The Canadian Journal of Program Evaluation Vol. 17 No. 3 Pages 39–56
ISSN 0834-1516

THE PITFALLS AND THE POTENTIAL OF EARLY EVALUATION EFFORTS: LESSONS LEARNED FROM THE HEALTH SERVICES SECTOR

Karen L. Lawson
Heather D. Hadjistavropoulos
Regina Health District
Regina, Saskatchewan

Abstract: Evaluators often find themselves assuming a variety of roles as they examine programs and interact with the people connected to those programs. The present article proposes that this is especially true when attempting to conduct an impact evaluation very quickly after a new program is initiated. Given the increasing trends toward program accountability, administrators will often undertake evaluations very quickly after new programs are initiated, and evaluators are increasingly asked to determine the impact of a program that is not yet fully functioning. Using examples drawn from the experience of conducting an outcome evaluation of a major reorganization of a health service delivery system very soon after the changes were implemented, the unique challenges and benefits of evaluating a complex program in the early phases following implementation will be highlighted. Specifically, the varied roles that the evaluators were required to assume and the lessons that they learned from expanding their professional boundaries will be outlined. In addition to the diverse roles that evaluators often occupy (such as educator, consultant, and researcher), those conducting early impact evaluations may find themselves acting as protocol trainers, mediators, and/or therapists for program staff and administration as they attempt to evaluate the outcome of a program that has not been fully implemented.

Résumé: Les évaluateurs doivent souvent assumer une variété de rôles lorsqu'ils examinent des programmes et rencontrent les gens qui en sont reliés. Cet article propose que ceci est particulière-

Corresponding author: Heather Hadjistavropoulos, Associate Director, Clinical Research and Development Program, Regina Health District, 2180 - 23 Avenue, Regina, Saskatchewan S4S 0A5; <hhajistavropoulos@reginahealth.sk.ca>

ment approprié lorsqu'on essaie d'effectuer une évaluation d'impact peu après qu'un nouveau programme soit lancé. Etant donné les tendances de plus en plus importantes vers l'évaluation de l'imputabilité des programmes, les administrateurs entreprennent souvent des évaluations très tôt après qué les nouveaux programmes soient lancés, et des évaluateurs sont de plus en plus souvent invités à déterminer l'impact d'un programme qui ne fonctionne pas encore entièrement. En utilisant des exemples tirés de l'expérience d'effectuer une évaluation des résultats d'une réorganisation importante d'un système de prestation de services de santé très peu après que les changements aient été mis en application, on discute des défis et des avantages uniques d'évaluer un programme complexe tôt après la réorganisation. Spécifiquement, on décrit les rôles divers que les évaluateurs ont été requis de prendre et les leçons qu'ils ont appris à la suite d'augmenter leurs frontières professionnelles. En plus des rôles divers que les évaluateurs jouent souvent (comme éducateur, conseiller, chercheur), ceux qui sont responsables de la conduite des évaluations précoces de l'impact peuvent se trouver à agir en tant qu'entraîneurs, médiateurs et/ou therapeutes pour le personnel et l'administration de ces programmes pendant qu'ils essayent d'évaluer les résultats d'un programme qui n'a pas été entièrement mis en application.

In order to rapidly demonstrate the success of their efforts, administrators will often undertake evaluations very quickly after new programs or procedures are initiated. In today's climate of limited resources and evidence-based programming, their rush to document the worthiness of their activities is understandable. In fact, given that *offering* a service does not necessarily equate to *providing* the service, let alone providing a service that achieves quality *results* (Speer & Trapp, 1976), in many ways it is necessary for the acceptance, survival, and continued funding of new initiatives to demonstrate at the earliest possible date that desired outcomes are being achieved. However, the majority of stakeholders (including administrators, service providers, client groups, and funding agencies) may not be aware of the limitations inherent in early evaluation. Further, although it has long been recognized that evaluators require diverse skills (Caron, 1993; Posavac & Carey, 1997) and often take on diverse roles (Patton, 1997; Posavac & Carey, 1997; Preskill & Torres, 2000), many evaluators may not be prepared for the unique challenges they will face when attempting to evaluate a program in the early stages. The present article draws on examples from the authors' experience of evaluating the outcome of a major reorganization of a health service delivery system to demonstrate

the unique challenges and benefits of evaluating the impact of a complex program in the early phases following implementation.

