Introduction to Program Evaluation— Using CDC's Evaluation Framework

Ву:

Thomas J. Chapel, MA, MBA
Chief Evaluation Officer
Centers for Disease Control and Prevention
TChapel@cdc.gov

404-639-2116

Pre-Institute Workshop: Sunday



Why We Evaluate...

"... The gods condemned Sisyphus to endlessly roll a rock up a hill, whence it would return each time to its starting place. They thought, with some reason...



Why We Evaluate...

...there was no punishment more severe than eternally futile labor...."

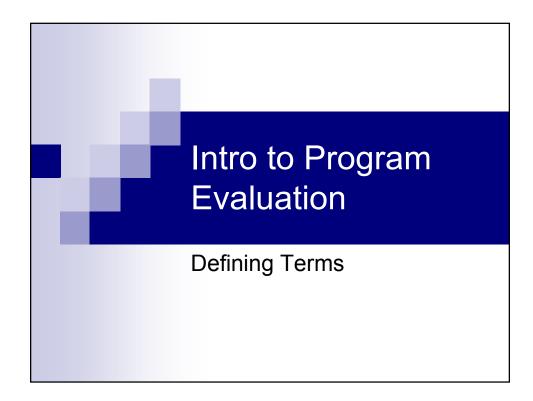
The Myth of Sisyphus

3



Today...

- CDC Evaluation Framework steps and standards
- Central role of "program description" and "evaluation focus" steps
- Create/use simple logic model(s) in evaluation
- Know/make informed decisions about design and data collection
- TIME PERMITTING: "Deep thoughts" about design

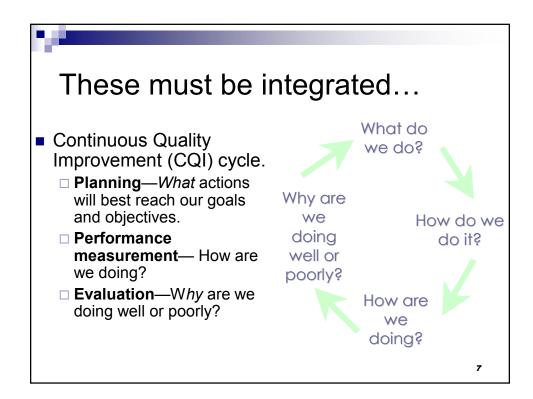


Defining Evaluation

Evaluation is the systematic investigation of the merit, worth, or significance of any "object"

Michael Scriven

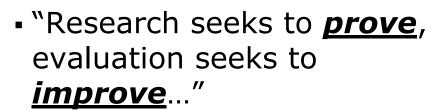
Program is any organized public health action/activity implemented to achieve some result





Research is...

 <u>Systematic</u> investigation, including research development, testing and evaluation, designed to develop or contribute to <u>generalizable</u> knowledge,



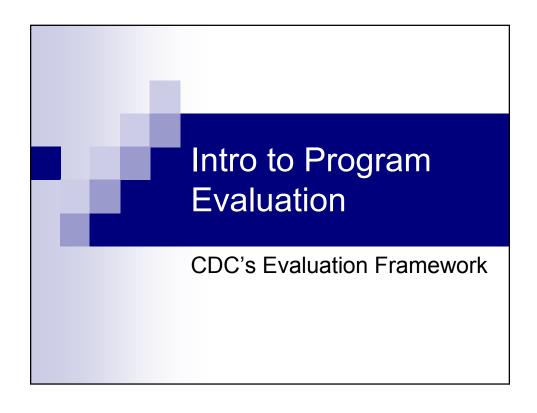
M.Q. Patton

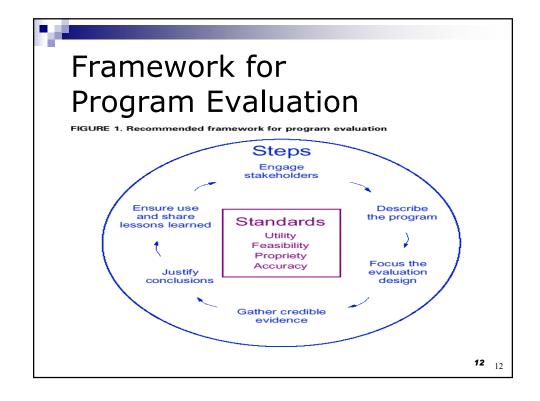
9

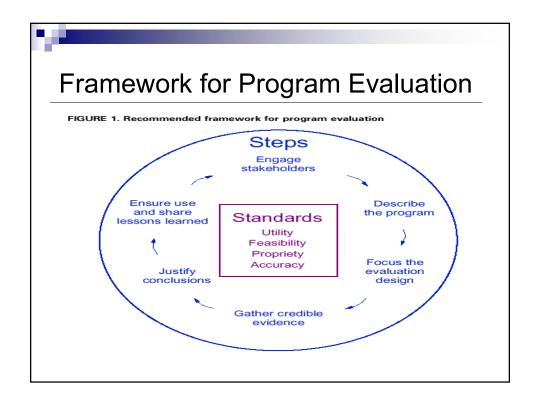


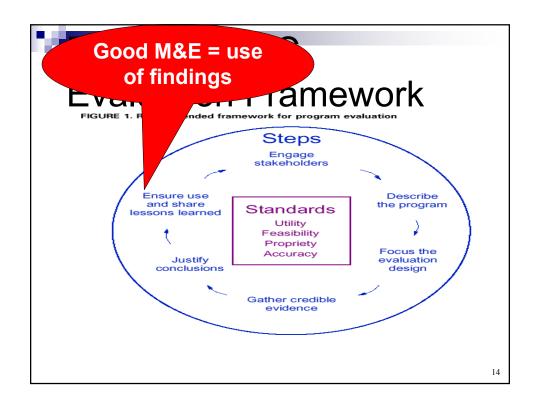
Surveillance...

