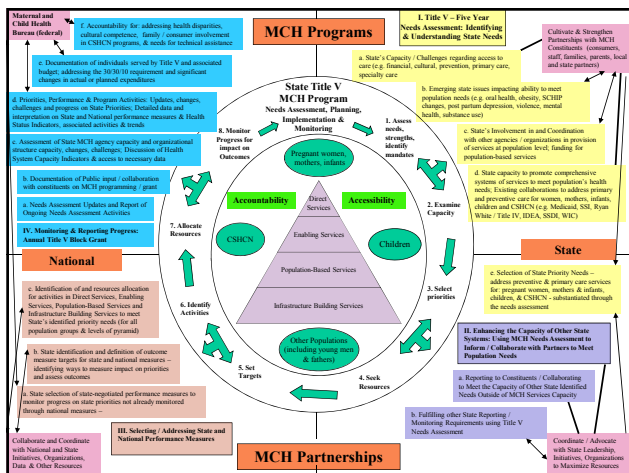# Program Evaluation

Donna J. Petersen
MCH Epidemiology Training Course
Minneapolis, Minnesota
Thursday June 2, 2011

Needs Assessment Planning Cycle

- What can you tell me about the planning cycle in Maternal and Child Health?
  - Yes, this is a quiz
- What are its origins?
- What is its purpose?
- What is its schedule?
- What does it require in the way of state capacity?

Planning Cycle

We gather data and information
  to build our knowledge base of community and population needs and assets
    in order to *make decisions* about
      how to best utilize our limited resources
        toward the best possible outcomes

Evaluation in the Planning Cycle

- So . . . How do we know if the *decisions* we made were good ones?
- We must include, in the planning cycle, plans for evaluation of our efforts, of our decisions, of our expenditures, of the consequences of our actions, intended or otherwise

Evaluation in the Planning Cycle

- We must create operating systems for the *monitoring* of our efforts and for the ultimate assessment of the outcomes of our efforts, the *evaluation*

Evaluation Outcomes

- Monitoring allows you to evaluate your process
  - Did I do what I intended to do in the ways in which I intended to do them?
  - If not, why not?
- Evaluation allows you to assess your outcomes
  - Did I achieve what I hoped to achieve through the processes that I put into place?
  - If not, why not?
  - Were there other unintended outcomes of these efforts?

Bottom Line: RESULTS

- Did your program make a difference?

- Funding agencies, legislators, the Governor, your agency head, other agencies, communities organizations, taxpayers, and families all want to know.

Bottom Line of the Bottom Line

- HOW did you achieve those results?
- By what means?
- Other state MCH programs, your colleagues, nonprofit organizations, academic institutions all want to know what you did to achieve your results

Evaluation is Essential

- For making mid-course corrections and changes in program implementation
- For determining if the program or policy has been effective
- For providing information for planning the next program or policy

MCH Program Evaluation

- We evaluate so we can *make decisions*
- If you will not be making decisions, it is a waste of time and money to evaluate
- Evaluation, done well, can be extremely informative; done poorly it can create more questions than answers
- Regrettably, "done well" means "done at the outset"

Evaluation Begins at the Beginning

- A program should begin with measurable objectives
- These objectives include intended targets from a baseline
  - Suggests that the program strategy was based on data
  - Suggests that data will be gathered throughout the life of the project

A Little Reality Check

- Is it ever the case that you are told to initiate a new program that is NOT based on your needs assessment nor even on any data?
- What is your level of responsibility to evaluate such efforts?
- How do you establish the baseline?
- What are appropriate objectives?

A Little More Reality Check

- Is it ever the case that you are asked to evaluate a program that has no baseline data, no objectives, not even a measurable purpose statement?
- What is your level of responsibility to evaluate these programs?
- How do you establish the baseline?
- What are the appropriate objectives?

Evaluation as Part of the Needs Assessment Planning Cycle

- What does this say about *stakeholder involvement* in evaluation?
- Is it appropriate to engage stakeholders in determining the questions to ask and the data to be gathered?
- Why or why not?
- Who are the stakeholders for evaluation?