REFORMING AND EVALUATING HEALTH CARE

As a result of the crisis in Canadian health care (Lewis, Donaldson, Mitton, & Currie, 2001), enormous attention is being given to health care reform in an attempt to provide health care services that are both more responsive to clients and more cost-effective to deliver. One reform option is to create "integrated health care systems" wherein the acute care services are linked in some fashion with other human service sectors (e.g., community services). The potential benefits that drive service integration include the improvement of health outcomes, client satisfaction, and system efficiency (Leggat & Leatt, 1997; Leutz, 1999). Integrating fragmented service sectors can help realize these improvements by reducing duplication of services, easing access to services, and coordinating services in order to meet the unique needs of clients. Because integration of acute and community services enhances communication and collaboration between the care providers in the various sectors and facilitates the transfer of care between the sectors, it is also seen as a primary avenue to effectively contain health care costs.

The Regina Health District (RHD) administration recognized the benefits of integrating the health service delivery system within their region. To facilitate service integration, the System-Wide Admission and Discharge Department (SWADD) was established in 1997. SWADD, formed by restructuring the duties and administrative accountability of existing acute and community care personnel into five distinct care teams, exists to usher clients through their episode of care (both within the hospital and upon their return to the community), to ensure that clients are receiving the appropriate level of care at any given point in time.

EVALUATION OF SWADD

In 1998 the Clinical Research and Development Program of RHD was approached by administration to undertake an evaluation of the impact SWADD was having on service delivery. An external consultant was contracted to manage the research process, and the evaluation began in March 1999. The primary objective of this evaluation was to examine whether integration of the acute and commu-

nity services through the re-organization of staff into SWADD teams was improving the admission/discharge processes within the district. Specifically, the effectiveness of SWADD in improving quality of care, access to services, integration of services, health outcomes, and cost-efficiency of service delivery was targeted for evaluation.

Although much attention has been given to evaluating the impact of service delivery reform (see Rowan, 2000, for a review), it has been argued that the literature regarding evaluations of care delivery models is sparse, fractured, and currently unable to inform policy decisions to any great degree (Abelson & Huchison, 1994, in Rowan, 2000). Conducting rigorous and systematic evaluations of service delivery reform not only is vital for the program in question, but has implications for service delivery on a larger scale.

Evaluation Methodology

The evaluation involved a series of studies that addressed various aspects of SWADD's effectiveness. The broad-based design used complementary methodologies to obtain information from all relevant stakeholder groups. The gathering of qualitative and quantitative information allowed a comprehensive examination of the impact SWADD was having on the delivery of health services in the RHD.

Study 1: Focus Groups/Key Informant Interviews

The purpose of the first study was to gather detailed information in a semi-structured format from all stakeholder groups regarding their impressions and opinions of the impact of SWADD on the target variables. The semi-structured discussion format allowed for the topics of interest to be addressed, but also allowed for the emergence of unanticipated themes. In total, 61 RHD clients from various service areas (Medicine, Emergency [ER], Surgery, Women's & Children's Health, Long Term Care [LTC], and Home Care [HC]), and 33 acute and community service providers (physicians, nurses, HC and LTC providers, social workers, and therapists) participated in focus groups or key informant interviews.

Study 2: Mail-Out Surveys

Based on the predominant themes that emerged from the focus groups, questionnaires were developed to distribute to larger, more representative samples drawn from the stakeholder groups. The

questionnaires targeted quality of admission and discharge services, access, integration, health outcomes, and cost-effectiveness. In total, 328 clients (response rate of 44%), 174 RHD service providers (response rate of 38%), and 29 service providers from partnering health districts (response rate of 44%) responded to the questionnaire.