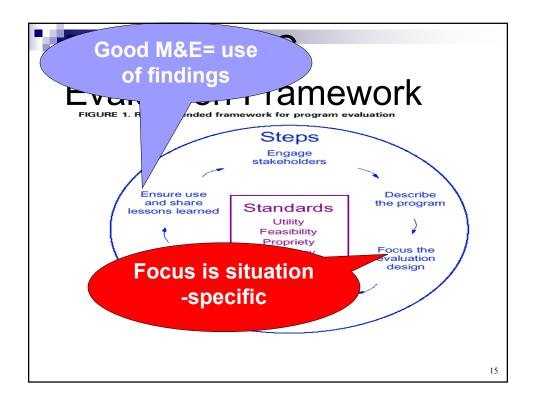
- Surveillance is continuous/routine data collection on various factors over regular intervals of time. Surveillance systems are:
 - □ <u>data source for program evaluation</u> especially of long-term and pop-based outcomes.
 - □ A resource for *formative* (*pre-implementation*) evaluation.

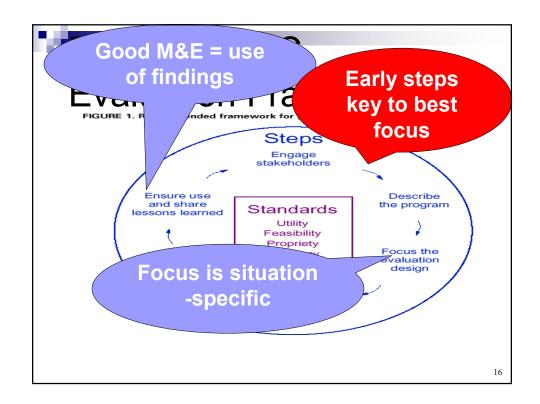














Step-by-Step

- Engage stakeholders: Decide who needs to be part of the design and implementation of the evaluation for it to make a difference.
- 2. <u>Describe the program</u>: Draw a "soup to nuts" picture of the program— activities and all intended outcomes.
- 3. **Focus the evaluation**: Decide which evaluation questions are the key ones

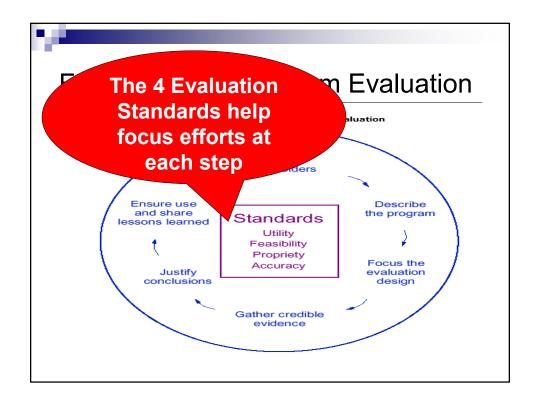
17



Step-by-Step

Seeds of Steps 1-3 harvested later:

- 4. **Gather credible evidence**: Write indicators and choose and implement data collection sources and methods
- 5. <u>Justify conclusions</u>: Review and interpret data/evidence to determine success of failure
- 6. <u>Use lessons learned</u>: Use evaluation results in a meaningful way.



٠,

The Four Standards

No one "right" evaluation. Instead, best choice at each step is options that maximize:

- <u>Utility</u>: Who needs the info from this evaluation and what info do they need?
- *Feasibility*: How much money, time, and effort can we put into this?
- **Propriety:** Who needs to be involved in the evaluation to be ethical?
- Accuracy: What design will lead to accurate information?

Intro to Program Evaluation

Step 2. Describing the Program



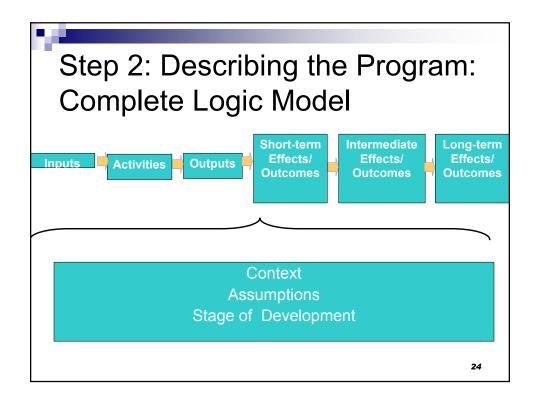
You Don't <u>Ever</u> Need a Logic Model, BUT, You <u>Always</u> Need a Program Description

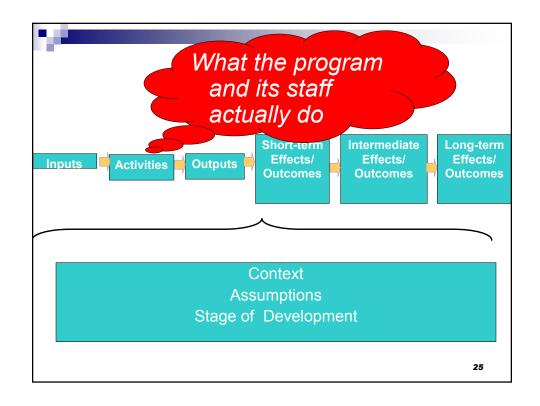
Don't jump into planning or eval without clarity on:

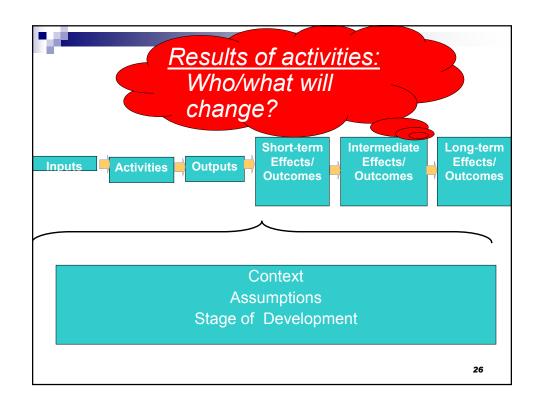
- The big <u>"need"</u> your program is to address
- The key <u>target group(s)</u> who need to take action
- The kinds of actions they need to take (your intended *outcomes* or objectives)
- Activities needed to meet those outcomes
- "Causal" <u>relationships</u> between activities and outcomes