Objectives

- A good plan relies on a set of *objectives*
  - to provide more specific direction
  - to frame your activities
  - to communicate your intentions
  - to ultimately enable you to evaluate your process and your outcomes

Measurable Objectives

- S   Specific
- M   Measurable
- A   Action-oriented
- R   Realistic
- T   Time-framed

Objectives

- Why measurable objectives?
- Because you need to know and be able to explain to others, what you intend to do by when, so that everyone knows your plan
- Because you need some kind of roadmap to guide your activities, the allocation of resources, staff assignments, etc
- Because you need a way to measure your success

No Objectives

- So, in the more common case where you are asked to evaluate something after the fact, you are still obligated to determine what it is you are measuring against what might have been the starting point, or the initiating incident or set of circumstances
- They are post hoc, but still essential

Objectives: Two Common Types

- Outcome – what you intend to achieve
  – Often the "what", indicates the desired state
- Process – how you intend to achieve it
  – Often the "how", the "when" and the "where"
- You need BOTH to measure success and others need BOTH to replicate what you've done

Process Versus Outcome Objectives

- You understand the difference . . .
- So, it should be obvious what are *process* objectives versus what are *outcome* objectives
- One reflects the *how*, the other reflects the *what*, right?

Choosing the right level

- An outcome objective can be a process objective and vice versa
- It all depends on your comfort level
- To what objective are you willing to be held accountable?
  - Reducing infant mortality?
  - Or enrolling women in a smoking cessation program?

Choosing the right level

- So, if you have determined that you are concerned about children's oral health . . .
- What is it you specifically want to achieve?
  - Reduce the incidence of dental caries in children under the age of 10 by 20% by 2015
  - Increase the number of 3rd grade children receiving dental sealants by 30% by 2014
  - Increase the number of municipal water systems with fluoridation by two by 2013

What can you realistically achieve?

- Continuing on . . .
  - Increase the number of dentists who will see children with public insurance from 50-75 by 2015
  - Increase the number of children on public insurance seen by dentists by 35% by 2015
  - Improve public insurance reimbursement rates for preventive dental visits by 20% by 2014

Can you be held accountable?

- And on . . .
  - Change dental practice laws to allow preventive care to be provided by licensed dental hygienists and pediatric nurse practitioners by 2015
  - Increase breastfeeding initiation rates from 70% to 85% by 2014
  - Eliminate sales of sweetened beverages from all elementary schools by 2013

Post Hoc Objectives

- So you walk into a "dental program" and are asked to evaluate it . . . What questions are you going to ask?

First Exercise

- Let's chew on this for a while
- Let's identify either programs you are developing that you would like to be able to evaluate and/or those you have lying around that you have been or will be asked to evaluate . . . .
- We'll need five good ones . . .

Let's Write Some Objectives!

- Take a few moments to write an outcome objective and a process objective around your program of choice from this list

Remember . . .

- S   Specific
- M   Measurable
- A   Action-oriented
  - Use action verbs, reduce, increase, etc
- R   Realistic
  - Based on data, literature, model programs, etc
- T   Time-framed

Program Evaluation

We (should) evaluate and monitor:
- to make decisions
- to assure fidelity to the plan or to make mid course adjustments
- to promote accountability
- to inform future plans
- and to advance the field (the long view)

Who has heard the term "Evidence-Based Practice"

- What does that mean?
- If you adopt an EBP do you have to evaluate your program?

Objectives and Evaluation

- We've already said objectives are critical
- It is also critical that we have the political will to make decisions based on evaluations
- Some scholars suggest conducting "pre-evaluations" or "evaluation readiness assessments"
- Might help shape the form of the evaluation

Evaluation

- Evaluation requires that you have a data gathering system in place for monitoring the process and a data gathering system in place for measuring the outcomes
- This implies that you have selected the measures needed to answer the questions posed
- Also important in post hoc evaluations

Evaluation as Applied Research

- The question being posed, or the "hypothesis" involves examining whether the program had any discernible effect on the problem being addressed beyond what would have happened by chance
- Can be measured at different levels

Evaluation Data Systems

- Depending on the design of the evaluation model, you may need measures on the major inputs, the activities, the intermediate outputs and the outcomes
- You also need measures that allow you to monitor performance along the way

Evaluation Measures

- Presumably, you have data sources that led you to develop your objectives in the first place
- It is important to determine all the measures needed to answer the ultimate evaluation question: did this program, intervention, or effort make a difference?

Data Sources

- Where will you find these data? Are they routinely collected? Are they housed in your agency or other agencies?
  - Are there issues with secondary data?
- Do you have to collect new data to answer the questions? What system will you put in place?
  - Are there issues with primary data collection?