Study 3: Retrospective Chart Audits

The final evaluation component involved a retrospective chart audit. A total of 900 charts, from the acute service areas most involved with SWADD interventions (ER, Medicine and Surgery), were retrospectively examined. The chart audit covered three time periods, with 1996 representing the time period prior to program implementation and 1998 and 1999 the time periods post program implementation. Three hundred charts from each time period were randomly selected. Level of acuity was coded using InterQual ISD criteria, which was chosen because it is an objective, well validated, and widely used tool for assessing acuity of both admission and days of stay (InterQual Products Group, 1996). The charts were also reviewed with respect to various other dimensions (e.g., readmission rate and ER visits subsequent to discharge, community referrals, case management, and barriers to discharge).

Evaluation Results Overview

The various evaluation phases converged on the conclusion that SWADD had made some strides toward achieving the proposed benefits of integration in certain areas, but not others (see Table1 for a summary of the findings; a full report on the evaluation findings is available from the authors). Specifically, the findings indicate that SWADD had a modest impact on quality, access, and integration of health services within the RHD. Although no positive impact on health outcomes could be demonstrated, neither was there an indication of any negative impact, dispelling the fears voiced by many that SWADD's focus on timely discharge from acute services was placing the health of RHD clients in jeopardy.

The most significant area in which the implementation of SWADD appeared to have a negative impact was on service provider morale and relations between management and staff. Staff members and service providers from all sectors felt uninformed about SWADD's rationale and procedures, angry that they had not been included in

Table 1
Summary of SWADD Evaluation Findings

	Physicians	SWADD Staff	Non-SWADD Staff	Clients	Chart Audit
Successful Aspects	• The availability/ responsiveness of the SWADD director • Increased ability to move clients through the acute care system • Savings to the acute care sector • One point of access to acute and community services	• One point of access to acute and community services • Increased ability to move clients through the acute care system • Enhanced communication and teamwork between health care providers from the different sectors • Increased ability to manage complex clients and facilitate timely access and follow-up	• Quicker and more appropriate long-term placements • Accessibility of SWADD teams • Better client management through discharge planning and team approach to service delivery	• Satisfied with admission procedures • Satisfied with access to community services • Satisfied with coordination of services across the sectors	• Non-acute admissions decreased (42.7% in 1996; 36.9% in 1999, $p < .05$) • Number of home care referrals at discharge increased (6.7% in 1996; 15.3% in 1999, $p < .01$) • Case management of acute clients increased (5.4% in 1996; 36% in 1999, $p < .01$) • No change in proportion of clients discharged while acute
Unsuccessful Aspects or No Change	• Physicians feel their practices are being monitored • Need to identify clients who could be cared for in the community earlier in their hospital stay	• Role ambiguity • Limited resources and staff • Lack of service provider education • Narrow focus on admission screening • Managing a system rather than individuals • Not enough focus on discharge planning	• Role ambiguity • Focus on the acute care sector to the detriment of community sector issues • SWADD staff uneducated about community issues • Lack of standards regarding client information received with community service referrals	• Less satisfied with discharge from hospital • Feel pressured to leave hospital • Protocol does not attend to emotional needs	• No change in average length of hospital stay • No change in number of non-acute days of stay • No change in the average delay of discharge • No change in readmission rate • No change in subsequent ER visit rate

the restructuring process, confused about their own roles and duties within SWADD protocol, unsure about role boundaries under the new system, frustrated that management was not taking their concerns seriously, and they were operating under the perception that SWADD had been implemented to meet the needs of management rather than to facilitate service delivery. Physicians were the most dissatisfied with the new procedures, especially acuity screening and utilization monitoring. They viewed screening as "bureaucratic meddling" and felt that the new protocol was questioning their clinical judgment and threatening the self-regulation of their practices.