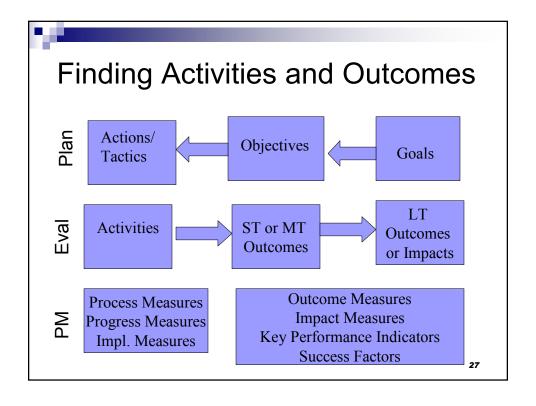
Logic Models and Program Description

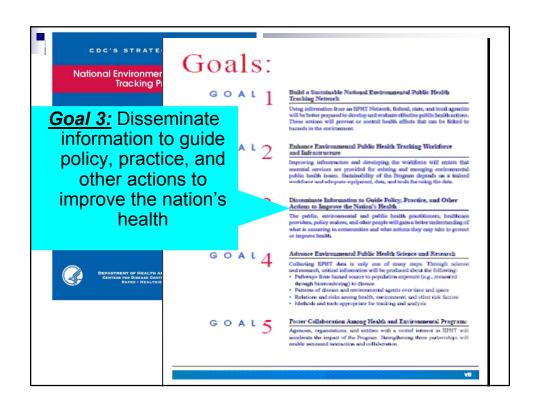
■ Logic Models: Graphic depictions of the relationship between your program's activities and its intended effects





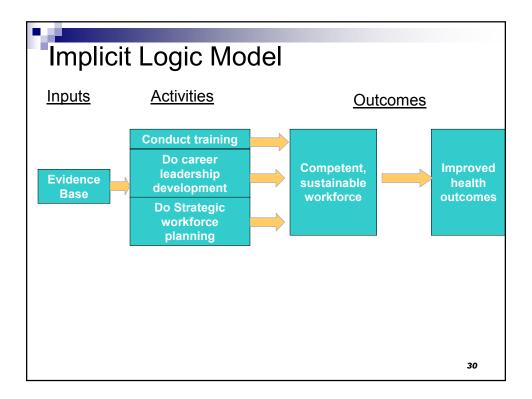


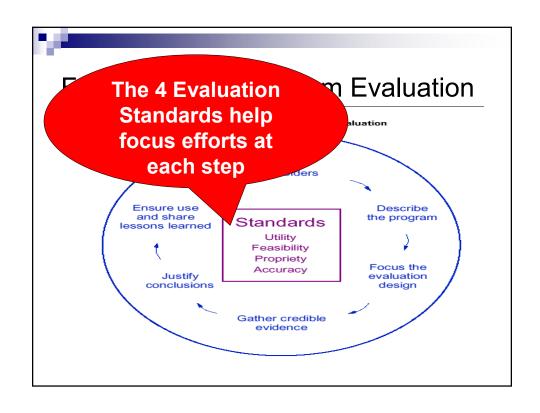


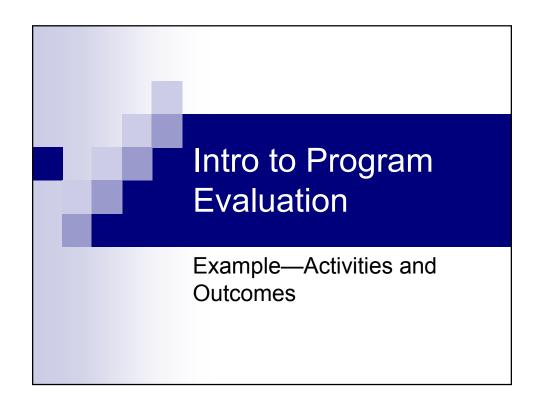


Finding Activities and Outcomes—OWCD Mission

To improve health outcomes by developing a competent, sustainable and diverse public health workforce through evidence-based training, career and leadership development, and strategic workforce planning.







Constructing Logic Models: Identify Activities and Outcomes by....

- Examining program descriptions, MISSIONS, VISIONS, PLANS, ETC and extracting these from the narrative, *OR*
- Reverse mapping
 —Starting with outcomes, ask "how to" in order to generate the activities which produce them, <u>OR</u>
- Forward mapping

 Starting with activities, ask "so what" in order to generate the outcomes that are expected to result

33



Then...Do Some Sequencing...

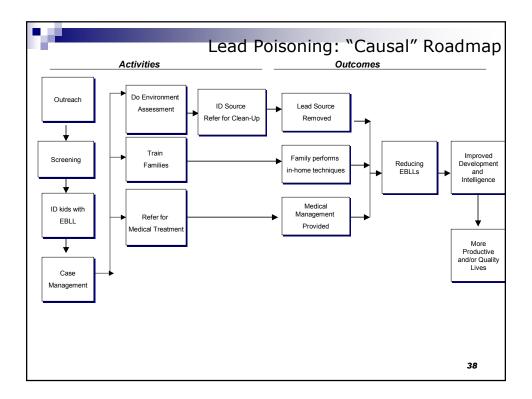
- Divide the *activities* into 2 or more columns based on their *logical* sequence. Which activities have to occur before other activities can occur?
- Do same with the *outcomes*. Which outcomes have to occur before other outcomes can occur?