Data Sources for Evaluation

- No different than the data sources for needs assessment!
  - Population data bases (census, vital records)
  - Surveys and surveillance systems
  - Program MIS data
  - Qualitative data
- Don't forget the literature – a good meta-analysis can save you a lot of heartache

Evaluation "Logic Models"

- It is helpful to create a logic model of the inputs, activities, outputs and outcomes in addition to intervening factors that may influence or affect the outcomes of interest
  - To assist in interpreting results
  - To be as thorough as possible in determining measures of interest
  - To have the needed data gathering systems in place

Logic Models

- Helpful in determining the appropriateness of evaluation questions, designs, measurements and data to be gathered
- What can reasonably be expected to effect the process and/or the outcome?
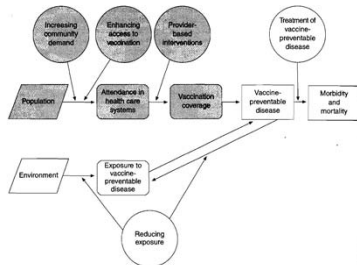- What characteristics of the environment, the population, the intervention should you take into account?

Figure 13-4 Evidence Model for Immunization Program Evaluations. *Source:* Adapted from Framework for Program Evaluation in Public Health, *Morbidity and Mortality Weekly Report,* 48 (No. RR-11), pp. 1–40, 1999, Centers for Disease Control and Prevention.

Measurement

- Develop specific indicators for each concept
  - Each identified concept may need to be measured in several ways
  - Remember to consider validity of the measures

Logic Models and Measures

- Measures should represent dimensions that are expected to change as a result of the program intervention
- Measures should also represent characteristics of program recipients (and controls), or of the program itself and of potential competing factors

---

First Small Group Exercise

- Pick one of the five topics we worked on earlier and develop a logic model for the evaluation
  - Consider the outcomes of interest
  - Inputs, activities and outputs
  - Any intervening variables that might be present
  - Characteristics of the program or the recipients
- Be sure to select a spokesperson to report back

Logic Model Discussion

- Themes?

Break for Lunch!

Second Small Group Assignment

- You were sent articles to review . . .
- We'll group by article
- Get in your small group and answer the questions about the articles
- Yes, this is intended to keep you awake after lunch!
- Select a recorder and a reporter

Reports Back

- Themes?

Types of Evaluation

- Formative Evaluation = process evaluation
  - Is the program implemented as intended
- Summative Evaluation = outcome evaluation
  - Are desired outcomes achieved
- Of course you need to do both to answer the question of interest

Formative Evaluation

- Includes assessment of fidelity to the plan
- ALSO includes an assessment of the relevance, completeness and quality of data being collected
  - Who is collecting the data?
  - How is feedback gathered from these sources?
- ALSO includes communication with stakeholders

Formative Evaluation

- Allows you to monitor the process of program or policy implementation
  - To make any adjustments necessary
  - To clarify expectations as needed
  - To document action steps taken for replicability
- Typically undertaken by the planners or program managers
- Supports iterative planning

Summative Evaluation

- Summative, or outcome, evaluation is typically undertaken by evaluators, often external to the program or organization
- The purpose is to determine the worth of the program, the value of the expenditure given the results achieved, and to make decisions on future investments or directions

Evaluation against what?

- Typically use some form of comparison against which to judge your outcomes
  - Where you were at baseline (i.e., the data that led you to identify this as a critical need in the first place)
  - The objective you set at the outset – did you reach it?
  - Another state or county that did not address this issue
  - National norms
  - The literature

Program - Evaluation

- Inputs
- Activities

- Process evaluation

- Outputs

- Intermediate outcomes

- Outcomes

- Outcome evaluation

Evaluation Design

- This links right back to the logic model we developed earlier
  - What are the outcomes of interest
  - How will you know that your intervention led to them
  - What were the steps along the way that contributed
  - What inadvertent results might emerge

Data Challenges

- Identifying a true control group
- Reaching agreement on data collection protocols with program staff, stakeholders and evaluators
- Determining what will be considered "successful" levels of effect (back to the objectives . . .)
- Data collection and use issues . . .(validity)

Evaluation as Applied Research

- Does this program benefit the public's health
  - Classic outcome evaluations
  - Economic evaluations (cost-benefit)
  - Process evaluations, formative evaluations
  - Intermediate and ultimate outcome evaluations

Evaluation as Applied Research

- Interested in efficacy
- But also interested in translation and applicability
- Can a program that works in a controlled setting work in the real world?
- Is an effective intervention acceptable to the public?

