PROGRAM EVALUATION

Workshop Module for
Transdisciplinary Understanding and Training on Research – Primary Health Care (TUTOR-PHC)

A collaborative initiative by

[Logos of Dalhousie University, McMaster University, Université de Montréal, and Western University]

Funding provided by

[Logo of CIHR IRSC]

Institutes of:
Aging; Aboriginal Peoples Health; Gender and Health; Health Services and Policy Research;
Nutrition, Metabolism and Diabetes; Population and Public Health

Compiled June 19, 2007
Program Evaluation
TUTOR-PHC Workshop Module

Objectives
- Knowledge: to obtain information on the five different components of evaluation, issues on causal inference and program logic models.
- Comprehension: to understand the problems with "black-box", outcomes only evaluations and the need for theory-driven, causal model evaluations.

Reading Assignments

Week 1
- Backgrounder prepared by the facilitator.

Week 2

Week 3

For Reference (not a requirement)

Tasks/Exercises:
- Week 1 – Discuss two case studies in program evaluation.
- Weeks 2-3 – Discuss the development of causal logic models as the basis for planning program evaluation. Participants apply principles of causal logic models to their own research.

Discussion Summary
Number of participants: 6
Total number of messages: 49
Number of facilitator posts and comments: 12

Week 1 – Original facilitator post
Please read the assigned pre-readings. While reading the papers, discuss among yourselves:
- What your initial thoughts are about the outcomes and conclusions?
What lessons were learnt?
What insights were gleaned?

Number of participant posts: 17
Number of facilitator comments: 5

Sample facilitator comments
- I will be reading everyone's comments, but won't be responding to each comment, as I would like for some discussion to ensue. In fact I find that by the time I have gone through 2-4 other postings, everyone has picked up on the issues, challenges, etc and pre-empted what I was going to say. This is actually good! So, I will be checking regularly but if you don't hear from me after every posting, don't worry. Be assured I am reading the postings and will comment after there has been a little discussion or when there is an issue that needs clarification or discussion from me.
- Regarding theory, I think this can be another challenge for some in clinical and behavioural medicine and health research. Again the social sciences are big on developing and using theories for any research they do, including student theses, yet for those coming from some other disciplines where theory is not an automatic expected requirement of any research question you have, it could be a little bit of a challenge. Much research ultimately focusses on behaviour change, whether it is with patients in trying to get them to engage in health promoting behaviours or regimen compliance, or with colleagues or organizations in trying to get them to uptake and utilize new knowledge (knowledge translation), there are many behaviour change, persuasion, knowledge diffusion, organizational change, etc., theories. This means going to these literatures to ferret out the theories.
  One good synopsis starting point is this website of NIH National Cancer Institute. I use it as a text. It is an excellent companion to have.

Week 2 – Original Facilitator post

One of the biggest challenges of evaluation is that more often than not, people think through an evaluation backwards. I don't know how many times I have seen/heard, we have data on, for example, the number of at risk women with low birth weight babies that come to this pre-natal advantage clinic, and I'll find a control group to see if the program made a difference.

What they are doing is identifying available DATA, and then thinking through the DESIGN, and maybe they might CONCEPTUALIZE, but generally not, how and why the program should work.

In fact it should be “CON-DE-DA”

It should always be done the reverse, conceptualize, then design and what data do you need.

The first step is to conceptualize the causal model - meaning what do you hypothesize will be the steps leading to get to your final outcome. This must
really come first. If you look at the Vingilis & Pederson paper with the generic causal model you will see that in addition to considering outcome and what "intervening" or "mediating" variables could be affecting outcome, one should also consider what "exogenous" variables (pre-existing conditions) could be affecting your outcome and "stochastic disturbance" is just a fancy word to error, which means have a large enough sample size to address power. (We in social science also have to come up with some Latin-Greek words).

