

## 4 Program Evaluation

CHRISTEL A. WOODWARD

*McMaster University*

### SUMMARY

*Program evaluation provides a set of procedures and tools that can be employed to provide useful information about medical school programs and their components to decision-makers. This chapter begins with contemporary views of program evaluation. It describes the various proposals of evaluation and the areas of decision-making that they can inform. Logic models are introduced as a tool to facilitate evaluation activities. The importance of including stakeholders in planning an evaluation and in clarifying the evaluation questions and goals is stressed. Evaluation requires careful planning, as this activity is always somewhat political in nature. Inadequate attention to planning may lead to questionable information and to poor uptake of results. A series of basic issues that must be considered in developing an evaluation strategy for medical education programs is described and illustrated with examples from recent program evaluation activities in undergraduate medical education. Recommendations are made about how we can continue to improve evaluation activities including some guidelines for those who wish to evaluate medical education programs.*

### INTRODUCTION

The professional education of physicians, as all educational programs, may be viewed as a social program. Society has a vital interest in how physicians are educated and wishes to ensure that medical schools produce graduates who will be able to fulfill societal needs (Hagenfeldt & Lowry, 1997). In most jurisdictions, the cost of medical education receives a significant subsidy by the larger society through direct and indirect grants from governments and other institutions and preferential tax treatment of educational institutions and hospitals. During the past 35 years, significant advances have occurred in our understanding of the tactics and methodological tools important to developing useful program evaluations. When

used in the evaluation of medical education, these strategies can help provide answers conducive to the improvement and continued refinement of physician education programs.

This chapter discusses some of these advances and examines their applicability to the evaluation of the education of physicians. It assumes that the reader is already familiar with basic research design and measurement issues that have been presented in previous chapters in this section. The approaches and tactics used in planning an evaluation are the focus of the discussion. Strategies that increase the likelihood that evaluation activities influence decision-making are highlighted.

Medical education is divided into three broad periods. Medical school, post-graduate or residency training and continuing education. Often each of these types of education is studied separately and considered a program. Yet, they may also be viewed as an educational continuum that has one over-arching program. They share some of the same broad implicit goals (to produce physicians who deliver high quality care in their chosen field). Because the administration and funding of education in these time periods are often different, and their specific objectives are different, they are more readily treated as separate programs. Given this complexity, for the remainder of this chapter, examples chosen to illustrate program evaluation issues in medical education will refer to the medical school period unless another time period is specifically mentioned.

## UNDERSTANDING PROGRAM EVALUATION

### *Contemporary views regarding program evaluation*

Program Evaluation as a discipline has matured from its early days in the 1960s. At that point, it could barely be distinguished from evaluation research: the use of standard quantitative research methods, preferably a randomized controlled trial, to answer questions about a program.

Not only does it now embrace all kinds of research methods but it also recognizes that good program evaluation produces information that is useful to and used by decision-makers (Patton, 1997; Weiss, 1998). Simply to generate information about a program is not good enough. The processes and procedures used should ensure that the information is also used to guide and improve programs. This has led to the development of approaches to program evaluation which are more inclusive and involve key players (called stakeholders in program evaluation language) from its very inception. Table 1 provides a summary of contemporary views regarding program evaluation.

The range of questions that can be asked about a program and be addressed in program evaluation is broad and may assist in program planning, implementation, improvement as well as providing information on program effects and impacts. Different stakeholders may have different perspectives on a program and have quite different concerns about it. They may not agree on what the most important

outcomes of a program are or see the need for specific program components. Some stakeholders may be as concerned about the unintended effects of a program or changes in it as they are in its intended effects. Unanimous agreement about which questions are most important is rarely seen; addressing only the concerns of one group will make it difficult to convince everyone that a fair and balanced approach to evaluation has been taken.

