoint.ca/kec/know.htm

4.3 Primary Intended User and Intended Use

4.3.1 Identify the Primary Intended User(s)

Primary intended users are those individuals who have responsibility to apply evaluation findings (to improve the program or to make judgements about the program) and implement recommendations.

To identify primary intended users consider:

- Who has requested the evaluation (e.g. funding agency, decision makers)?
- Who will use the results of the evaluation?
- Who are the key stakeholders that will be affected by the findings?
- Who would potentially be interested in the findings once they are known?

As stated previously, the practice of designing the evaluation according to the information needs of the primary intended users of the evaluation is known as the practice of **utilization-focused evaluation**. [29]

4.3.2 Identify the Intended Use(s)

The intended use of the evaluation findings influences the way that an evaluation is designed, and therefore the primary intended user(s) must agree on the intended use in the evaluation planning stage.[30]

Related to the intended uses of evaluation, two primary reasons to conduct evaluation are to: [31]

- **Improve programs** and inform decisions about improving process or design.

- **Make judgements** about the overall merit, worth and value of the program; often involving examination of its success (intended and unintended outcomes) and used to inform decisions about a program (should it be continued or terminated?).

The common evaluation approach to improve programs is described as **formative evaluation**. The common evaluation approach to make judgements is described as **summative evaluation**.

---

[29] Patton (2008)

[30] While the literature often distinguishes between "intended use" and "purpose" of evaluation, these terms are considered the same for the purpose of this document.

[31] Patton, M.Q. (2008) describes four additional purposes of evaluation: accountability reporting, monitoring systems, generating generic knowledge and developmental evaluation (p.140).

4.4 Evaluation Questions

4.4.1 Potential Evaluation Questions

To identify potential evaluation questions, interview stakeholders to determine different perspectives. If possible, allow flexibility for adding questions during the evaluation.

If the purpose of the evaluation is for **learning and to improve the program**, consider these evaluation questions:[32]

- Are the activities being implemented as planned?
- What works and what does not work? Strengths and weaknesses?
- Participants' reactions?
- What works for whom in what ways and under what conditions?
- How can outcomes and impacts be increased?
- How can costs be reduced?
- How can quality be enhanced?

If the purpose of the evaluation is to **judge the overall value and to inform major decision-making** about the value and future of the program and model, consider these evaluation questions:[33]

- Does the program meet participants' needs? Is there a gap between the intended and actual population served?
- To what extent does the program have merit or worth?
- How do outcomes and costs compare with other options?
- To what extent can outcomes be attributed to the intervention?
- Is the program theory clear and supported by findings?
- Is this an especially effective practice that should be funded and disseminated as a model program?

❖ **For Consideration**

Program relevance and **performance (effectiveness and efficiency)** are the core evaluation issues in the Government of Canada's Policy on Evaluation. Evaluation questions related to program relevance and performance could be:

- Is there a continued need for the program (?)
- Are the program goals aligned with government (organizational) priorities?
- Has the program achieved the expected outcomes?
- Has the program operated in an efficient and cost-effective manner?

The Government of Canada's Policy on Evaluation (2009) is available at: http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=15024

❖ **For More Information**

See Appendix C "Examples of Evaluation Questions"

---

[32] Patton, M.Q. (2008).
[33] Patton, M.Q. (2008).

4.4.2 Prioritize and Select Evaluation Questions

There will typically be many questions about a program, more than can be answered given limits on available resources for the evaluation (time, money, and expertise). First consider questions identified by the primary intended user(s) and expect some negotiation regarding the choice.

Consider the following factors that influence the choice of evaluation questions:[34]

- **Age of the Program:** Questions about improvement and implementation are important for a new program. Established programs are more likely to be able to investigate success and impact.

- **Resources:** Some questions may be more costly to answer than others. Lack of data or expertise (personnel) may prohibit answering some questions.

- **Knowledge and Values:** Knowledge about the program and the value placed on the program's success will influence the choice and priority of evaluation questions.

- **Consensus:** Are there questions that all stakeholders agree should be answered? This approach promotes group participation, but important questions may be avoided.

- **Result Scenarios:** Will the answers to these questions change people's beliefs, attitudes, or behaviour about the program?

- **Funding:** What questions does the funder of the evaluation want to pay for? This approach may leave key questions unanswered.

4.5 Evaluation Methodology

Evaluation methodology refers to what data is collected and the research design that defines how it is collected.

4.5.1 Data Collection

Data used for evaluation may be quantitative, qualitative, or both (mixed-methods). For each evaluation question, use multiple sources of information (triangulation of evidence) to answer an evaluation question. **Triangulation** is a best practice because it improves the internal validity of evaluation findings and recommendations.

Quantitative data (numerical) can be collected by using:

- Rating scales;
- Tests;
- Chart Audits;

[34] Grembowski, D. (2001)

- Administrative data;
- Descriptive statistics (summarize/describe a collection of data);
- Performance measurement data (See Appendix A).

Statistical analysis may be necessary to draw conclusions about quantitative data. Regression analysis, for instance, is used for prediction (including forecasting), inference, hypothesis testing, and understanding causal relationships.

Qualitative data (words and images) is used to explain, understand or interpret the people or situations being studied. It can be collected by:

- Observation;
- Conversation analysis;
- Individual interviews (e.g. participants, experts, stakeholders);
- Surveys with open-ended questions;
- Focus groups;
- Unstructured diary-keeping and journals (secondary data);
- Documents (literature review, documentary analysis) (secondary data).

A mixed-method approach combines qualitative and quantitative sources of information. Combining qualitative and quantitative data improves the evaluator's understanding and therefore strengthens the internal validity of evaluation findings. For instance, an evaluation may use qualitative data collected through a focus group to better inform issues that should be explored in a survey of program participants.

❖ **Additional References**

The Research Methods Knowledge Base is a comprehensive web-based textbook that addresses all social research methods. This resource is available at: http://www.socialresearchmethods.net/kb/introval.php

Qualitative analysis can include action research, grounded theory, ethnography and narrative analysis. See "Introduction to Qualitative Data Analysis": http://onlineqda.hud.ac.uk/Intro_QDA/index.php

The (UK) National Centre for Social Research Criteria offers criteria to assess the quality of four qualitative research methods (in-depth interviews, focus groups, observation and documentary analysis). See "Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence" available at: http://www.cabinetoffice.gov.uk/strategy/publications.aspx

Examples of qualitative research software are offered by the American Evaluation Association: http://www.eval.org/Resources/QDA.htm

See the "User-Friendly Handbook for Mixed Method Evaluations" produced by the U.S. National Science Foundation, available at: http://www.nsf.gov/pubs/1997/nsf97153/start.htm

---

4.5.2 Evaluation (Research) Design

There are three basic types of research design, each of which may be used in evaluation: experimental, quasi-experimental, and non-experimental.

Experimental and quasi-experimental designs seek to describe an association between the program and its outcomes. These research designs can provide evidence **to judge and make decisions about the program**.

- **Experimental**
  - Compares treatment (client) group with a control group;
  - Uses randomization;
  - Can estimate the probability that there is a causal link between the program and observed outcomes;

- **Quasi-experimental**
  - Compares a treatment group with a comparison group;
  - Does not use randomization;
  - May include natural groups not under control of the program;
  - Contributes to an understanding of the causal link between program and observed outcomes;

Non-experimental design includes descriptive design, which is used to answer evaluation questions to **learn about the program and improve the program**; it does not seek to describe a causal association between observed outcomes.

- **Non-experimental**
  - Control or comparison groups are not used;
  - Data analysis (e.g. regression), correlation, trend analysis, descriptive statistics;
  - Involves measuring the effects of a program after it has been implemented;
  - Descriptive design (e.g. using qualitative data) of the program;
  - May contribute to an understanding of the causal link between program and observed outcomes.

Quasi-experimental and non-experimental (including descriptive designs) are most often used to evaluate health care initiatives.[35]

❖ **For More Information**
Appendix D offers more information about experimental, quasi-experimental and non-experimental design and examples in healthcare.

❖ **For Consideration**
Economic evaluation is a type of analysis that explicitly connects the cost of programs to their outcomes to assist decision makers in all resource allocation.

[35] Linden et al. (2006)

Causality between costs and outcomes must be determined prior to economic evaluation. Cost-benefit analysis, cost-effectiveness analysis and cost-utility analysis are the three types of economic evaluation.