Considering the current climate of fiscal restraint surrounding health care in Canada, the high cost of providing non-necessary acute care (Health Services Utilization and Research Commission [HSURC], 1994), and SWADD's motto of providing the "right care, at the right time, at the right place," it should not be surprising that the presence of non-acute patients being cared for within a hospital facility was a primary performance indicator of interest. The chart review revealed that the proportion of non-acute patients in the hospital did decrease slightly since the inception of SWADD (see Table1). However, this change is only marginally significant, and the 1999 review revealed that over one-third of patients being admitted to the hospital were not in need of acute care following SWADD implementation. Further, the proportion of non-acute days of stay did not change after the implementation of SWADD. These figures, while consistent with other Canadian findings (see Kalant, Berlinguet, Diodati, Dragatakis, & Marcotte, 2000 for a review), indicate that the objective of deterring non-acute admissions and improving the cost-efficient use of acute care resources during hospital stays does not appear to have been realized by SWADD efforts.

CHALLENGES TO CONDUCTING AN EARLY EVALUATION

Although the evaluation indicated that some objectives of SWADD were not being met, it has to be reiterated that this evaluation was conducted very soon after implementation. In fact, although the program was officially implemented in 1997, at the time of the evaluation many aspects of the protocol had not yet been practically implemented. This may account for some of the lack of demonstrated changes. But, in line with the philosophy that the "journey is more important than the destination," the challenges encountered while conducting this evaluation held many hidden benefits for the program stakeholders and the evaluators alike. In our opinion, many of these

hold the potential to be more beneficial to the program in the long run than definitive conclusions regarding SWADD's success.

Conducting research in the real world comes with inherent limitations and challenges that all program evaluators must contend with. All evaluators must take into consideration issues such as multiple stakeholder groups, time and financial restraints, an inability to control implementation efforts, lack of an appropriate control group, absence of baseline data, outside constraints on evaluation design, and so on (Posavac & Carey, 1997). Further, all evaluators find themselves assuming multiple roles, such as "facilitator, problem solver, educator, coach, and critical friend" (Caracelli, 2000, p.103). While conducting this outcome evaluation at such an early stage of program implementation, we discovered that it was necessary for us to draw on skills and assume unanticipated roles, over and above those usually required of evaluators (Caron, 1993; Patton, 1997). As stakeholders are increasingly motivated to demonstrate the efficacy of their programs at ever earlier stages, more evaluators will be faced with many of the same challenges encountered by our research team. The challenges faced by us in our efforts to assess the impact of reform involving widespread changes are outlined below, along with the lessons we learned along the way and the roles that we were required to assume.

Lack of Program Integrity

Evaluating a program at an early stage may mean trying to determine the success of an initiative that is not fully implemented. Programs are rarely implemented exactly as planned at the onset, and it falls to the program evaluator to determine how closely the "program in action" matches the "program in theory." Breaches of program integrity jeopardize the success of the program efforts, but can also render evaluation findings ambiguous at best or misleading at worst by resulting in Type III errors (errors of interpretation due to the inadequate implementation of program activities; Steckler, 1989).

Lesson Learned: In this program, the service providers did not know how to carry out the processes or components involved in the initiative. It not only involved changes in policies, procedures, and protocols, but also entailed major staff reorganization. We quickly realized that the majority of the program components either were not yet implemented or were implemented partially or incorrectly at the time of the evaluation. In such a case, it is unrealistic to judge

the success of the program by the degree to which the stated objectives are being met. Rather, the evaluator must be flexible enough to determine the discrepancies between the "program in theory" and the "program in action," and interpret the findings accordingly.

Evaluator Role – Protocol Trainer: When the efforts to inform and involve the stakeholders in change efforts have not been adequate, confusion over how to function under the new system can occur. During the SWADD evaluation, many focus group participants relayed the sentiment that the evaluation team members were the first people to not only inform them of SWADD protocol, but also take the time to try to answer their questions. Although we did not design or implement SWADD, it fell to us to relay information regarding the new program, address the questions, and forward the concerns to the RHD administrative body overseeing SWADD.

Vague Goals and Objectives

Early stages of program efforts are often characterized not only by vague roles and boundaries, but also by ill-defined program goals. Administrators and service providers may differ greatly on what they perceive to be the services that are targeted for change, and what they define as the intended outcomes, of both the program itself and the evaluation of the program (Posavac & Carey, 1997). Trying to evaluate a program that is not guided by agreed-upon goals is not an easy task (Mercer & Goel, 1994). How can successful attainment of program objectives be measured when the program objectives are not consistently defined?