**Table 1. Contemporary views regarding program evaluation**

1.	Many different research strategies may be employed
2.	The information gained should be useful to decision-makers
3.	Stakeholder involvement increases the likelihood that findings will be used
4.	Questions should inform program planning and activities
5.	Programs are not static
6.	Evaluation is important to program accountability
7.	Evaluation should assist in understanding how a program produces the effects seen

There is also a recognition that programs are not static and the developmental phase of the program should also influence the types of questions that are posed. It is senseless to ask about program outcomes at a point when implementation issues are still unresolved and the program is still evolving. Generating information about the need for a program (or the need for changes in it) is also an important activity that may promote program development (or change). This is the first stage in Figure 1.

Needs assessment can lead to a phase of planning for change (or a new program). As change is implemented, evaluation activities can address how well implementation is proceeding. Evaluation is also useful to monitor program performance and assess outcomes. These activities may identify new needs which begin the cycle of activities again.

The image of program evaluation has changed from something dreaded and imposed to something important to program development and accountability. To some extent, this change has occurred because stakeholders have been given a greater role in decision-making about the questions that are important to address. Also it is rare that a program has been scrapped based on its evaluation (unless it was a pilot project). Rather, the information gained has been used to make changes in programs. Such incremental changes are much more common than the disappearance of a program.

Finally, and perhaps most important, program evaluators have become more concerned with understanding how a program works and why it produces the effects seen. They have sought to make explicit assumptions (the theory or theories) underlying a program as some of our assumptions, when examined carefully, are naive at best, and often find no support in the empirical literature. Often these assumptions are buried in program criteria or activities and, when explicated, we quickly see that they could not possibly be expected to have the outcomes we say the program intends.