In my paper I present another example of "black-box" evaluation whose conclusions (outcome) were not causally or conceptually linked to the intervention. It was described in the newsletter of the Canadian Health Services Research Foundation, a governmental granting agency for health services research. The byline of the newsletter article reads "Research shows hospital cuts don't hurt health, access to care." The actual report by the Health Services Utilization and Research Commission (HSURC) in Saskatchewan concluded that closing small rural hospitals and cutting acute care hospital beds in major cities actually reduced death rates (if you are really interested you can check the actual study which is on-line: Health Services Utilization and Research Commission. Assessing the Impact of the 1993 Acute Care Funding cuts to Rural Saskatchewan Hospitals Summary Report No. 13, Sept. 1999, www.sdh.sk.ca/hsurc). In 1993, when 52 small hospitals were closed or converted to health centres because of spending cuts, the Commission evaluated the impact of the closures. First the Commission analyzed hospital use and mortality data in Saskatchewan before and after the cuts, from 1990 to 1996 which they stated would "determine whether access to hospital care or the health of rural residents was affected" (p.2). To determine whether or not these cuts affected the health of rural residents, they examined death rates as a proxy measure for health status. They compared hospital use and death rates among four groups: 1) communities that stopped receiving acute care funding to keep their hospitals open, 2) rural communities that never had hospitals, 3) rural communities that still maintained their rural hospitals, and 4) the rest of Saskatchewan. They tabulated the rates of heart attack, stroke, total, premature and motor vehicle injury (MVI) deaths per 100,000 population using a simple pre-post design (1990-1992 vs 1993-1996) for communities that never had small hospitals, for communities that kept small hospitals open and for communities that closed small hospitals in 1993. They concluded that, "Health status as measured by death rates, improved throughout the province during the study period. Communities that experienced the 1993 acute care funding cuts had the largest overall improvement in mortality rates; communities that still have small hospitals, the smallest. Communities that never had a hospital had the lowest mortality rates throughout the study period" (p.3). Death rates per 100,000 population showed pre-post hospital closure declines for heart attacks, strokes and all causes in all three communities (where hospitals closed, where hospitals remained open and where no hospitals had existed) and for the rest of Saskatchewan. However, after the 1993 hospital closures, deaths per 100,000 due to motor vehicle injuries and premature deaths increased for communities which still kept hospitals open, and decreased for the other two communities with no and closed hospitals, and for the rest of Saskatchewan. In addition, they found hospital use declined throughout the province: communities with hospitals whose funding for acute care was cut had the sharpest decline in rates, while communities which still maintained their rural hospitals continued to have the highest hospital use rates.
Second the Commission interviewed residents through telephone surveys and focus groups about their opinions of their hospital closures. The telephone survey found that 18% were dissatisfied with health services prior to the 1993 funding cuts while 54% were dissatisfied after the funding cuts. Reasons for concerns were lack of local emergency services, decreased availability of health services, doctors and travel distance to access care. Yet the report also indicated that of those who were telephone-surveyed, 89% stated that the funding cuts had no direct effect on their personal health and 64% were in good health. The conclusions of the study were that,

"Cutting the acute care funding to 52 rural Saskatchewan hospitals has not adversely affected the health status of residents in these communities. .... Community residents' perceptions were consistent with the mortality data. Despite widespread fears that health status would deteriorate, respondents overwhelmingly reported that the loss of acute care funding did not adversely affect their own health or their family's health. In addition, two out of every three people surveyed rated their current health status as either good or very good. ...Although we have no data to confirm that removing acute care funding from communities improved residents' health status, some of the data are intriguing" (p.9).

The implicit causal model here is that hospitals increase deaths.

Unfortunately a simple pre-post design, within a "black-box" approach was not appropriate for conducting this evaluation. First, the three community groups had different population-based death rates even before the intervention, ranging from 680 to 880/100,000 population. The highest death rates pre- (and post-) intervention existed in communities with hospitals, while the lowest pre- (and post-) death rates existed in communities which never had hospitals. Using the generic model in the paper, clearly there are pre-existing conditions (exogenous variables) which could affect the results, and which needed to be accounted for, in the research design and statistical analysis. Moreover, no mediating (intervening variables) factors were discussed and included to explain the findings. That is, the causal mechanisms by which the changes or, lack thereof occurred, has not been measured, so that the HOW and WHY are not understood.

Of particular concern should be the findings that deaths due to MVI actually decreased in communities that closed hospitals and in which average driving distance to nearest hospital had increased to more than 50 km, yet increased in communities which kept hospitals open. Empirical evidence indicates that most critically injured patients can be saved if definitive surgical intervention is provided within one hour (the "Golden Hour" for emergency treatment). It is hard to believe that an injured patient's chance of survival would not be affected by having to endure an additional 30-60 minute drive. To explain the findings of this study and to eliminate other possible causal factors, one would need to assess the exogenous pre-conditions, the actual implementation of the closures, and mediating variables. In other words, HOW and WHY did some death rates decrease in the experimental communities and increase in the communities that still kept their hospitals open? The report does not attempt to explain the increase in MVI mortality in the still-open hospital communities. Rather it states: "The data do not suggest that the affected
Program Evaluation

communities have suffered... as a result of the 1993 acute care funding cuts; if anything, death rates for heart attacks and motor vehicle trauma, is no different and possibly better". Had they conducted a theory-based, casual model evaluation, they would have questioned the use of a population health approach to evaluate hospital closure outcomes over a short time period and a priori, presented the causal mechanisms for expected outcomes. As it stands, this black-box evaluation begs more questions than it answers.