Such was the situation encountered while conducting the evaluation of SWADD. It became apparent that different stakeholder groups had different perceptions of what SWADD was implemented to accomplish. SWADD objectives, as defined by administrators, primarily revolved around facilitating access to services and enhancing continuity of care. The primary objectives, as perceived by the majority of service providers, focused on cutting costs in the acute care sector by stringently monitoring the use of acute care resources.

Lesson Learned: When program goals or objectives are ambiguous or ill-defined, an evaluator can find him/herself at a loss on whether to evaluate the program based on the stated objectives of program administrators or on the pragmatic objectives perceived by front-line service providers. There may be pressures on administrators to

document ambitious or politically desirable goals that either the program as delivered cannot hope to achieve, or that are not the actual underlying objectives for the program. A program whose success is judged against the attainment of goals that the implemented changes do not target is doomed to appear non-successful.

Evaluator Role – Program Archeologist: An evaluator must determine the appropriate evaluation criteria by not only considering the documented program goals, but also by examining the changes that have been implemented and discussing with service providers what unstated desired outcomes drove such changes. In our case this involved delving beneath the surface of the documented goals and objectives. The stated goals of the program were non-specific (e.g., to better meet the health needs of clients, to use health care resources more efficiently). Our first task was to clearly operationalize these ultimate goals into more concrete and measurable short-term objectives (e.g., increase community referrals, increase case management of acute care clients, decrease non-acute admissions to hospital, decrease readmission and subsequent ER visit rate) and to examine how these were causally linked to program components. This was accomplished through detailed discussions with various stakeholder groups during the focus sessions.

In short, our very important first step was to develop a *program logic model*, which functions to identify poorly defined program components, program goals, and causal linkages (Rutman, 1980; Weiss, 1997; see Rowan, 2000, for a review of the importance of logic models in conducting evaluations of health care reform). Through this process the discrepancy between "program in theory" and "program in practice" became better understood, as the causal linkages between many program components and the goals were shown to be tenuous at best. Ideally, a formative evaluation (Scriven, 1967) involving the development of a logic model during the program planning stage should be conducted, in order to determine if the proposed theory of change is valid (Weiss, 1997) before changes are implemented.

Dissension Amongst Stakeholder Groups

That people fear change is a tried and true axiom. What is true for individuals is usually true at a grander scale for organizations that, after all, are made up of many people. Given the large number of people and professional groups involved in the SWADD reorganization, resistance to change was not surprising.

Service providers in the various health sectors expressed the unanimous sentiment that SWADD was designed to meet management needs and was implemented in an unrealistically short period of time. Most importantly, they felt that management had given no consideration to the staff who had to carry out the new procedures. They felt that the concerns of both service providers and clients had been ignored in the design and implementation process of both the reform and the evaluation. These sentiments stemmed from and reinforced the steady deterioration of service provider morale that has been experienced within health care as part of the crisis in Canadian health care over the recent years (Innes, 1997; Mulder, 2001).

As a result of these diverse factors, the evaluators faced conducting their investigation in an emotionally strained socio-political climate. The stakeholders (administration, service providers, and patients) expressed various degrees of animosity and distrust toward each other, and this dissension between the stakeholders resulted in many challenges for the evaluation team. First, the focus groups became a forum for each group to air their case to the evaluators. The evaluators were often called upon to relay concerns to management, or arbitrate disagreements between stakeholders. Further, it was not uncommon for emotions to run high as individuals recounted their frustration with SWADD and with employment within the health care sector in general. The evaluators quickly realized that they would have to be able to appropriately address the emotional needs of the focus group participants.

Lesson Learned: Stakeholder dissension and frustration with health care in general may have tainted the information being relayed to the evaluation team. Consciously or unconsciously it was possible that stakeholders were using the focus groups as a forum for venting their frustrations about health care in general or as a means of furthering the interests of their professional group within the district, rather than focusing on the impact of SWADD. When this possibility was probed, many focus group participants admitted that it was difficult for them to disentangle their dissatisfaction with SWADD from their dissatisfaction with the climate of health care in general. Thus the evaluation team continuously had to be aware that the current socio-political climate of health care and dissension between the stakeholder groups held the potential to bias the information and ultimately the evaluation findings.