Listing Activities and Outcomes: **Lead Poisoning** Activities Effects/Outcomes □ Outreach □ Lead source identified □ Screening □ *Families* adopt in-home □ Case management techniques □ **Referral** for medical tx □ **Providers** treats EBLL □ Identification of kids with kids elevated lead (EBLL) ☐ <u>Housing Authority</u> □ Environmental assessment eliminates lead source □ *Referral* for env clean-up □ EBLL reduced Family training □ Developmental "slide" stopped 35 □ Q of L improved

| Glob | al Logic Model: Childhoo | od Lead Poisoning Prog | ram |
|----------------------------|---|--|----------------------------|
| Early Activities If we do | Later Activities And we do | Early Outcomes | Later Outcomes And then |
| Outreach | Refer EBLL kids for medical treatment | EBLL kids get medical treatment | |
| Screening | | | EBLL reduced |
| ID of elevated kids | Train family in in- home techniques | Family performs in-home techniques | Develop'l slide stopped |
| Case mgmt of EBLL kids | Assess environment of | Lead source identified | Quality of life improves |
| | EBLL child | Environment gets cleaned up | |
| | Refer environment for clean-up | Lead source removed | |
| | , | | 36 |

For Planning and Evaluation "Causal" Arrows Can Help

- Not a different logic model, but same elements in different format
- Arrows can go from:
 - □ Activities to other activities: Which activities feed which other activities?
 - □ Activities to outcomes: Which activities produce which intended outcomes?
 - □ Early effects/outcomes to later ones: Which early outcomes produce which later outcomes



Note!

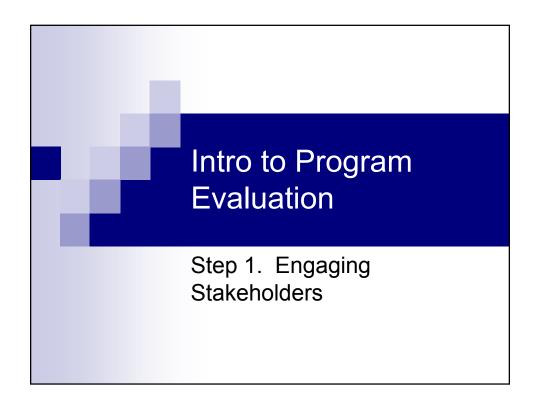
Logic Models make the program theory <u>clear</u>, not <u>true!</u>

39



Logic Models Take Time...So Be Sure to Use Them

- Not worth it as "ends in themselves"
- But can pay off big in evaluation:
 - □ Clarity with stakeholders
 - □ Setting evaluation focus



Which S'holders Matter Most?

- *Affected* by the program?
- *Involved* in program operations?
- Intended <u>users</u> of evaluation findings?

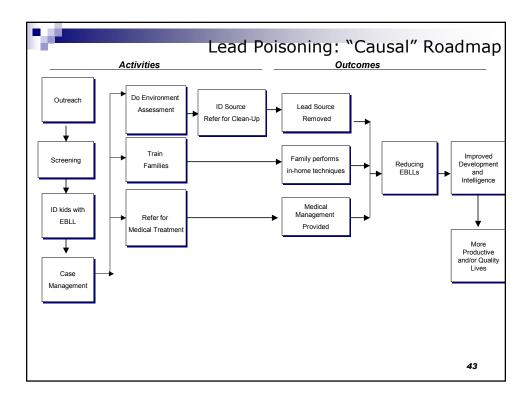
Of these, who do we most need to:

Enhance *credibility?*

Implement program changes?

Advocate for changes?

Fund, authorize, expand program?

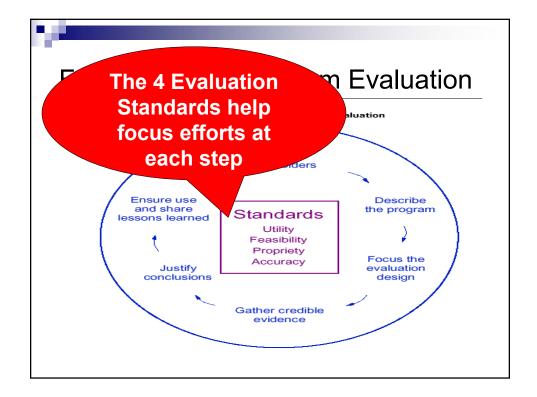




Using the Logic Model with Stakeholders

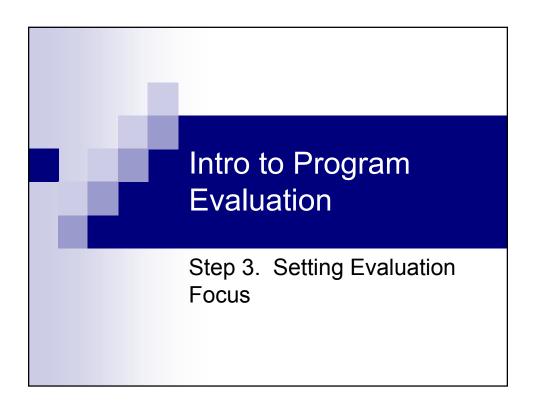
Do they agree/disagree with:

- The activities and outcomes depicted?
- The "roadmap"?
- Which outcomes = program "success"?
- How much progress on outcomes = program "success"?
- Choices of data collection/analysis methods?