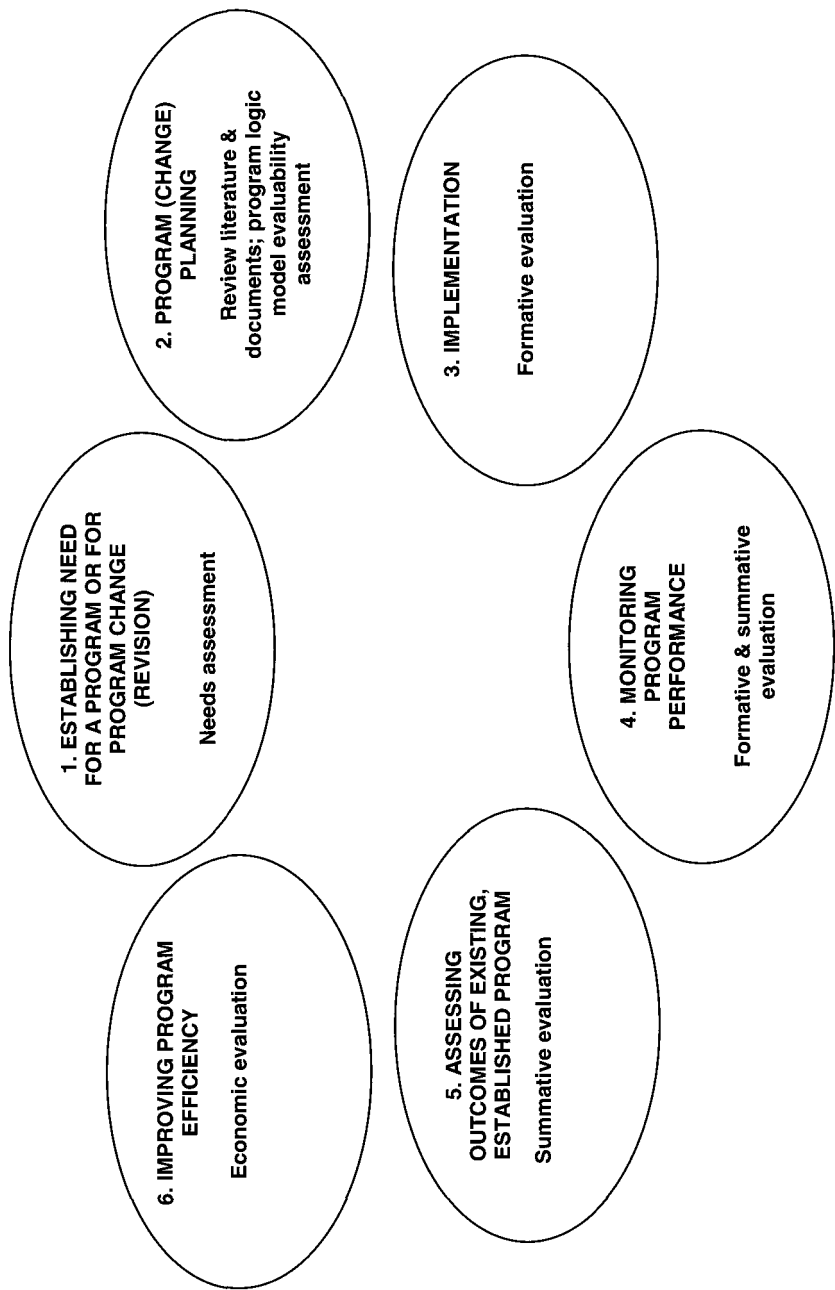


Figure 1. Stages in program development and associated program evaluation activities

### *Types of program evaluation and their purposes*

There are numerous reasons for doing program evaluation (see Figure 1) and they differ considerably in the kinds of information that is of interest, how the evaluation is structured and who are the most important stakeholders and audiences.

First, we can evaluate the need for a program, called needs assessment. Given that medical education programs have existed for a long time, when thinking about medical education, we can evaluate the need for a change in the program or whether it should be expanded or contracted. Establishing need is sometimes important in “selling” any changes that we wish to make and garnering support for those changes. For example, when the Robert Wood Foundation was interested in promoting changes in medical education in the United States (US), it commissioned a survey by Louis Harris and Associates of medical schools deans, department chairs and faculty members which helped establish “need” (Cantor, Cohen, Barker, Shuster, & Reynolds, 1991). The findings suggested that a clear majority of these medical educators thought that US medical education needed fundamental changes and slightly over half agreed that this was also true of medical education at their own institutions. Many believed that medical education had not kept pace with the way that medicine was currently practiced. Without such evidence, it is unlikely that the Foundation would have invested in a program of funding “reform” which was to be led by medical educators. The most important audience for this information and a key stakeholder group were medical educators themselves, so they chose to survey only this group. Information about the need for change from the general public would likely have been less compelling and not provided enough impetus to sustain change.

Evaluation also plays an important role in program planning and implementation. Often an evaluator will assist program planner and managers by developing a logic model. Logic models are a simplified, diagrammatic description of the program including its goals, the component activities needed to accomplish them, their outputs (countable by products of each component), short and longer term outcomes (direct results or accomplishments) and impacts (effects for which the program can claim only partial responsibility). Figure 2 shows a diagram of logic model components. Usually, a program logic model can be drawn on one page and provides an overview of the program. Once a logic model has been developed, it can be used to examine the assumptions about how the program (or changes in it) is to produce the effects intended. Such an exercise is useful in the planning stage as it may point out that no activities have been planned that would allow the program to have some of the effects intended. Conversely, it may point out that activities are planned that are not linked to program objectives. It can also help identify all the tasks that need to be accomplished to implement the program.

A formal evaluability assessment may accompany program planning or be done in planning for program evaluation at any later stage when evaluation is planned. An evaluability assessment develops the terms of reference for the evaluation and

<http://www.springer.com/978-1-4020-0466-7>

International Handbook of Research in Medical  
Education

Norman, G.R.; van der Vleuten, C.P.M.; Newble, D.I.  
(Eds.)

2002, XIII, 1106 p. In 2 volumes, not available  
separately., Hardcover

ISBN: 978-1-4020-0466-7