Evaluator Role – Mediator and Therapist: Evaluators are usually called upon to examine the process and the impact of change (Posavac

& Carey, 1997). Evaluators must be cognizant that resistance to change and conflicting agendas of stakeholders may bias evaluation efforts. Beyond this, evaluators must also be aware of the emotional impact that can accompany change. Researchers conducting early evaluations must be particularly attuned to these issues, which are likely to be at a maximum level shortly following reform. When the stakeholders feel that no one else is listening to their concerns or championing their cause, it is likely that the evaluator will be implicitly expected to fulfil each of these roles. Although evaluators taking on an empowering role or supporting the program they are evaluating are not novel situations (Patton, 1997), we found ourselves being asked to support not the program and its goals, but the service providers in their attempt to empower themselves to resist the implementation of a program that they viewed as dysfunctional and unfair.

Evaluators should also be prepared for the possibility that early evaluations may interfere with program development. Ongoing evaluation may be used as an argument against making any immediate changes that might facilitate program efforts or increase staff morale (Posavac & Carey, 1997). Although it became apparent very quickly that certain roles should be clarified immediately in order to increase staff satisfaction and facilitate efficient service delivery, the administration decided to delay any modifications until the entire evaluation was completed in order to determine the scope of changes necessary. This was troublesome to staff members, some of whom came to view the evaluation as a means by which the administration could stall making refinements, leading of course to further frustration directed toward the evaluation team.

Interpretation of Evaluation Findings

Health evaluations are usually conducted to determine if programs are positively impacting care and/or service delivery. However, for all the reasons listed above, it is unlikely that a program will achieve its true potential or demonstrate a large measure of success soon after implementation. That being said, the question becomes “What is the use of conducting an impact evaluation in the early phases following implementation?”

If the evaluation framework is properly designed and all stakeholders are aware of the limitations, early outcome evaluations can provide valuable information to guide program development. Although the

RHD administration initially conceived of a "summative" evaluation (Scriven, 1967) wherein the results are used to make decisions about the continuation of a program, the research team was able to help them perceive the evaluation effort as more "formative," or as a means to gain information to drive program development. While the wisdom of conducting an early summative evaluation is questionable, the recommendation is always to conduct formative evaluations as soon as possible (Posavac & Carey, 1997). Formative evaluations provide an indication of whether the changes are addressing the target goal and help to guide program development to enhance the likelihood that the target goals will be reached.

To maximize the useful information stemming from the investigation, the SWADD evaluation design not only examined performance indicators as measures of success (as was the initial sole focus of administration), but also included an investigation of process and systemic variables that may be acting as obstacles to success. As a result, we were able to move beyond concluding that the program did not appear to be meeting all of its stated or intended objectives, to examining why this was the case. The latter is often the more important question in the early stages of a program's life. A formative evaluation framework allows for proper interpretation of early outcome evaluation findings.

Many potential barriers can threaten health care reform efforts (Saskatchewan Association of Registered Nurses, 1998), and the evaluation revealed that both process and systemic barriers were functioning to hinder success (see Table2). Specifically, it was concluded that both system and process issues were preventing the protocol from being implemented as planned. Given these obstacles, it was not surprising that SWADD had not been entirely successful in meeting all its stated objectives. If the evaluation team had conducted the simple summative evaluation examining the bottom-line outcomes initially desired by the administration, the results would have led to the conclusion that the program was not achieving a significant degree of success, but would have provided no clue as how to improve the situation. Instead, the health district was given a clear picture of what aspects of the program had achieved some positive impact, as well as the factors that were impeding SWADD's success and pragmatic recommendations regarding how to address these factors.