❖ **An Example in Healthcare**

The Michael Smith Foundation for Health Research funded a series of training modules that discuss **economic evaluation**, critical appraisal of economic evaluations, cost effectiveness analysis, budget impact analyses, patient reported outcomes and health related quality of life.

These modules are available at:

- http://spectra.tosm.ttu.edu/econeval
- http://spectra.tosm.ttu.edu/econeval2
- http://spectra.tosm.ttu.edu/econeval3

4.5.3 Considerations for the Evaluation Methodology

The evaluation methodology should be designed to provide the information to answer the evaluation questions of intended users. Both complex and simple evaluations should be equally rigorous in relating the methodology to the intended use of the evaluation.[36]

For instance, if the intended use of the evaluation is to inform a decision-making process (e.g. continue, expand or terminate funding), then the complexity of the evaluation methodology needs to reflect the impact of the decisions that will be taken.

In planning the evaluation, consider:[37]

- What information is sufficient to address the most important evaluation questions?
- Of this information, how much can be collected and analyzed in a low-cost and practical manner, using questionnaires, surveys and checklists?
- How accurate will the information be?
- Will the methods get all of the needed information?
- What additional methods should and could be used if additional information is needed?
- Will the information appear as credible to decision makers, to funders or management?
- Will the nature of the audience conform to the methods of data collection? For instance, will they fill out questionnaires carefully, engage in interviews or focus groups, or allow their documentation to be examined?
- Who can administer the data collection tools? Is training required?

[36] Habicht, J. P., Victora, C. G., & Vaughan, J. P. (1999).
[37] Reproduced from McNamara, C. (2008).

- How can the information be analyzed?
- What are the limitations of the selected methodology?

The evaluation methodology must be **affordable**. Evaluation costs increase rapidly with complexity and therefore a compromise must be made between rigor and cost. The resources required to develop and implement the evaluation plan could be employed in other ways (e.g. program delivery) and so the perceived benefit of answering evaluation questions must outweigh the costs.

- What are the estimated costs for data collection (e.g. human resources, software licensing), including provisions to ensure data integrity?
- What are the estimated costs for undertaking the evaluation?

Ideally, the evaluation methodology will:
- Provide the appropriate type of information to answer the evaluation questions;
- Balance the information needs with the cost of seeking that information;
- Consider how to use existing data and secondary data;
- Use triangulation (multiple lines of evidence) to ensure reliability of findings and conclusions.

❖ **For Consideration**

Evaluation designs can often be modified to meet resource constraints:
- Simplify evaluation design;
- Reconstruct baseline data;
- Reduce the costs of data collection.

See "Conducting Quality Impact Evaluations under Budget, Time and Data Constraints". http://www.worldbank.org/oed/ecd/conduct_qual_impact_eval.html

❖ **An Example in Healthcare**

Health Canada created a "Performance Measurement and Evaluation Framework" for the Strategic Training Initiative in Health (STIHR), which is available at http://www.cihr-irsc.gc.ca/e/27737.html.

Below is an excerpt that summarizes the methods used to collect data for one evaluation question that relates to a broader evaluation issue.

Evaluation Issue: Implications of STIHR selection processes for the potential success of funded STIHR programs		
Evaluation Questions	**Indicators**	**Data Sources**
What is the nature of the interplay between strategic and excellence-driven funding within the strategic research areas with STIHR funding, and what are the implication of this for overall likelihood of program success?	Degree of successful resolution of tension Effectiveness of reconciliation strategies	Survey of grantees Survey of Institutes and partners

❖ **An Example in Health Care**

The Canadian Nurses Association has developed a toolkit for implementation and evaluation of a nurse practitioner program. This resource describes the interdependency of the planning, performance measurement and evaluation.

See "Implementation and Evaluation Toolkit for Nurse Practitioners in Canada", http://23072.vws.magma.ca/CNA/practice/advanced/initiative/default_e.aspx

---

4.6 Develop a Communications Plan

One main goal of evaluation is to produce and disseminate information that is useful for primary intended users. The process to develop "useful" information started when primary intended users and other stakeholders were engaged in identifying the intended use of the evaluation and the evaluation methodology.

The likelihood that evaluation findings are used is improved when evaluation findings are communicated directly with intended users of the evaluation (e.g. managers, decision-makers).

Develop a communications plan to ensure that primary intended users receive the evaluation findings in a timely and appropriate format, according to their needs.

During the evaluation process, consider:

- Interim evaluation reports;
- Written reports with ongoing oral communication.

The communications plan may need to accommodate for primary intended users who are internal and external to the organization. Consider the intended users and, more broadly, the evaluation stakeholders of the evaluation findings, and then specify:

- What types of findings would be relevant to each audience?
- Who should deliver the message?
- How should the message be delivered?
- What is the expected impact or intended use of the message?

The level of detail and presentation style of the information will likely need to be adapted to meet the needs of different audiences. For example, decision makers may prefer an executive summary that describes the "bottom line", in which case it is the evaluator's job to summarize data analysis and technical language into concise sentences that can be easily understood.

Consider planning (and budgeting) for other communication activities, including:

- Feedback workshops among program staff to review and apply findings and implement recommendations;
- Follow-up with the evaluator to interpret and apply findings.

4.7 Checklist of Evaluation Standards – Develop an Evaluation Plan

The following Evaluation Standards and checkpoints apply to developing an evaluation plan:[38]

- **Information Scope and Selection:** Will the information collected address pertinent questions about the program and be responsive to the needs of primary intended user(s)?

  - Understand the intended use of the evaluation findings and priority evaluation questions;
  - Interview primary intended user(s) as well as other stakeholders to determine different perspectives;
  - Assign priority to questions identified by the primary intended user(s);
  - Plan to obtain sufficient information to address the most important evaluation questions.

- **Plan for Evaluation Impact**: Is the evaluation planned in ways that encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased?

  - Involve stakeholders throughout the evaluation planning;
  - Discuss with stakeholders how they might use the findings in their work;
  - Create a communications plan for internal and external audiences;
  - Consider hosting feedback workshops to go over and apply findings;
  - Consider making arrangements for the evaluator to provide follow-up assistance in interpreting and applying the findings.

- **Political Viability:** Is the evaluation planned with anticipation of the different positions of stakeholder groups, so that their cooperation may be obtained, and so that possible attempts by any of these groups to curtail evaluation operations or to bias or misapply the results can be averted or counteracted?

  - Anticipate different positions of different stakeholder groups;
  - Plan to avert or counteract attempts to bias or misapply the findings;
  - Foster cooperation among stakeholders;
  - Involve stakeholders throughout evaluation planning;
  - Agree on editorial and dissemination authority;
  - Plan to report to "right-to-know" audiences.

[38] There are 30 standards in total (relating to utility, feasibility, propriety, accuracy). The standards indicated above were selected because they are relevant to this section of the document. The checkpoints are reproduced from Stufflebeam, D. L. (1999).

- **Practical Procedures:**
  - Tailor methods and instruments to information requirements;
  - Plan to minimize disruption during data collection;
  - Plan to minimize the data burden during data collection;
  - If applicable, choose data collection procedures that the staff are qualified to carry out;
  - Choose data collection procedures in light of known constraints;
  - Make a realistic schedule;
  - As appropriate, plan data collection and other evaluation procedures a part of routine events.

- **Cost Effectiveness:** Is the evaluation expected to be efficient and able to produce information of sufficient value, so that the resources expended can be justified?

  Consider how the evaluation is planned to:
  - Make use of in-kind services;
  - Produce information that inform decisions or fosters program improvement;
  - Generate new insights about the program;
  - Help spread effective practices within the program;
  - Minimize disruptions during data collection;
  - Minimize time demands on program personnel during data collection.

- **Service Orientation**: Is the evaluation designed to address and effectively serve the needs of the full range of targeted participants? (may not apply to all evaluations)

  - Assess needs of the program's customers;
  - Assess program outcomes against targeted customers' assessed needs;
  - Identify program strengths to build on;
  - Identify program weaknesses to correct;
  - Give interim feedback for program improvement;
  - Expose harmful practices;
  - Inform all right-to-know audiences of the program's positive and negative outcomes.

- **Rights of Human Subjects:** Is the evaluation designed to respect and protect the rights and welfare of human subjects?

  - Make clear to stakeholders that the evaluation will respect and protect the rights of human subjects;
  - Clarify intended uses of the evaluation;
  - Keep stakeholders informed;
  - Understand values of human subjects;
  - Respect diversity;
  - Honour confidentiality/anonymity agreements;
  - Do no harm to human subjects.