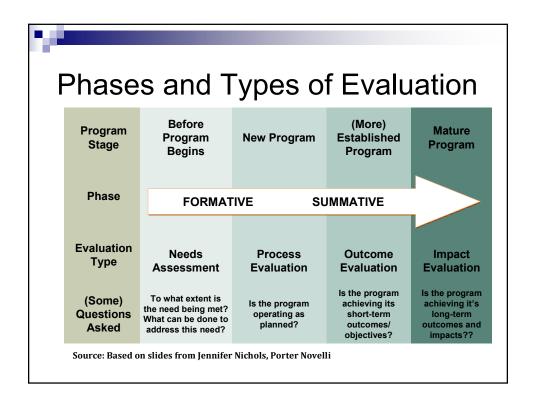
Case Exercise—Stakeholders

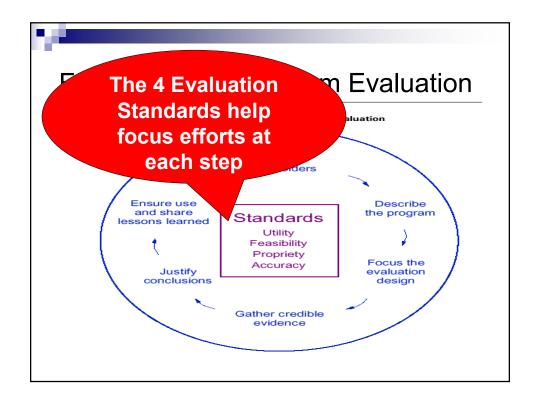
- We need [this stakeholder]...
- To provide/enhance our [any/all of: credibility, implementation, funding, advocacy]...
- And, to keep them engaged as the project progresses...
- We'll need to demonstrate [which selected activities or outcomes].



Evaluation Can Be About Anything

- Evaluation can focus on any/all parts of the logic model
- Evaluation questions can pertain to
 - □Boxes---did this component occur as expected
 - □Arrows---what was the relationship between components







Setting Focus: Some Rules

Based on "utility" standard:

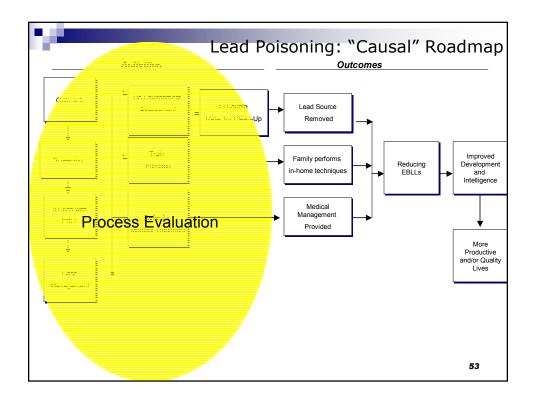
- Purpose: Toward what end is the evaluation being conducted?
- <u>User:</u> Who wants the info and what are they interested in?
- <u>Use:</u> How will they use the info?

51



(Some) Potential Purposes

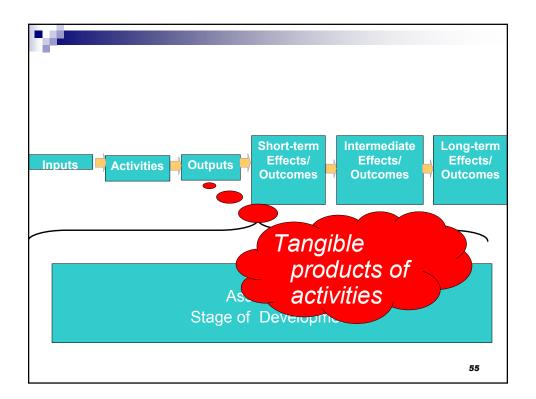
- Test program implementation
- Show accountability
- "Continuous" program improvement
- Increase the knowledge base
- Other...
- Other...





Process Evaluation

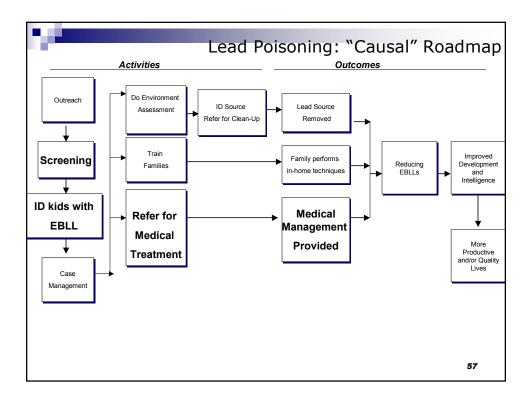
- ■The type and quantity of services provided
- ■The number of people receiving services
- What actually happens during implementation
- ■How much money the project costs
- ■The staffing for services/programs
- The number of coalition activities and meetings
- Assessment of program fidelity



•

Lead Poisoning: Sample Outputs

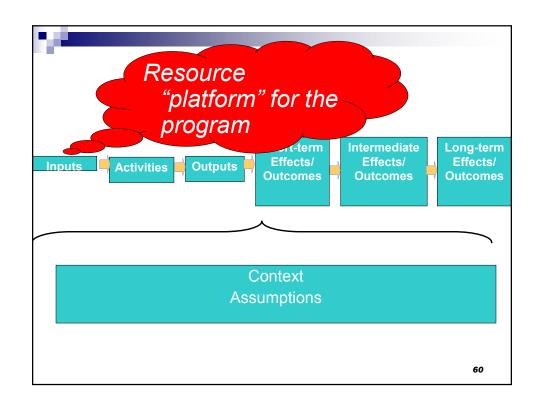
- Pool (#) of eligible kids
- Pool (#) of screened kids
- Referrals (#) to medical treatment
- Pool (#) of assessed homes
- Referrals (#) for clean-up



Lead Poisoning: "Upgraded" Outputs

- Pool (#) of screened kids (meeting likely risk profile)
- Pool (#) of eligible kids (with lead level >XXd/ul)
- Referrals (#) to (qualified or willing) medical treatment providers
- Pool (#) of assessed ("leaded") homes
- Referrals (#) for clean-up (to qualified or willing orgs)

| Global | Logic Model: Child | lhood Lead Poiso | ning Program | |
|---------------------------|--|---|--|--------------------------------|
| Early Activities | Later Activities | Outputs | Early Outcomes— | Later Outcomes |
| Outreach | | (#) of eligible kids meeting risk profile | EBLL kids get medical treatment | EBLL reduced |
| Screening | | (#) screened kids with lead < threshold | | Develop'l |
| ID of elevated kids | Refer for medical tx | (#) referrals to qualified medical tx | Family performs in- home techniques | slide stopped Quality of |
| Do case mgmt | Train family in in-home techniques | (#) of families completing training | Lead source identified | life improves |
| | Assess environ't | (#) of "leaded" homes | Environ cleaned up | |
| | Refer house for clean-up | (#) referrals to qualified clean-up | Lead source removed | 59 |