Regardless of results, simple input/output or "black box" type of evaluation often draws conclusions that are less than satisfactory. On the one hand, "black-box" based positive findings of intervention success may be difficult for policymakers or practitioners to apply. For example, the Alberts et al. study, mentioned in the paper, evaluated a multi-faceted intervention of great causal complexity. In this study, even if the causal link that the researchers claim to have established was valid, one cannot determine which and how much of the above mentioned interventions resulted in the desired positive outcome. Thus this type of evaluation typically provides no information about which combination of components and treatment strength of components are necessary to obtain the desired outcome.

Unfortunately, poor evaluations do not necessarily sit on shelves. They can and have been used to drive policies, services, etc

In weeks 2-3 (the materials can be read now) will focus on how to think through causal logic models as the conceptualization of any intervention needs to come first. At this website:

http://www.gov.ns.ca/health/primary-care/ is Strengthening Primary Care in Nova Scotia, An Evaluation Initiative. You should have already read their program logic model (reading # 5) now you can read their progress report, in particular the evaluation section. Might I suggest that you challenge yourselves in thinking about what went wrong with their evaluation, how you would improve both the logic model and/or the intervention and evaluation. The different readings can help plus, you all seem to have much insight into the issues already.

Think about your own research. Is it an evaluation question or can you develop an evaluation question from your area? Can you think through the basic causal model? Can this causal model then be turned into a logic model? Importantly, though, logic models are rarely completed by just one person. So it will be important for all to provide insights, feedback and challenge the hypothesized causal links and theories of program and implementation.

Number of participant posts: 13
Number of facilitator comments: 3
Sample comments

- Interesting and astute comments. Indeed, logic models have taken on a cult status and many people do them to do them. Like vision statements of a bigone era, they are crafted and them promptly ignored, forgotten, etc.

  There is no single correct type of logic model. Logic models need to be tools to be used in different ways; they are blueprints. I have made
detailed ones with indicators and step by step processes included to help focus the implementation and I have done simple ones for board of directors to get a snap shot of what the program is about. Additionally, I have never done them on my own, but only acted as a facilitator with a large group of stakeholders, including the clients to use the program. Suffice it to say that good, useful models with causal links established are rarely done in a week. Normally I facilitate them through many meetings, discussions and hashing through exactly what is the program about, what it is trying to do, exactly how will it work and how will we know if we are meeting all our implementation, shorter and longer term objectives.

Week 3 – Original Facilitator post
Hi everyone - just a reminder on the week's task
In weeks 2-3 we will focus on how to think through causal logic models as the conceptualization of any intervention needs to come first. At this website: http://www.gov.ns.ca/health/primary-care/ is Strengthening Primary Care in Nova Scotia, An Evaluation Initiative. You should have already read their program logic model (reading # 5) now you can read their progress report, in particular the evaluation section. Might I suggest that you challenge yourselves in thinking about what went wrong with their evaluation, how you would improve both the logic model and/or the intervention and evaluation. The different readings can help plus, you all seem to have much insight into the issues already. Think about your own research. Is it an evaluation question or can you develop an evaluation question from your area? Can you think through the basic causal model? Can this causal model then be turned into a logic model? Importantly, though, logic models are rarely completed by just one person. So it will be important for all to provide insights, feedback and challenge the hypothesized causal links and theories of program and implementation.

Number of participant posts 9
Number of facilitator comments: 3
Sample comments
 o So one tries to push for more time and IMPORTANTLY, collaborate, collaborate, collaborate. I have rarely done an evaluation on my own. Rather I bring together a team, like yourselves, because even if you have a shorter time frame, the hashing through of the logic model, thinking through the "if-then" assumptions, identifying indicators, etc, is always better done with a group and with consultation from as many stakeholders as possible, including the end users.

Additional References provided during the workshop
5. For those interested in the issues, methods and techniques for implementation, your 
might be interested in this study on the topic. Go to: 
http://www.lawson.on.ca/publications.html an click on the knowledge diffusion paper.
http://www.cdc.gov/std/program/progeval/3-PGprogeval.htm
Interprofessional perception scale (Duncanis & Golin, 1981) Interprofessional Team 
effectiveness scale, Jefferson Survey of attitudes towards physician-Nurse collaboration

Participant Feedback
Average workshop rating (1=poor, 5=excellent): 4.1
Sample comments

Comments related to the instructor
I thoroughly enjoyed the comments and materials provided by [the facilitator] – she 
always had insightful, relevant comments and made an effort to respond to queries in a 
timely manner.

The most useful part of this workshop was
Learning about relevant literature that I was not familiar with.
Application to real life situations and our own research.

Suggestions for improvement
I think an exercise might have provided a common thread as a focus.

This workshop module was offered in June 2005. For more information, please 
contact the TUTOR-PHC program manager at tutor@uwo.ca.