- **Human Interactions:** Have individuals involved in the evaluation shown respect for human dignity, so that no person is threatened or harmed?

    - Consistently relate to all stakeholders in a professional manner;
    - Honour time commitments;
    - Be sensitive to stakeholders' and participants' diversity of values and cultural differences;
    - Be evenly respectful in addressing different stakeholders;
    - Do not ignore or help cover up any individual's incompetence, unethical behaviour, fraud, waste, or abuse.

- **Program Documentation**: Is the program described and documented clearly and accurately?

    - Collect descriptions of the intended program from various written sources;
    - Collect descriptions of the intended program from the client and various stakeholders;
    - Describe how the program was intended to function;
    - Maintain records from various sources of how the program operated;
    - As feasible, engage independent observers to describe the program's actual operations;
    - Describe how the program actually functioned;
    - Analyze discrepancies between the various descriptions of how the program was intended to function;
    - Analyze discrepancies between how the program was intended to operate and how it actually operated;
    - Ask the client and various stakeholders to assess the accuracy of recorded descriptions of both the intended and the actual program;
    - Produce a technical report that documents the program's operations.

- **Context Analysis**: Is the context in which the program exists examined in enough detail, so that its likely influences on the program can be identified?

    - Use multiple sources of information to describe the program's context;
    - Describe the program technical, social, political, organizational, and economic features;
    - Analyze how the program's context is similar to or different from contexts where the program might be adopted;
    - Report those contextual influences that appeared to significantly influence the program and that might be of interest to potential adopters;
    - Estimate effects of context on program outcomes;
    - Identify and describe any critical competitors to this program that functioned at the same time and in the program's environment;
    - Describe how people in the program's general area perceived the program's existence, importance, and quality.

- **Described Purposes and Procedures**: Are the purposes and procedures of the evaluation monitored and described in enough detail, so that they can be identified and assessed?

  - Record the primary intended users of the evaluation;
  - Describe the intended uses of evaluation findings;
  - Monitor and describe how the evaluation's purposes change over time;
  - Identify and assess points of agreement and disagreement among stakeholders regarding the evaluation's purposes;
  - Plan to describe the evaluation's purposes and procedures in the summary and full-length evaluation reports.

- **Defensible Information Sources**: Are the sources of information used in a program evaluation described in enough detail, so that the adequacy of the information can be assessed?

  - Obtain information from a variety of sources;
  - Use pertinent, previously collected information once validated;
  - As appropriate, employ a variety of data collection methods;
  - Document and report information sources;
  - Document, justify, and report the criteria and methods used to select information sources;
  - Include data collection instruments in a technical appendix to the evaluation report;
  - Document and report any biasing features in the obtained information.

- **Valid and Reliable Information**: Has the evaluation methodology been developed to assure that the interpretation of the information is valid and reliable for the intended use?

  - As appropriate, plan to triangulate information to address each question;
  - As feasible, choose measuring devices that in the past have shown acceptable levels of reliability for their intended uses.

5.0 Conduct Evaluation and Disseminate Findings

5.1 Review and Implement the Evaluation Plan

If a significant amount of time elapses between developing the evaluation plan and undertaking an evaluation, consider meeting the Evaluation Team, key stakeholders or the Steering Committee to determine:

- Are intended activities, outputs and outcomes still relevant to the target population?
- Are resources available to conduct the evaluation as planned?
- Does the Communications Plan need to be updated?
- Do any other sections of the Evaluation Plan need to be updated?

❖ **For Consideration**

Evaluability assessment describes a review (before undertaking evaluation, after an evaluation plan is complete) to determine whether the requirements for conducting an evaluation can be met (e.g. adequately defined program goals, available resources, timeline, data availability and data verification).

An evaluability assessment also asks the fundamental questions:

- Should an evaluation proceed?
- Are there limitations in the way the plans are expected to be implemented?

❖ **For More Information**

See Appendix E for more information about how to conduct an "Evaluability Assessment" to ensure that the evaluation can proceed as planned.

5.2 Developing Recommendations

Evaluation recommendations are statements of proposed change, based on the evidence from evaluation findings. [39]

Recommendations should meet the following criteria:

- **Defensible**: Link each recommendation to the evaluation findings and empirical evidence.
- **Timely**: Ensure recommendations are ready to be used in the decision making process.
- **Realistic**: Be reasonable. Recommendations that appear unfeasible may be ignored.
- **Targeted**: Indicate who has the authority to approve/disapprove each recommendation and who is responsible for implementation.

[39] Grembowski, D. (2001).

- **Simple**: Use clear, simple language to ensure recommendations are understood.

- **Specific**: Addresses only one idea in each recommendation. Organize recommendations into specific types of tasks or actions.

Best practices in developing recommendations include:[40]

- **Investing time:** Recommendations are often the most important part of the report, and therefore adequate time on their development is required.

- **Starting early:** At the beginning of the evaluation, it may be appropriate to propose anticipated evaluation findings and discuss potential recommendations with decision makers. When the evaluation is complete, the intended users are more likely to be familiar with the findings.

- **Consider all issues:** Consider recommendations for all types of evaluation findings – do not narrow the scope to a certain type of finding or area of concern.

- **Consider all sources:** Good recommendations can arise from difference sources. Consider previous evaluations, other reports, program staff and clients.

- **Work closely with decision makers and program staff:** Well-informed recommendations require input from program staff and management. By working closely and honestly with decision makers and program staff, the evaluator can build acceptance of the recommendations.

- **Consider whether recommendations should be general or specific:** It may be easier to achieve consensus on general recommendations that allow program management some control over how they are implemented. On the other hand, a recommendation may not be implemented without precise details about who/when/how to proceed.

- **Consider fundamental or incremental change:** Fundamental changes (e.g. changing program goals) is often complex. Recommending incremental changes, which lead to the same type of change, may be more feasible.

- **Consider program context:** Recommendations must fit with the program's political, cultural, social, and organizational contexts.

- **Describe potential costs and benefits:** If the potential costs and benefits can be described, a recommendation may seem more feasible.

- **Avoid recommending another evaluation:** Design the evaluation such that it is rigorous enough to answer the evaluation questions posed. Think forward, in planning the evaluation, to evaluation findings that may raise more questions than answers.

[40] Grembowski, D. (2001).

5.3 Reporting Style and Format

The types of reports (e.g. written or oral) should be defined in the Communications Plan. The purpose of this section is to present ideas about style, format, content, and the process of reporting information. These characteristics also influence the utility of evaluation findings.

Here are some general tips:

- Charts/graphics are essential to capturing attention and communicating quickly;
- Tone, content, and language of a key message needs to be appropriate for its intended audience;
- Communicate sensitive information carefully;
- Develop clear, simple, action-oriented messages.

Reports on the evaluation findings could follow a number of formats (written and oral). In fact, written and oral delivery could be combined, as appropriate.[41]

Formats for written reports include:

- Executive summary, followed by a full report;
- Executive summary, followed by a few key tables, graphs, and data summaries;
- Executive summary only, and make data available for those interested;
- Newsletter article for dissemination;
- Press release.

Formats for oral presentation include:

- Oral presentation with charts;
- Short presentation followed with question/answer period;
- Discussion groups based on prepared written material;
- Retreat-like session with intended users;
- Video or audio taped presentation;
- Debate session regarding certain conclusions/judgements;
- Involve selected primary users in reporting and facilitating any of the above modes of oral presentation.

❖ **Additional References**

The Treasury Board of Canada Secretariat offers a template for an evaluation report, the "Guide for the Review of Evaluation Reports" is available at: http://www.tbs-sct.gc.ca/eval/tools_outils/4001752_e.asp

Canadian Health Services Research Foundation (CHSRF) recommends "1:3:25" reader-friendly writing.

- Start with one page of main messages;
- Follow with a three-page executive summary;
- Present the evaluation report in 25 pages or less.

Find the CHSRF Communications Notes regarding "Reader-friendly writing" at http://www.chsrf.ca/knowledge_transfer/resources_e.php

[41] Adapted from Menu 13.2 in "Evaluation Reporting Menu" in Patton, M. Q. (2008)

❖ **For More Information**

The Children's Hospital of Eastern Ontario has developed a Toolkit on Knowledge Exchange to help assess that the right methods are used at the right time and involve the right people. The Knowledge Exchange Checklist is available at: www.onthepoint.ca/kec/documents/Kechecklist.pdf

---

5.4 Using the Evaluation Findings and Recommendations

The use of evaluation findings (which may include implementation of recommendations) is likely more of a process than a single event. The purpose and expected use of evaluation findings is explored as part of the evaluation planning process and the Communications Plan should disseminate the information according to their intended use.