Lead Poisoning: Sample Inputs

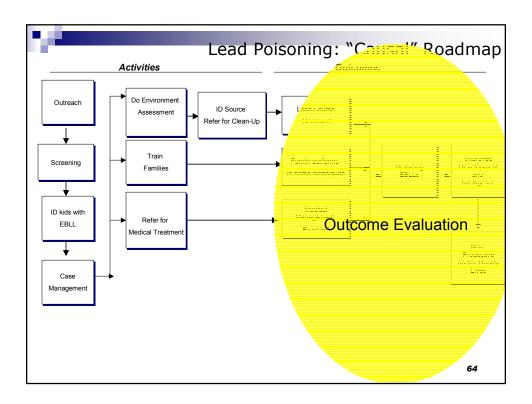
- Funds
- Trained staff
- Legal authority to screen
- Relationships with orgs for med tx and env cleanup

| | Global Logic | Model: Childho | ood Lead Poisor | ning Program | |
|------------------------------------|---------------------------|------------------------------------|---|---------------------------------------|--------------------------------|
| Inputs | Early Activities | Later Activities | Outputs | Early Outcomes— | Later Outcomes |
| Funds Trained staff | Outreach Screening | | (#) of eligible kids meeting risk profile (#) screened | EBLL kids get medical treatment | EBLL reduced |
| R'ships with orgs for med tx | ID of elevated kids | Refer for medical treatment | kids with lead < threshold (#) referrals to qualified | Family performs inhome techniques | Develop'l slide stopped |
| and clean up Legal | Do case mgmt | Train family in in-home techniques | medical tx (#) of families completing training | Lead source identified | Quality of life improves |
| authority | | Assess environ't | (#) of "leaded" homes | Environ cleaned up | |
| | | Refer house for clean-up | (#) referrals to qualified clean-up | Lead source removed | 62 |



(Some) Potential Purposes

- Test program implementation
- Show accountability
- "Continuous" program improvement
- Increase the knowledge base
- Other...
- Other...





Outcome Evaluation

- ■Results of program services
- Changes in individuals
 - □ Knowledge/awareness
 - □ Attitudes
 - □Beliefs
- Changes in the environment
- Changes in behaviors
- Changes in disease trend

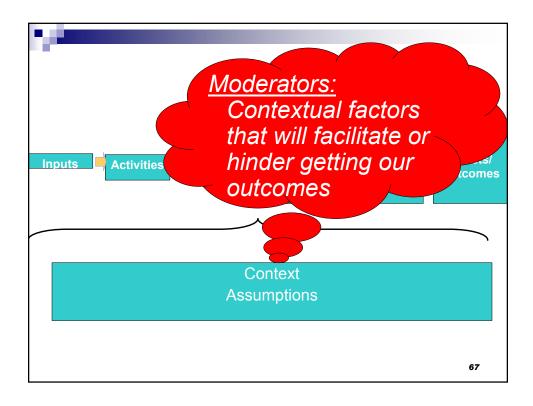
65



"Reality Checking" the Focus

Based on "feasibility" standard:

- Stage of Development: How long has the program been in existence?
- Program Intensity: How intense is the program? How much impact is reasonable to expect?
- Resources: How much time, money, expertise are available?

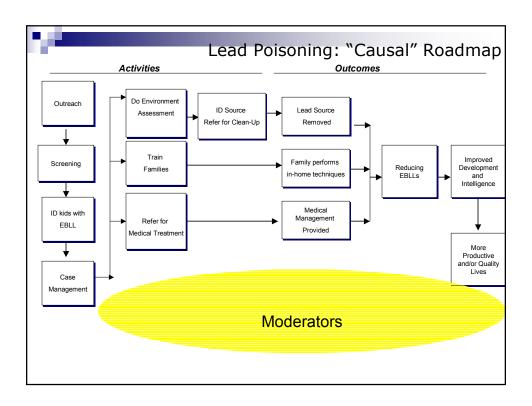


Moderators/Contextual Factors

- <u>P</u>olitical
- <u>E</u>conomic
- <u>S</u>ocial
- <u>T</u>echnological

Moderators—Lead Poisoning

- Political—"Hazard" politics
- <u>E</u>conomic— Health insurance
- <u>T</u>echnological— Availability of hand-held technology





Some Evaluation Scenarios

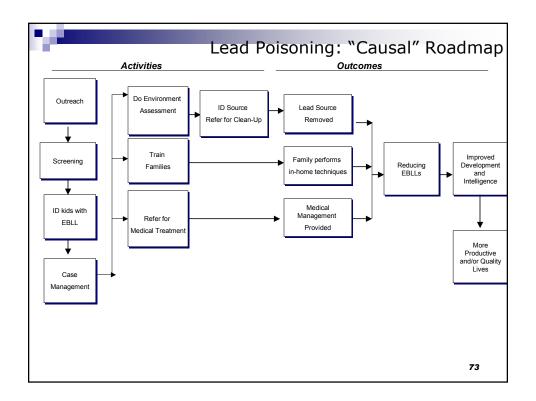
Scenario I: At Year 1, other communities want to adopt your model but want to know "what are they in for"

71



Scenario 1:

- Purpose: Examine program implementation
- *User:* The "other community"
- Use: To make a determination, based on your experience, whether they want to adopt this project or not



Some Evaluation Scenarios

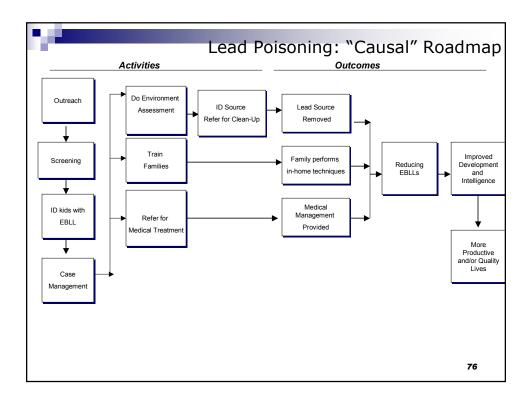
Scenario II: At Year 5, declining state revenues mean you need to justify to legislators the importance of your efforts so as to continue funds.