Lesson Learned: Although stakeholders may desire an impact evaluation in order to demonstrate the efficacy of their programs, they

often do not recognize the importance of including an evaluation of how the program is being delivered. If an early evaluation, which focuses on outcome variables alone, is to be undertaken, stakeholders need to be made aware of the inherent limitations of such an endeavor. Essentially, early outcome evaluations are more likely to falsely label a program as non-successful, simply because the program is not yet being fully implemented as intended.

Evaluator Role - Educator: Because individuals who do not routinely conduct applied research cannot be expected to know these constraints, it falls to the evaluators to educate the stakeholders regarding what they can reasonably expect from early evaluation efforts. Further, it is also the responsibility of the evaluator to properly interpret evaluation findings in the context of limitations and constraints, to enhance the likelihood that the findings will be used appropriately. It is also reasonable to expect that the evaluator may find him/herself educating the stakeholders regarding the benefits of including an examination of process issues in order to facilitate the interpretation of the findings.

Table 2
System and Process Obstacles to SWADD's Success

System Issues	• Lack of community-based alternative levels of care results in inability to discharge non-acute patients from hospital • No follow-up of non-acute admissions during their stay • Service delays (i.e., diagnostic tests, consultations) result in inability to discharge non-acute patients from hospital

Process or Implementation Issues	• Service provider concern over narrow definition of health impeded implementation – focus on acuity may be to the detriment of client emotional well-being
	• Lack of education and communication regarding SWADD procedures or protocol
	• Role ambiguity
	• Physician non-compliance with screening/discharge protocol
	  • View screening as "bureaucratic meddling"
	  • Challenges physicians' clinical judgement
	  • Threatens self-regulation of practice
	  • Physician ultimately responsible for patient's health
	  • No incentive for change
	  • Lack of enforcement procedures

Managerial Instability

Efforts to make large-scale changes within already complex organizations can be plagued by staff instability. During the time period that the evaluation was being conducted, both the Director of SWADD and the CEO of the RHD, who originally envisioned SWADD, resigned their posts. At the time of dissemination of the evaluation results and recommendations, SWADD was without a director. As a result, although many concrete recommendations regarding changes that should occur to enhance SWADD efforts were put forth by the evaluation team (see Table3 for a summary of recommendations), there was no person who had the mandate and authority to enact such changes.

Lesson Learned: Change breeds change, and it is not unlikely that a change in management may accompany a change in protocol, especially when the new initiatives are not popular amongst many stakeholders. If the new management were not involved in the evaluation plan they may be less invested in the evaluation. In our case,

Table 3
Recommendations Stemming from SWADD Evaluation

General Recommendations	• Clearly define SWADD structure, activities, and target objectives • Actively include all stakeholders in decision-making and program development efforts • Inform all stakeholders of changes – active and repeated dissemination and education • Continue with regular evaluation efforts to monitor implementation
System Recommendations	• Review need for increased resources in community sector • Educate physicians regarding community care options • Educate public regarding the changing role of acute care
Program Recommendations	• Focus SWADD intervention efforts on elderly/medically complex clients • Continue to enhance efforts to allow information sharing between the integrated health sectors (with a focus on accuracy, timeliness, and quality) • Determine placement of program within existing administrative structure • Clearly define the utilization screening process • Clarify role boundaries under SWADD protocol

the major frustration lay in the fact that lack of management translated into lack of ability to act on the recommendations stemming from the evaluation. This frustration led to asking ourselves the question, "What role does the evaluation team have in becoming the agents of change?"

Evaluator Role – Change Agent: An evaluator may find him/herself taking responsibility and attempting to become the driving force behind enacting the evaluation recommendations. We found ourselves responsible for disseminating the results to the stakeholder groups, and not only presenting, but supporting, the recommendations suggested by the evaluation findings. We became invested in the recommendations, and found ourselves involved in working committees where we lobbied for specific changes. While it does detract from the "objective evaluator" stance of an external evaluator (Posavac & Carey, 1997), evaluators have to be ready to take responsibility and take action, if we truly believe that program success is at stake.

CONCLUSION

As seen from the above discussion, many potential challenges await those who undertake impact evaluations very shortly following the implementation of large-scale and complex programs. But, on a positive note, many benefits potentially stem from beginning evaluations early in a program's development as well.