Different purposes of evaluation lead to different uses of evaluation findings. Evaluation findings can be used immediately in two ways:[42] [43]

- **Conceptual use**: The evaluation produces new information about the program and this information changes how people understand the program and how it works (e.g. how it serves the intended target population). This information may be used to change the program (e.g. make adjustments to better meet needs of target population), but are not directed at a particular decision about the future of the program.

- **Instrumental use**: Evaluation findings are directed at a particular decision for a specific program at a concrete point in time (e.g. end or expand a program).

For the purpose of this document, conceptual use is expected to arise from an evaluation designed to learn about and improve a program. Instrumental use is expected to result from an evaluation designed to make judgements about a program.

In the longer term, evaluation findings and the evaluation process can be used to **influence audiences** who have similar interests but are not connected to the program.

There are many factors that influence how (and if) evaluation findings are used (e.g. existing knowledge, beliefs, values, budget and time constraints). It is more likely that evaluation findings are used (and recommendations implemented) when:

- Intended users and use is accurately identified;
- Evaluation questions are answered in a clear way;
- Findings are accurate and relevant to intended users;
- Evaluation findings are communicated directly with intended users of the evaluation (E.g. managers, decision-makers).

---

[42] Grembowski, D. (2001).

[43] Patton, M.Q. (2008) describes process use or "the changes resulting from engagement in the evaluation process" as a third direct use of evaluation (p.112).

5.5 Final Checklist of Evaluation Standards – Conduct Evaluation, Report Findings

The following Evaluation Standards and checkpoints apply to conducting the evaluation, documentation and reporting results:[44]

- **Evaluator Credibility:** Are the persons conducting the evaluation both trustworthy and competent to perform the evaluation, so that the evaluation findings achieve maximum credibility and acceptance?

  - Engage competent evaluators;
  - Engage evaluators whom the stakeholders trust;
  - Engage evaluators who can address stakeholders' concerns;
  - Engage evaluators who are appropriately responsive to issues of gender, socioeconomic status, race, and language and cultural differences;
  - Help stakeholders understand and assess the evaluation plan and process;
  - Attend appropriately to stakeholders' criticisms and suggestions.

- **Values Identification:** Are the perspectives, procedures, and rationale used to interpret the findings carefully described, so that the bases for value judgments are clear?

  - Consider alternative sources of values for interpreting evaluation findings;
  - Provide a clear, defensible basis for value judgments;
  - Reference, as appropriate, the relevant institutional mission;
  - Reference the program's goals;
  - Take into account the stakeholders' values;
  - As appropriate, present alternative interpretations based on conflicting but credible value bases.

- **Report Clarity:** Did the evaluation report clearly describe the program being evaluated, including its context, and the purposes, procedures, and findings of the evaluation, so that essential information is provided and easily understood?

  - Clearly report the essential information;
  - Issue brief, simple, and direct reports;
  - Focus reports on contracted questions;
  - Describe the program and its context;
  - Describe the evaluation purpose, methodology, and findings;
  - Support conclusions and recommendations;
  - Avoid reporting technical jargon;
  - Report in the language(s) of stakeholders.

[44] There are 30 standards in total (relating to utility, feasibility, propriety, accuracy). The standards indicated above were selected because they are relevant to this section of the document. The checkpoints are reproduced from Stufflebeam, D. L. (1999).

- **Report Timeliness and Dissemination:** Were significant interim findings and evaluation reports disseminated to intended users, so that they could be used in a timely fashion?

  - Make timely interim reports to intended users;
  - Deliver the final report when it is needed;
  - Have timely exchanges with the intended users and other stakeholders;
  - Employ effective media for reaching and informing the different audiences;
  - Keep the presentations appropriately brief;
  - Use examples to help audiences relate the findings to practical situations.

- **Political Viability:** Was the evaluation conducted with anticipation of the different positions of stakeholder groups, so that their cooperation may be obtained?

  - Understand different positions of different stakeholder groups;
  - Avert or counteract attempts to bias or misapply the findings;
  - Foster cooperation among stakeholders;
  - Involve stakeholders throughout the process of conducting the evaluation;
  - Adhere to agreement regarding editorial and dissemination authority;
  - Report to "right-to-know" audiences.

- **Practical Procedures:**
  - Minimize disruption during data collection;
  - Make a realistic schedule;
  - As appropriate, undertake data collection and other evaluation procedures as part of routine events.

- **Evaluation Impact**: Was the evaluation conducted and evaluation findings reported in ways that encourage follow-through by stakeholders, so that the likelihood that the evaluation will be used is increased?

  - Involve stakeholders throughout the evaluation;
  - Encourage and support use of the findings;
  - Show stakeholders how they might use the findings in their work;
  - Provide interim reports;
  - Make sure that reports are open, frank, and concrete;
  - Supplement written reports with ongoing oral communication.

- **Human Interactions:** Have individuals involved in the evaluation shown respect for human dignity, so that no person is threatened or harmed?

  - Consistently relate to all stakeholders in a professional manner;
  - Honour time commitments;
  - Be sensitive to stakeholders' and participants' diversity of values and cultural differences;
  - Be evenly respectful in addressing different stakeholders;

- Do not ignore or help cover up any individual's incompetence, unethical behaviour, fraud, waste, or abuse.

- **Complete and Fair Assessment:** Was the evaluation complete and fair in its examination and recording of strengths and weaknesses of the program being evaluated, so that strengths can be built upon and problem areas addressed?
  - Assess and report the program's strengths;
  - Assess and report the program's weaknesses;
  - Report on intended outcomes;
  - Report on unintended outcomes;
  - Give a thorough account of the evaluation's process;
  - As appropriate, show how the program's strengths could be used to overcome its weaknesses;
  - Have the draft report reviewed (e.g. by the Evaluation Steering Committee);
  - Appropriately address criticisms of the draft report;
  - Acknowledge the final report's limitations;
  - Estimate and report the effects of the evaluation's limitations on the overall judgment of the program.

- **Valid and Reliable Information**: Has the evaluation methodology been implemented to assure that the interpretation of the information is valid and reliable for the intended use?

  - Document and report the data collection conditions and process;
  - Assess and report the comprehensiveness of the information provided by the methodology in relation to the information needed to answer the evaluation questions;
  - In reporting reliability of an instrument, assess and report the factors that influenced the reliability, including the characteristics of the participants, the data collection conditions, and the evaluator's biases;
  - Check and report the consistency of scoring, categorization, and coding;
  - Pilot test new instruments in order to identify and control sources of error;
  - Acknowledge reliability problems in the final report;
  - Estimate and report the effects of unreliability in the data on the overall judgment of the program.

- **Systematic Information**: Was the information collected, processed, and reported in an evaluation systematically reviewed?

  - Establish protocols for quality control of the evaluation information;
  - Train the evaluation staff to adhere to the data protocols;
  - Systematically check the accuracy of scoring and coding;
  - When feasible, use multiple evaluators and check the consistency of their work;
  - Verify data entry;
  - Proofread and verify data tables generated from computer output or other means;
  - Systematize and control storage of the evaluation information;

- Define who will have access to the evaluation information;
- Strictly control access to the evaluation information according to established protocols;
- Have data providers verify the data they submitted.

- **Disclosure of Findings:** Did the formal parties to the evaluation ensure that the full set of evaluation findings along with pertinent limitations are made accessible to the persons affected by the evaluation and any others with expressed legal rights to receive the results ("right-to-know" audience)?

  - Define the right-to-know audiences;
  - Report all findings in writing;
  - Report relevant points of view of both supporters and critics of the program;
  - Report balanced, informed conclusions and recommendations;
  - Show the basis for the conclusions and recommendations;
  - Disclose the evaluation's limitations;
  - In reporting, adhere strictly to a code of directness, openness, and completeness;
  - Assure that reports reach the appropriate audiences (intended users, "right-to-know" audiences, others).

- **Analysis of Quantitative Information**: Was quantitative information appropriately and systematically analyzed so that evaluation questions are effectively answered?

  - Begin by conducting preliminary exploratory analyses to assure the data correctness and to gain a greater understanding of the data;
  - Choose procedures appropriate for the evaluation questions and nature of the data;
  - For each procedure specify how its key assumptions are being met;
  - Report limitations of each analytic procedure, including failure to meet assumptions;
  - Employ multiple analytic procedures to check on consistency and replicability of findings;
  - Examine variability as well as central tendencies;
  - Identify and examine outliers and verify their correctness;
  - Identify and analyze statistical interactions;
  - Assess statistical significance and practical significance;
  - Use visual displays to clarify the presentation and interpretation of statistical results.