So

Scenario 2:

<u>Purpose:</u> Determine program impact <u>User:</u> Your org and/or the legislators **Use:**

- □ You want to muster evidence to prove to legislators you are effective enough to warrant funding, or
- □<u>Legislators</u> want you to show evidence that proves sufficient effectiveness to warrant funding





Steps 4-5. Gather Credible Evidence and Justify Conclusions

What is an indicator?

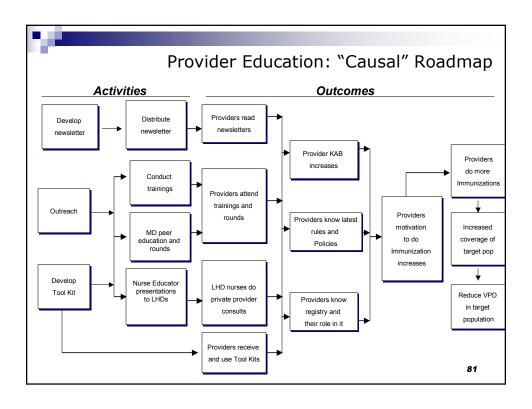
Specific, observable, and measurable characteristics that show progress towards a specified activity or outcome.

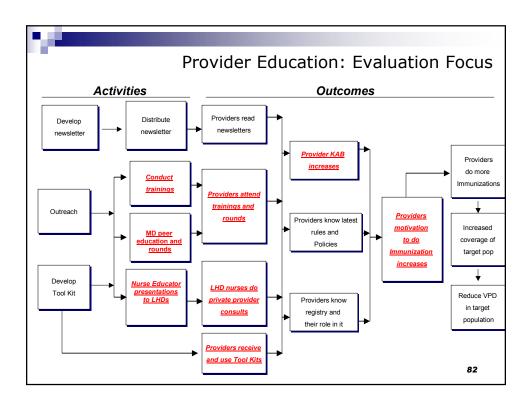


Selecting Indicators

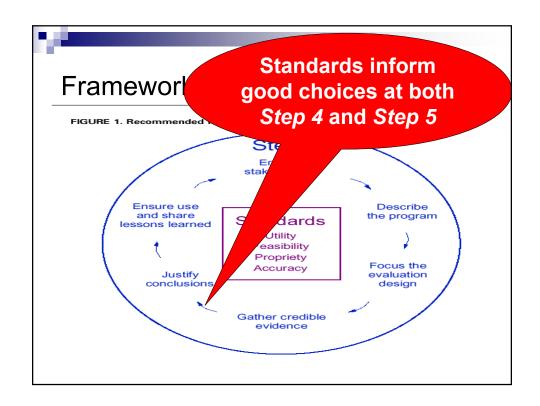
- Focused and measure an important dimension of the activity or outcome
- Clear and specific in terms of what it will measure
 - □ NOT components of the activity/outcome
 - □ NOT "fruits" of the activity/outcome
- At least one indicator for each activity or outcome of interest; but may need multiple ones

| Global Logic Model: Provider Education | | | |
|--|----------------------------|--------------------------|-----------------------|
| Early Activities | Later Activities | Early Outcomes | Later Outcomes |
| Do outreach to | Distribute newsletter | Provs read newsletters | KAB increases |
| providers | Conduct | Provs attend | Know policies |
| Develop newsletter | immuno trainings | trainings and rounds | Know registry |
| Develop | Nurse Develop educator LHD | Provs receive and | Motivation increases |
| Tool Kit | presentations | use tool kits | Do more immuno |
| | Physician peer ed | LHD nurses do private | |
| | rounds | prov consults | Coverage increases |
| | | | VPD reduced |





easurement Table: Scenarios 1-2: Provider Education Program **Eval Focus Components** Conduct immuno trainings A series of 3 trainings will be conducted in all 4 regions of the state Nurse educator LHD presentations Nurse educators will make presentations to 10 largest LHDs Physician peer ed rounds Physicians will host peer ed rounds at 10 largest hospitals Trainings will be well-attended and reflect good Provs attend trainings and rounds mix of specialties and geog representation Provs receive and use tool kits 50%+ of providers who receive tool kit will report use of it (or, "call-to-action" cards will be received from 25% of all providers receiving toolkit LHD nurses do private prov Trained nurses in LHDs will do provider consults with largest provider practices in county consults KAB increases Providers show increases in KAB on key immunization items such as [THESE] Motivation increases Provider intent to immunize increases 83





Not "Collect Data", BUT "Gather Credible Evidence"

Narrowing from 100s of ways to collect data:

- Utility: Who's going to use the data and for what?
- Feasibility: How much resources?
- Propriety: Ethical constraints?
- Accuracy: How "accurate" do data need to be?

Q.



Not "Analyze Data", BUT "Justify Conclusions"

- Utility: Who's going to use the data and for what?
- Feasibility: How much resources?
- Propriety: Ethical constraints? What does "ethical" mean?
- Accuracy: How "accurate" do we need to be? What does "accurate" mean?