Evaluation as Applied Research

- Interest is in the *unique* or *net effect* of a program above and beyond what might have occurred due to myriad other factors
- Requires a research design that controls for these "myriad other factors"
  - Typically some type of control group

Evaluation as Applied Research

- Process: were services provided as intended to those intended to be served
- Intermediate Outcome: of those served, how many achieved the desired outcome
- Ultimate Outcome: among the population affected by the problem, how many were positively affected
- Economic Analysis: at what cost was this benefit achieved

Evaluation Designs

- Historical controls, pretest-posttest
- Comparison group
  - Could be the nation, other states
- Experimental group versus a control group
  - Could be counties, neighborhoods, clinics, schools
- Randomly assigned participants to a case or a control group

Considerations

- How control or comparison groups are formed influences validity of inference
- Sample size must be sufficient to detect differences
- Stakeholders must agree that what is being measured is relevant, important and can reflect change over time
- Evidence that program was implemented as planned and that the control group's experience was sufficiently different

Non-Experimental Designs

- Anecdotal case reports
- Intervention without a control
- Intervention with "literature controls"

Quasi-Experimental Designs

- Intervention with historical controls
- Case-control observational study
- Intervention against existing databases

Experimental Designs

- Simple randomized controls
- Randomized control trials

Classic Design

- The Solomon 4 Group
- Random assignment to a case or a control group; both given pre-test
- After the intervention, case group and control group both given the post-test
- Allows you to deal with various threats to internal and external validity

Threats to Internal Validity

- To what extent are the program or intervention effects really due to the program or intervention rather than competing explanations?
  - Maturation
  - Self-selection
  - Changes in instrumentation
  - Historical influences unrelated to the intervention

Threats to External Validity

- To what extent can the results be generalized to other situations?
  - Again, the controlled situation versus the real world
  - And/or, different populations, different communities, different circumstances may result in different outcomes for the same intervention

Fidelity to Program Design

- The quality of the process evaluation is very important to the validity of the outcome evaluation
- An elegant study of a poorly implemented program will likely indicate that the program did not work as intended – a good process evaluation would have told you that before you invested in the outcome evaluation

Last Group Exercise

- Let's brainstorm on the appropriate/feasible design for the programs we've been working on
- Also, consider some possible control or comparison groups
- Then, answer the questions on the hand-out
- Select a recorder and a reporter

Components of Evaluation Designs

- Characteristics of the Context
  - Sociopolitical environment, i.e. things happen while you plan
- Characteristics of the Participants
  - Demographic, SES, attitudes, behaviors, etc.
- Characteristics of the Program
  - Activities, services, staffing, materials, processes, etc.
- Characteristics of the outcomes
  - Anticipated and unanticipated
- Characteristics of the costs
  - Direct, indirect and opportunity costs

Measurement and Evaluation

- Again, the purpose of the evaluation will dictate the appropriate measures
  - If it is a classic outcome evaluation, you need outcome data
  - If it is an economic analysis, you need data on services and their associated costs, together with results data
  - If is a process evaluation, you need detailed information on program implementation

Evaluation Foci

- Again, important to decide at the outset what the focus of the evaluation will be
  - What decisions are you intending to make based on this evaluation?
- Then you can determine what data you need to collect or have available
  - Primary data collection
  - Using secondary data sets

Stakeholders!

- And to whom is the evaluation meaningful?
  - Who cares about the result?
  - Who has a vested interest in what you find?
  - Brings us right back to where we started . . .
- Remember, that evaluation is a political exercise and must be undertaken and understood in that context
- Whether you designed this correctly at the beginning or inherited it after the fact . . .

CDC Framework for Program Evaluation

Standards for effective evaluation:

Utility: must serve the information needs of intended users

Feasibility: must be realistic, prudent, diplomatic and frugal

Propriety: must be legal, ethical and have regard for the welfare of those involved and affected

Accuracy: must reveal and convey technically correct information

CDC Framework for Program Evaluation

1. Engage stakeholders
2. Describe the program
3. Focus the evaluation design
4. Gather credible evidence
5. Justify conclusions
6. Ensure use and share lessons learned

Closing Words of Advice

- Good evaluations have a clear purpose
  - what is the decision you intend to make?
- And a clear plan for action
  - is there political will? or is there politics?
- Good evaluations begin with clear objectives
  - even if you inherit it after the fact, take the time to frame the purpose, intent and process

Closing Words of Advice

- Good evaluations have clear, valid measures
  - and valid/reliable data systems to collect them
- Good evaluations are based on a defensible logic model
  - that allows you to answer the question – did this make a difference
- Good evaluations are transparent

Closing Words of Advice

- And, of course, good evaluations are part and parcel of your overall *planning model*
  - They link back to needs and capacity
  - They facilitate efficiency
  - They promote accountability
  - They can help contribute to our overall knowledge base
    - (please publish your results in the MCH Journal ☺)

Thank you!

Good luck and God speed