- **Analysis of Qualitative Information**: Was qualitative information appropriately and systematically analyzed so that evaluation questions are effectively answered?

- Focus on key questions;
- Define the boundaries of information to be used;
- Obtain information keyed to the important evaluation questions;

- Verify the accuracy of findings by obtaining confirmatory evidence from multiple sources, including stakeholders;
- Choose analytic procedures and methods of summarization that are appropriate to the evaluation questions and employed qualitative information;
- Derive a set of categories that is sufficient to document, illuminate, and respond to the evaluation questions;
- Test the derived categories for reliability and validity;
- Classify the obtained information into the validated analysis categories;
- Derive conclusions and recommendations and demonstrate their meaningfulness;
- Report limitations of the referenced information, analyses, and inferences.

- **Justified Conclusions**: Were the conclusions explicitly justified, so that stakeholders can assess them?

- Focus conclusions directly on the evaluation questions;
- Accurately reflect the evaluation procedures and findings;
- Limit conclusions to the applicable time periods, contexts, purposes, and activities;
- Cite the information that supports each conclusion;
- Identify and report the program's side effects;
- Report plausible alternative explanations of the findings;
- Explain why rival explanations were rejected;
- Warn against making common misinterpretations;
- Obtain and address the results of a review by intended users of the draft evaluation report;
- Report the evaluation's limitations.

- **Impartial Reporting**: Did reporting procedures guard against distortion caused by personal feelings and biases of any party to the evaluation, so that evaluation reports fairly reflect evaluation findings?

  - Engage the primary intended user to determine steps to ensure fair, impartial reports;
  - Establish appropriate editorial authority;
  - In addition to the primary intended users, determine the audiences that have a "right-to-know" about the evaluation findings;
  - Establish and follow appropriate plans for releasing findings to all right-to-know audiences;
  - Safeguard reports from deliberate or inadvertent distortions;
  - Report perspectives of all stakeholder groups;
  - Report alternative plausible conclusions;
  - Describe steps taken to control bias;
  - Participate in public presentations of the findings to help guard against and correct distortions by other interested parties.

- **Meta evaluation**: Was the evaluation itself evaluated, so that its conduct is appropriately guided and, on completion, stakeholders can closely examine its strengths and weaknesses?

- Designate or define the standards to be used in judging the evaluation;
- Assign someone responsibility for documenting and assessing the evaluation process and products;
- Budget appropriately and sufficiently for conducting the metaevaluation;
- As feasible, contract for an independent metaevaluation;
- Evaluate the instrumentation, data collection, data handling, coding, analysis and communication of findings to stakeholders against the relevant Evaluation Standard.

6.0 Appendices

Appendix A: Performance Measurement Plan

A Performance Measurement Plan guides performance reporting or trend analysis to support decision-making, accountability and transparency.

A Performance Measurement Plan helps to ensure that **performance indicator** information has been collected systematically and routinely. A performance indicator is composed of a number that provide the magnitude (how much) and a unit that gives the number its meaning (what).

There may be a **target** set for performance indicators. A target is a measurable performance or success level that an organization, program or initiative plans to achieve within a specified time period. For instance, if an immunization campaign aims to vaccinate 90% of a target population then "90% of the target population" is a performance target.

A performance indicator may be assessed against a **benchmark**, which refers to the performance that has been achieved in the recent past by other comparable organizations, or what can be reasonably inferred to have been achieved in the circumstances. For instance, if laboratory services sets that the "standard" for receiving results is 3 days, then "3 days" is the benchmark for wait time.

A performance indicator may also be compared to a **baseline** (indicator) that describes the situation prior to an intervention, against which progress can be assessed or comparisons made.

The Performance Measurement Plan describes:

- Indicators that describe activities, outputs and outcomes of the program;
- Baselines and targets for indicators (if applicable);
- Data collection method and source of information;
- Methods for analysis and reporting on the information (if applicable);
- Individuals responsible to collect and report on the performance indicators.

Implement a Performance Measurement Plan at the start of the program to ensure:

- Baseline data is collected;
- Systems and processes are established to support data collection and reporting.

Performance indicators should be selected judiciously. Ask these questions:

- What are the expected outcomes of the program (goals)?
- What indicators are necessary to meet reporting commitments for the program?
- What indicators are necessary to support organization goals or strategic plans?
- Which indicators are meaningful to stakeholders or program management?
- What indicators will contribute to answering evaluation questions?

The performance indicators in the Performance Measurement Plan should reflect a 'balance' among program components (inputs, activities, outputs and outcomes).

❖ **Additional Reference**

The (UK) National Health Service offers five criteria to assess the quality of a performance indicator:

- Relevance: Does the indicator address a sufficiently important issue?
- Validity: Does this indicator actually measure what it is claiming to measure?
- Possibility: Is it possible to populate the indicator with reliable data?
- Meaning: What would the results indicate? What results are important?
- Implications: What are the implications of a change in the indicator?

"The Good Indicators Guide: Understanding How to Use and Choose Indicators" is available at: http://www.apho.org.uk/resource/item.aspx?RID=44584.

Finally, the resources and costs of establishing and/or maintaining a performance measurement system have to be considered. These resources could be employed in other ways (e.g. program delivery) and so the perceived benefit of reports on outcome indicators should outweigh the cost of such a system.

- What current systems (e.g. information or operational systems) are in place to support data collection and reporting?
- What are the estimated costs for performance measurement activities (e.g. human resources, software licensing etc.) including provisions to ensure data integrity?
- Who is responsible for data collection and reporting?
- What frequency of data collection and reporting for each performance indicators?
- How, when and by whom will the Performance Measurement Plan be reviewed and adjustments made?

❖ **Additional References**

"*Splash Ripple* Using Outcomes to Design & Manage Community Activities Plan" describes how to identify and collect information necessary to show change as a result of the program. This document is available at: http://www.smartfund.ca/docs/smart_outcomes_guide.pdf

❖ **An Example in Health Care**

Health Canada's Strategic Training Initiative in Health includes a "Performance Measurement and Evaluation Framework" (see http://www.cihr-irsc.gc.ca/e/27737.html).

Below is a modified excerpt from the performance measurement plan for this initiative, showing an example for one output and one short-term outcome.

Program Element	Performance Indicators	Data Sources	Responsibility for data collection / reporting	Frequency of Measurement
Output: High quality trainees recruited	Mean undergraduate standing of trainees	Progress reports	Grantees	Annual
Short-term outcome: Increased number of students in health research training	Numbers of students in targeted research areas	Survey of Institutes and partners	CIHR	At three and six years

Appendix B: Sample Evaluation Charter

In this example, the Evaluation Charter describes the creation of an Evaluation Plan.

1. Introduction

E.g. The purpose of this 'Evaluation Charter' is:

1) to outline our preliminary understanding of the requirements to complete the project.
2) to raise key questions that require a response before embarking on the project.

2. Purpose of the Project

E.g. The purpose of this project is to undertake an evaluation planning process for the program. The evaluation planning process will involve two main activities:

1) Regular consultation with staff (managers, supervisors, coordinators)
   - 6 monthly meetings (3 hours each)
   - 6 monthly teleconferences (1 hour each)

2) Creation of an evaluation plan, which will include:
   - Program description and logic model
   - Evaluation questions and proposed methodology
   - Recommendations and next steps

3. Constraints

E.g. The scope and purpose of this project is constrained by the fact that the plan must be complete before XXXX.

4. Intended Use, Users and Audience of the Evaluation Plan

E.g. The Evaluation plan will be useful to program staff (Executive Director, Directors, Managers, Supervisors and NSS Coordinators).

5. Working Group Roles and Responsibilities

Name	Role in Working Group
Program staff	E.g. Participate in meetings and teleconferences - Offer guidance and expertise - Provide feedback on documents - Identify sources of information - Preparation for each meeting may require up to 2 hours.
Evaluator (internal to organization)	E.g. Facilitate discussion to complete Evaluation Plan - Offer related evaluation expertise - Coordinate document review/ feedback - Preparation for each meeting may require up to 6 hours.

6. Work Plan

Describe major milestones, meetings, goals and dates.

7. Project Sign-off

Include signatures to indicate understanding of the materials contained in the document and agreement to the goals, deliverables, and responsibilities described.