The evaluation of SWADD while it was still a young program allowed for its level of success to be tentatively gauged. But more importantly, it allowed for areas of improvement to be identified that, if addressed, would enhance the likelihood of SWADD achieving its full potential in the future. When they are young, programs are also relatively flexible, making further change or restructuring easier than at later stages following implementation when the new protocols have become familiar or ingrained.

Despite the inherent challenges, we found that conducting an early impact evaluation was worth the effort. While the efforts may be premature in relation to making summative statements regarding the ultimate worth or success of a program, they can serve to provide a recommended road-map for the future and enhance the probability of success. A framework that stresses the formative nature of the investigation, and evaluators prepared to assume multiple

and varied roles, can produce a vivid picture of where the potential of a program lies, and what factors might be impeding that potential from being fully realized.

REFERENCES

Caracelli, V.J. (2000). Evaluation use at the threshold of the twenty-first century. *New Directions for Evaluation, 88,* 99–111.

Caron, D.J. (1993). Knowledge required to perform the duties of an evaluator. *Canadian Journal of Program Evaluation, 8*(1), 59–78.

Health Services Utilization and Research Commission. (1994). *Barriers to community care: Final report.* Saskatoon: Author.

Innes, G. (1997). The Canadian health care crisis and how to survive it. *Journal of Emergency Medicine, 15,* 410–415.

InterQual Products Group. (1996). *The InterQual Review System*. Marlborough, MA: InterQual Inc.

Kalant, N., Berlinguet, M., Diodati, J.G., Dragatakis, L., & Marcotte, F. (2000). How valid are utilization review tools in assessing appropriate use of acute care beds? *Canadian Medical Association Journal, 162*(13), 1809–1813.

Leggat, S.G., & Leatt, P. (1997). A framework for assessing the performance of integrated health delivery systems. *Healthcare Management Forum, 10,* 11–18.

Leutz, W.N. (1999). Five laws for integrating medical and social services: Lessons from the United States and the United Kingdom. *Milbank Quarterly, 77*(1), 77–110.

Lewis, S., Donaldson, C., Mitton, C., & Currie, G. (2001). The future of health care in Canada. *British Medical Journal, 323,* 926–929.

Mercer, S.L., & Goel, V. (1994). Program evaluation in the absence of goals: A comprehensive approach to the evaluation of a population-based breast cancer screening program. *Canadian Journal of Program Evaluation, 9*(1), 97–112.

Mulder, D.S. (2001). Current health care crisis: A Canadian perspective. *Archives of Surgery, 136,* 169–171.

Patton, M.Q. (1997). *Utilization-focused evaluation*. London: Sage.

Posavac, E., & Carey, R. (1997). *Program evaluation: Methods and case studies* (5th ed.). Upper Saddle River, NJ: Prentice Hall.

Preskill, H., & Torres, R. (2000). The learning dimension for evaluation use. *New Directions for Evaluation, 88,* 25–35.

Rowan, M.S. (2000). Logic models in primary care reform: Navigating the evaluation. *Canadian Journal of Program Evaluation, 15*(2), 81–92.

Rutman, L. (1980). *Planning useful evaluations: Evaluability assessment.* London: Sage.

Saskatchewan Association of Registered Nurses. (1998, October). Facilitating an integrated approach to primary health care in Saskatchewan. *Primary Health Care Newsletter,* 26–29.

Scriven, M. (1967). The methodology of evaluation. In R. Tyler, R. Gagne, & M. Scriven (Eds.), *Perspectives of curriculum evaluation* (pp. 39–83). Chicago: Rand-McNally.

Speer, D.C., & Trapp, J.C. (1976). Evaluation of mental health service effectiveness. *American Journal of Orthopsychiatry, 46,* 217–228.

Steckler, A. (1989). The use of qualitative evaluation methods to test internal validity: An example in a work site health promotion program. *Evaluation and the Health Professions, 12,* 115–133.

Weiss, C. (1997). Theory based evaluation: Past, present and future. *New Directions for Evaluation, 76,* 41–55.