Appendix C: Examples of Evaluation Questions

The following list of evaluation questions are intended to give examples of language and the types of questions that may be asked. There is overlap among several questions as they have been compiled from multiple sources. [45] [46] [47]

- **Implementation:** Were the program activities carried out as originally intended? Is the program being delivered as it was designed?

- **Relevance:** Are the program goals addressing the needs that motivated the creation of the program? Does the program continue to be consistent with organizational priorities?

- **Success**: Is the program achieving its intended outcomes, within budget and without unwanted outcomes?

- **Adequacy:** Were the outcomes sufficient to meet the needs of the target population?

- **Effectiveness:** Effectiveness describes the extent to which the program outcomes were achieved. Effectiveness is a 2 stage question:
  - **Attribution:** Can the observed outcomes be attributed to the program? (includes incremental effect, unintended effects)
  - **Compliance and Accountability:** Is the program achieving the goals it was intended to accomplish? Did the activities take place as planned?

- **Efficiency:** Efficiency describes how resources/inputs are converted to results. There are three types of efficiency (technical, productive and allocative) and all apply to healthcare. [48] Is the maximum possible improvement obtained from a set of resource inputs (technical efficiency)?

- **Affordability:** Is the program affordable? If not, could the program stop?

- **Role of Organization:** Is there a legitimate and necessary role for this organization (e.g. government) in delivering this product or service?

- **Continuous Improvement**: Are there ways to improve program delivery from an effectiveness, efficiency or affordability perspective?

- **Learning:** Has the program established best practices? Are there lessons learned from implementation or analysis of success of the program?

---

[45] Kusek & Risk (2004)
[46] U.S. Centers for Disease Control and Prevention
[47] Treasury Board of Canada Secretariat "Results-Based Management Lexicon"
[48] Palmer, S., & Torgerson, D. J. (1999), http://www.bmj.com/cgi/content/full/318/7191/1136

Appendix D: Evaluation (Research) Designs and Examples

Experimental Design

Experimental design is used to definitively establish the link between the program and observed outcomes (cause and effect). In experimental design, often described as a "randomized control trial" (RCT), people are randomly assigned to a treatment group (program recipients) or control group (people that do not receive the program services) and then outcomes are compared for the two groups.

If the treatment and control groups have been randomly assigned, it is assumed that the only difference between the two groups is the treatment (that the "pre" stage is the same for the whole group), and therefore "after-only" experimental design is sufficient.

Randomization must be designed at early stages of the program to best ensure the difference between the program and control group were due to confounding factors, bias or chance. The same random assignment must be maintained through the period of experimentation. In this situation, RCTs have the highest internal and external validity and therefore are the ideal research designs ('gold standard') for testing causal relationships.

In clinical health care settings (e.g. Drug trials), experimental design is essential to determine the efficacy of the causal chain between the program recipient and the outcome and where results may be safely extrapolated to other settings.

Challenges to using Experimental Design

Random assignment may not be feasible or ethical. In reality, random assignment (of people to policies or policies to people) may not be possible. For instance, evaluation tobacco control policies could not use an RCT.

- It can be costly to identify, track and collect data from members of a control group. As a result, a randomized experiment may not be affordable.

- A randomized experiment is best suited for established programs that have established clear and consistent set of activities and warrant the cost involved, after implementation issues have been addressed.

- Random assignment may not be ethical, for instance, if the service is necessary for the entire target population.

- The pool of potential participants may be too small to fill both a treatment and a control group (e.g. in a small community).

Quasi-experimental design

In the case of quasi-experimental designs, treatment group (program recipients) and comparison group (those who didn't receive the program) are not randomly assigned. As a result, quasi-experimental design can inform discussions of cause-and-effect, but cannot prove that a program causes a change in outcomes.

Quasi-experimental design is the "next best" option when experimental design is not practical or feasible. Quasi-experimental design can yield important information about the program, such as:

- Is there evidence to support the idea of a causal link between the program and observed outcomes?
- Who is being served by the program?
- Is the target population being served?
- Are expected outcomes being achieved?
- Are there unintended outcomes?
- Which sub-groups of the target population are achieving these outcomes?

There are three common types of quasi-experimental design:

- Compare groups of program recipients across similar communities.
- Compare individual program recipients with individuals who have similar characteristics.
- A "pre-test/post-test" or "before-after" experimental design with the individual program recipient as his or her own comparison.

Challenges to using Quasi-experimental Design

The selection and feasibility of selecting a comparison group is one of the main challenges of this design:

- The comparison population may not be similar to the population being served. For example, if the comparison population is more advantaged than the population being served, then outcomes for program participants may seem less positive than they really are.

- Conducting a comparison group study is sometimes both demanding and costly. For example, in the situation that pre-test data and post-test data are collected for the study population and for the comparison population, it may be worth considering an experimental study as such a design (while not necessarily more cost-effective) would allow for conclusions to be drawn about the casual link between the program and observed outcomes.

❖ **An Example in Healthcare**

In March 1999, the publicly-funded drug benefit program in British Columbia (PharmaCare) introduced a drug benefit restriction policy with a randomized delayed control group: 10% of general practices in BC were granted a 6-month optional delay in the policy. The delayed control group was part of the research design to evaluate the new drug policy.[49]

---

[49] Maclure, M., Carleton, B., & Schneeweiss, S. (2007). Prior studies included Carleton, B., & Maclure, M. (2001).

Non-experimental design

The introduction of **non-experimental design** in program evaluations has been the most significant change to the field of evaluation since the 1970s.

Non-experimental (implicit) designs are frequently used in evaluation. It is often the only design that can be used in situations when:

- No pre-program measures exist;
- There is no obvious control group available;
- It is not reasonable to assign interventions on a random basis.

Non-experimental design is characterized by:

- "Naturalistic inquiry" in the 'real-world' rather than in manipulated settings;
- Consideration for the social context in which the program operates;
- Development (inductive analysis) rather than testing of hypotheses (deductive analysis);
- The use of non-standardised, semi-structured or unstructured methods which are sensitive to the social context of the study.

Non-experimental design can be used in both process and impact evaluation for both exploratory and descriptive purposes, including to:

- Examine an issue or problem that is poorly understood;
- Inform the kind of intervention required;
- Identify factors that contribute to successful or unsuccessful delivery;
- Identify outcomes (intended or unintended) and how they occur;
- Examine the nature of requirements of different groups within the target population;
- Explore the contexts in which a program operates;
- Explore organizational aspects of delivery.

Challenges to using Non-experimental Design

There are, however, more challenges in attributing impacts to specific interventions with non-experimental designs as compared to quasi-experimental or experimental designs (the strongest for attributing impacts).

❖ **An Example in Healthcare**

The Victoria AIDS Respite Care Society (VARCS) has provided respite care to people living with HIV/AIDS since 1991. In 1998, an evaluation of VARCS services was undertaken to gain a better understanding of this model of community-based care.[50]

The evaluation was comprised of four components:

- An examination of the historical evolution of VARCS;
- An account of the community development processes used;

[50] Stajduhar, K.I, E. Lindsey, and L. McGuinness. (2002).

- A description of VARCS services;
- An analysis of the impact and outcomes of VARCS services.

The impact and outcomes component of this study was guided by **participatory action research.** In participatory action research, people join in partnership with researchers to design and implement research studies that are relevant and meaningful to them.

Data were collected by individual, in-depth interviews and focus group interviews. Open-ended questions were used to fully explore participants' experiences. Interview data were subjected to **qualitative thematic analysis**, which involved the following steps:

- Listening to audiotape recordings while concurrently reading transcriptions;
- Identifying units of meaning;
- Coding into early themes;
- Identifying patterns and meta-themes;
- Synopsis of the impact and outcomes.

Eight themes emerged from the analysis that appeared to speak to the qualitative impact and outcomes of VARCS service, including the provision of appropriate care (by health professionals and volunteers).

The findings from these eight themes indicated a high level of support for VARCS's model of respite care. The findings lead to recommendations for program changes. For instance, clearly stated guidelines and policies related to the role of volunteers were suggested to improve communication patterns between VARCS staff and volunteers.

The report notes that a sampling bias may have existed as many of the participants held close associations with VARCS or were clients and, therefore, may have felt inhibited to openly discuss any criticisms. Sampling bias was minimized by actively seeking out participants who had been but were no longer associated with the Society.

Appendix E: Evaluability Assessment

Evaluability assessment describes the early review (before undertaking evaluation) to determine whether the requirements for conducting an evaluation can be met. These requirements include adequately defined goals, available resources, timeline, data availability and data verification.

An evaluability assessment is conducted after a "Plan for Program Evaluation" is complete. Assessing the "evaluability" of a program requires managers and staff to show evidence (e.g. data collection and preliminary reports) that the evaluation plan can be implemented.

An evaluability assessment asks the fundamental questions:[51]

- Could an evaluation proceed?
- Are there limitations in the way the plans are implemented?
- Should an evaluation proceed?
- Does the timing of the evaluation meet the needs of the prospective audience for the report?

If an acceptable evaluation cannot be conducted within the constraints of the evaluation plan, then the resources and time frame may be renegotiated, the scope and goals of the evaluation revised, or the evaluation cancelled.

❖ **Examples in Healthcare**

CHSPR (2006) examined the evaluability of the logic model of the **Primary Health Care System in Canada** (as presented in Step 2) to determine:

- How existing data sources could be used to describe PHC;
- Identify gaps in the data sources that hinder PHC reporting;
- Recommend how gaps in data sources could be filled.[52]

Program improvement may be the most significant outcome of an evaluability assessment. An evaluability assessment for the **Calgary Cross-Cultural Mental Health Consultation Project** clarified four barriers to successful evaluation:[53]

- Defining the target clients;
- Gaps in the logic model between activities and outcomes;
- Shifting from a process to outcome orientation;
- Realistic goal development.

---

[51] McDavid & Hawthorn (2006)
[52] Broemeling, A. M., Watson, D. E., Black, C., & Reid, R. J. (2006)
[53] Thurston, W. E., Graham, J., & Hatfield, J. (2003)

7.0 References

7.1 References for Evaluation Terminology

The Evaluation Centre at Western Michigan University:
http://ec.wmich.edu/Glossary/glossaryList.htm

OECD "Glossary of key terms in evaluation and results based management" (2002).
http://www.oecd.org/dataoecd/29/21/2754804.pdf

Treasury Board of Canada Secretariat "Results-Based Management Lexicon",
http://www.tbs-sct.gc.ca/rma/lex-lex_e.asp

7.2 References Used in the Guide

Bamberger, M. (2006). Conducting Quality Impact Evaluations Under Budget, Time and Data Constraints. Independent Evaluation Group, World Bank, Washington DC. Retrieved from: http://www.worldbank.org/ieg/ecd/conduct_qual_impact_eval.html

Broemeling, A. M., Watson, D. E., Black, C., & Reid, R. J. (2006). Measuring the Performance of Primary Health Care. Retrieved from: http://www.chspr.ubc.ca/files/publications/2006/chspr06-21S.pdf

Canadian Nurse Practitioner Initiative. (2006). Implementation and Evaluation Toolkit for Nurse Practitioners in Canada. Retrieved from: http://www.cnpi.ca/documents/pdf/Toolkit_Implementation_Evaluation_NP_e.pdf

Carleton, B., & Maclure, M. (2001) Randomized Controlled Trial of Pharmacare's Nebulizer to Inhaler Conversion Policy. Canadian Health Services Research Funding Final Report. Retrieved from: www.chsrf.ca/final_research/ogc/pdf/carleton_e.pdf

Centre for Health Services and Policy Research. (2003). What a Tangled Web We Weave: Improving Performance Reporting and Accountability in BC. *Proceedings of the 16th Annual Health Policy Conference of the Centre for Health Services and Policy Research.* Retrieved from: http://www.chspr.ubc.ca/hpc/archive

Coyne, K. and P. Cox. (2004). "Splash Ripple" Using Outcomes to Design & Manage Community Activities. Retrieved from: http://www.smartfund.ca/docs/smart_outcomes_guide.pdf

Fetterman, D. M., Kaftarian, S. J., & Wandersman, A. (1996). Empowerment Evaluation: Knowledge and Tools for Self-assessment & Accountability. Sage Publications Inc.

Grembowski, D. (2001). The Practice of Health Program Evaluation. Sage Publications.

Issel, L. M. (2004). Health Program Planning and Evaluation: A Practical, Systematic Approach For Community Health. Jones and Bartlett Publishers.

Joint Committee on Standards for Educational Evaluation. (1994). The Program Evaluation Standards, 2nd ed. Thousand Oaks, CA: Sage

Kusek, J. Z., & Rist, R. C. (2004). Ten Steps to a Results-based Monitoring and Evaluation System: A Handbook for Development Practitioners. World Bank Publications. Retrieved from: http://www.preval.org/documentos/00804.pdf

Linden, A., W. Trochim, J. Adams. (2006). Evaluating Program Effectiveness Using the Regression Point Displacement Design. *Evaluation and the Health Professions, 29* (4), 407-423.

Love, A. J. (1991). Internal Evaluation: Building Organizations From Within. Sage Publications.

Maclure, M., Carleton, B., & Schneeweiss, S. (2007). Designed Delays Versus Rigorous Pragmatic Trials: Lower Carat Gold Standards Can Produce Relevant Drug Evaluations. *Medical Care*, 45(10), S44.

McNamara, C. (2008). Basic Guide to Program Evaluation. Retrieved from: http://www.managementhelp.org/evaluatn/fnl_eval.htm#anchor1665834

Mainz, J. (2003). Defining and Classifying Clinical Indicators for Quality Improvement. *International Journal for Quality in Health Care*, 15(6), 523.

McDavid, J. C., & Hawthorn, L. R. L. (2006). Program Evaluation & Performance Measurement: An Introduction to Practice. Sage Publications Inc.

National Health Service. The Good Indicators Guide: Understanding How To Use and Choose Indicators. Retrieved from: http://www.apho.org.uk/resource/item.aspx?RID=44584

Palmer, S., & Torgerson, D. J. (1999). Definitions of efficiency. *British Medical Journal, 318*(7191), 1136-1136. Retrieved from: http://www.bmj.com/cgi/content/full/318/7191/1136

Patton, M. Q. (2008). Utilization-focused Evaluation. Los Angeles, CA: Sage Publications.

Research Councils UK. (2005). Evaluation: Practical Guidelines. Retrieved from: http://www.rcuk.ac.uk/aboutrcuk/publications/corporate/evaluationguide.htm

Rossi, P. H., Lipsey, M. W., & Freeman, H. E. (2004). Evaluation: A Systematic Approach. Thousand Oaks, CA: Sage

Rutman, L. (1977). Evaluation Research Methods: A Basic Guide. Beverly Hills, Calif.: Sage Publications.

Spencer, L., J. Ritchie, J. Lewis and L. Dillon. (2003). Quality in Qualitative Evaluation: A Framework for Assessing Research Evidence. National Centre for Social Research. UK: Government Chief Social Researchers Office. Retrieved from: http://www.cabinetoffice.gov.uk/strategy/publications.aspx

Stajduhar, K. I., Lindsey, E., & McGuinness, L. (2002). A Qualitative Evaluation of an HIV/AIDS Respite Care Service in Victoria, Canada. *Evaluation & the Health Professions*, 25(3), 321.

Stufflebeam, D. L. (1999). Evaluation Contract Checklist. Retrieved from http://www.wmich.edu/evalctr/checklists/checklistmenu.htm

Stufflebeam, D. L. (1999). Program Evaluations Metaevaluation Checklist. Retrieved from http://www.wmich.edu/evalctr/checklists/checklistmenu.htm

Thurston, W. E., Graham, J., & Hatfield, J. (2003). Evaluability Assessment. A Catalyst for Program Change and Improvement. *Evaluation & the Health Professions*, 26(2), 206-221.

Treasury Board of Canada Secretariat. (2004). Guide for the Review of Evaluation Reports. Retrieved from: http://www.tbs-sct.gc.ca/eval/tools_outils/4001752_e.asp

Treasury Board of Canada Secretariat. (2007). Performance Reporting Good Practices Handbook. Retrieved from: http://www.tbs-sct.gc.ca/rma/dpr3/06-07/handbk-guide/

Van Marris, B., & King, B. (2006). Evaluating Health Promotion Programs. Health Communication Unit, Centre for Health Promotion, University of Toronto. Retrieved from: http://www.thcu.ca/infoandresources/evaluation.htm

Volkov, B.B. and J.A. King. (2007). A Checklist for Building Organizational Evaluation Capacity. Retrieved from: http://www.wmich.edu/evalctr/checklists/ecb.pdf

Watson, D., Broemeling, A., Reid, R., Black, C. (2004). A Results-based Logic Model for Primary Health Care. Vancouver: Centre for Health Services and Policy Research, University of British Columbia. Retrieved from: http://www.chspr.ubc.ca/research/phc/logicmodel

W.K.K. Foundation. (2004). Logic Model Development Guide. Battle Creek (MI): WK Kellogg Foundation. Retrieved from: http://www.wkkf.org/

W.K.K. Foundation. (1998). WK Kellogg Foundation Evaluation Handbook. Battle Creek (MI): WK Kellogg Foundation. Retrieved from: http://www.wkkf.org/

7.3 Additional Evaluation References

Adair, C. E., Simpson, E., Casebeer, A. L., Birdsell, J. M., Hayden, K. A., and Lewis, S. (2006). Performance Measurement in Healthcare: Part II-State of the Science Findings by Stage of the Performance Measurement Process. *Healthcare Policy*, 2(56-79).

Aucoin, P. (2005). Decision-Making in Government: The Role of Program Evaluation. Treasury Board Secretariat, Government of Canada. Retrieved from: http://www.tbs-sct.gc.ca/eval/tools-outils_e.asp

Bevan, G. (2006). Setting Targets for Health Care Performance: Lessons From a Case Study of the English NHS. *National Institute Economic Review*, 197(1), 67.

Blalock, B. (1999). Evaluation Research and the Performance Management Movement: From Estrangement to Useful Integration? *Evaluation*, 5(2), 117.

Bowen, S. and S.A. Kreindler. (2008). Indicator Madness: A Cautionary Reflection on the Use of Indicators in Healthcare. *Healthcare Policy, 3*(4), 41-48.

British Columbia. Ministry of Health. (1996). The Evaluation Framework Reference Guide. Victoria: Author.

British Columbia. Ministry of Health. (1996). Health Service Outcome Reporting: A Guide for Health Agencies. Victoria: Author.

British Columbia. Office of the Auditor General. (2003). Performance Reporting Principles for the British Columbia Public Sector. Retrieved from: http://www.bcauditor.com/

British Columbia. Office of the Auditor General. (2007). Follow-up of 2004/05 Report 3: Preventing and Managing Diabetes in British Columbia. Retrieved from: http://www.bcauditor.com/

British Columbia. Office of the Auditor General. (2008). Strengthening Accountability in British Columbia: Trends and Opportunities in Performance Reporting. Retrieved from: http://www.bcauditor.com/

Brown, A.D. and Veillard, J. (2008). Commentary: Indicators with a Purpose-Meaningful Performance Measurement and the Importance of Strategy. *Healthcare Policy, 3*(4), 49-52.

Bryant-Lukosius, D., & DiCenso, A. (2004). A Framework for the Introduction and Evaluation of Advanced Practice Nursing Roles. *Journal of Advanced Nursing*, 48(5), 530-540.

Calgary Health Region. (2005). Regional Evaluation Guide. Calgary: AB: Calgary Health Region. http://www.calgaryhealthregion.ca/programs/research/pdf/evaluation_guide.pdf

Canadian Institute for Health Information. (2008). Health Indicators 2008. Retrieved from: http://secure.cihi.ca/cihiweb/products/HealthIndicators2008_ENGweb.pdf

Carleton, B. and M. Maclure. (2001). Randomized Controlled Trial of Pharmacare's Nebulizer to Inhaler Conversion Policy. Canadian Health Services Research Funding Final Report. Retrieved from www.chsrf.ca/final_research/ogc/pdf/carleton_e.pdf

Clark, M., & Sartorius, R. (2004). Monitoring and Evaluation: Some Tools, Methods, and Approaches. Washington, DC: The World Bank. Retrieved from: http://www.worldbank.org/ieg/ecd/me_tools_and_approaches.html

Fredericks, K.A., M. Deegan and J.G. Carman. (2008). Using System Dynamics as an Evaluation Tool: Experience from a Demonstration Project. *American Journal of Evaluation*, 251(29)

Gauthier, B., S. Borys, N. Kishchuk, and S. Roy. (2006). Evaluation Practice in Canada: Results of a National Survey. *Canadian Journal of Program Evaluation*, 21(3), 1-42.

Godfrey, M. M., Nelson, E. C., Wasson, J. H., Mohr, J. J., & Batalden, P. B. (2003). Microsystems in Health care: Part 3. Planning Patient-centered Services. *Joint Commission Journal on Quality and Safety*, 29(4), 159-170.

Green, C.J. and Moehr, J.R. (2000). Performance Evaluation Frameworks for Vertically Integrated Health Care Systems: Shifting Paradigms in Canada. Proc AMIA Symp, 315, 9. Retrieved from: http://www.amia.org/pubs/symposia/D200651.PDF

International Fund for Agricultural Development. (2003). A Methodological Framework for Project Evaluation. EC 2003/34/W.P.3

Jarvis, H. (2000). Differentiating Among Research, Evaluation and Measures to Assure Quality. *Healthcare Management Forum*, 13(4), 34-36.

Kettl, D. (2000). Performance Management: The State of the Field. Association for Public Policy and Management, 22nd Annual Research Conference, Seattle, Washington.

Kring, D. L. (2008). Research and Quality Improvement: Different Processes, Different Evidence. MEDSURG Nursing—June 2008—Vol. 17/No. 3, 17(3), 162-169.

Levy, A. K. (2005). Categorizing Outcomes of Health Care Delivery. *Clinical and Investigative Medicine.* 28(6), 347-350.

Love, A. (1993). Internal Evaluation: An Essential Tool for Human Service Organizations. *The Canadian Journal of Program Evaluation*, 8(2), 1-15.

Mark, B. A., Salyer, J., & Geddes, N. (1997). Outcomes Research. Clues to Quality and Organizational Effectiveness? *The Nursing Clinics of North America*, 32(3), 589-601.

Mathison, S. (2005). Encyclopedia of Evaluation. Sage Publications.

Mayne, J. (2001). Addressing Attribution Through Contribution Analysis: Using Performance Measures Sensibly. *The Canadian Journal of Program Evaluation,* 16(1), 1-24.

Mayne, J. (2004). Reporting on outcomes: Setting performance expectations and telling performance stories. *Canadian Journal of Program Evaluation*, 19(1), 31-60.

Mayne, J. (2008). Addressing Cause and Effect in Simple and Complex Settings Through Contribution Analysis. In R. Schwartz, K. Forss, and M. Marra (Eds.) (Ed.), Evaluating the Complex. Transaction Publishers (forthcoming).

Mayne, J., & Zapico-Goñi, E. (1997). Monitoring Performance in the Public Sector: Future Directions from International Experience. New Brunswick, NJ: Transaction Publishers.

McDavid, J. (2001). Program Evaluation in British Columbia in a Time of Transition: 1995-2000. *Canadian Journal of Program Evaluation*, 3-28.

McEwan, K. L., & Bigelow, D. A. (1997). Using a Logic Model to Focus Health Services on Population Health Goals. *Canadian Journal of Program Evaluation*, 12(1), 167-174.

Norman, R. (2004). Recovering From a Tidal Wave: New Directions for Performance Management in New Zealand's Public Sector. *Public Finance and Management*, 4(3), 429-447.

Popp, J., L Heureux, L., Dolinski, C., Adair, C., Tough, S., Casebeer, A., et al. (2005). How do you Evaluate a Network? A Canadian Child and Youth Health Network Experience. *Canadian Journal of Program Evaluation*, 20(3), 123.

Porteous, N. L., B. J. Sheldrick, P.J. Stewart. (1997). Program evaluation toolkit: A blueprint for public health management. Public Health Research, Education and Development Program, Ottawa-Carleton Health Department.

Thurston, W. E., & Potvin, L. (2003). Evaluability Assessment: A Tool for Incorporating Evaluation in Social Change Programmes. *Evaluation*, 9(4), 453.

U.S. Department of Health and Human Services. Centers for Disease Control and Prevention. Office of the Director, Office of Strategy and Innovation. (2005). Introduction to Program Evaluation for Public Health Programs: A Self-Study Guide. Atlanta, GA: Centers for Disease Control and Prevention.

Victora, C. G., Habicht, J. P., & Bryce, J. (2004). Evidence-based Public Health: Moving Beyond Randomized Trials. *American Journal of Public Health*, 94(3), 400-405.

Weiss, H. (1999). The Interface Between Evaluation and Public Policy. *Evaluation*, 5(4), 468.

Zelman, W. N., Pink, G. H., & Matthias, C. B. (2003). Use of the Balanced Scorecard in Health Care. *Journal of Health Care Finance*, 29(4), 1-16.