Characterizing Data and Methods

Data can be:

Nominal vs. ordinal vs. interval

Data collection/methods can be:

- Primary vs. secondary
- Obtrusive vs. unobtrusive
- Quantitative vs. qualitative



These Ways to Gather Evidence...

- Written survey
- Personal interview
 - □ individual, group
 - structured, semi-structured, conversational
- Observation
- Document analysis
- Case study
- Group assessment
 - brainstorming, delphi, nominal group, fishbowl
 - □ Role play, dramatization
- Expert or peer review
- Portfolio review
- Consensus modeling

- Testimonials
- Perception tests
- Hypothetical scenarios
- Storytelling
- Geographical mapping
- Concept mapping
- Freelisting
- Sociograms
- Debriefing sessions
- Cost accounting
- Photography, drawing, art, videography
- Diaries/journals
- Logs, activity forms, registries



Cluster Into These Six Categories...

- Surveys
- Interviews
- Focus groups
- Document review
- Observation
- Secondary data analysis

89



Choosing Data Collection Methods

- Function of *context*:
 - □Time
 - □ Cost
 - □ Ethics
- Function of *content* to be measured:
 - ☐ Sensitivity of the issue
 - □ "Hawthorne effect"
 - □Validity
 - □ Reliability

Choosing Methods—Cross-Walk to Eval Standards

- Function of *context*:
 - ☐ Time [FEASIBILITY]
 - □ Cost [FEASIBILITY]
 - □ Ethics [PROPRIETY]
- Function of *content* to be measured:
 - ☐ Sensitivity of the issue [ALL]
 - ☐ "Hawthorne effect" [ACCURACY]
 - □ Validity [ACCURACY]
 - □ Reliability [ACCURACY]

91 91

Trade-offs of Different Data Collection Methods

| Method/Factor | Time | Cost | Sensitive Issues | Hawthorne Effect | Ethics |
|--------------------|------|------|---------------------|---------------------|--------|
| Survey: Mail | | | | | |
| Personal Interview | | | | | |
| Focus Groups | | | | | |
| Document Review | | | | | |
| Survey: Phone | | | | | |
| Observation | | | | | |
| Secondary Data | | | | | |



Examples—What's Best/Worst Method?

- Point-in-time estimate—sexual behavior of high school males
- Understanding context—intimate partner violence
- Adoption of housekeeping and nutrition behaviors to reduce lead burden



Quantitative and Qualitative

- Quantitative methods... produce data that can be counted or expressed numerically
- Qualitative methods... produce data that do not indicate ordinal (or beyond) values

Source: Adapted from Nkwi, Nyamongo & Ryan

Data Collection Methods Overlap "Types"

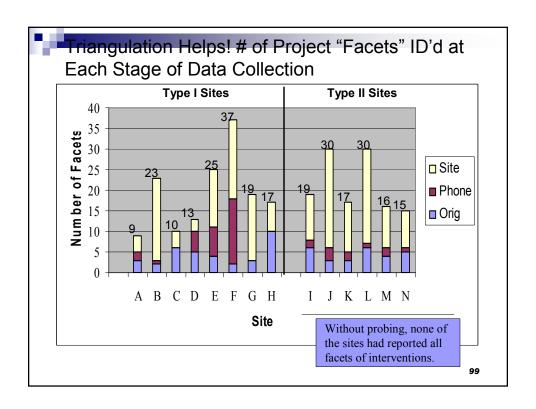
| Method | Quantitative | Qualitative |
|--------------------|--------------|-------------|
| Survey: Mail | | |
| Personal Interview | | |
| Focus Groups | | |
| Document Review | | |
| Survey: Phone | | |
| Observation | | |
| Secondary Data | | |

95 95

Less Structure Participant Observation **Focus Groups IDIs Flexibility Precision QL Diaries Inductive Good existing** power Free Lists data **EDM Ability to Pile Sorts Comparability** capture unknown Surveys **Direct Observation** QT Diaries How does your situation fit? **More Structure** Source: Gregory Guest, PhD

| Measurement Table: Scenarios 1-2 : Provider Education Program | | |
|---|---|--|
| Eval Focus Components | Indicators | |
| Conduct immuno trainings | A series of 3 trainings will be conducted in all 4 regions of the state | |
| Nurse educator LHD presentations | Nurse educators will make presentations to 10 largest LHDs | |
| Physician peer ed rounds | Physicians will host peer ed rounds at 10 largest hospitals | |
| Provs attend trainings and rounds | Trainings will be well-attended and reflect good mix of specialties and geog representation | |
| Provs receive and use tool kits | 50%+ of providers who receive tool kit will report use of it (or, "call-to-action" cards will be received from 25% of all providers receiving toolkit | |
| LHD nurses do private prov consults | Trained nurses in LHDs will do provider consults with largest provider practices in county | |
| KAB increases | Providers show increases in KAB on key immunization items such as [THESE] | |
| Motivation increases | Provider intent to immunize increases 97 | |

| Measurement Table : Provider Education Program | | |
|---|---|--|
| Indicators | Methods/Sources | |
| A series of 3 trainings will be conducted in all 4 regions of the state | Training logs | |
| Nurse educators will make presentations to 10 largest LHDs | Training logs | |
| Physicians will host peer ed rounds at 10 largest hospitals | Training logs | |
| Trainings will be well-attended and reflect good mix of specialties and geog representation | Registration info | |
| 50%+ of providers who receive tool kit will report use of it (or, "call-to-action" cards will be received from 25% of all providers receiving toolkit | Survey of providers Analysis/count of call-to- action cards | |
| Trained nurses in LHDs will do provider consults with largest provider practices in county | Survey of nurses, survey or providers, or training logs | |
| Providers show increases in KAB on key immunization items such as [THESE] | Survey of providers, or focus groups, or intercepts | |
| Provider intent to immunize increases | Same 98 | |



On "Justifying Conclusions"

"It is not the facts that are of chief importance, but the light thrown upon them, the meaning in which they are dressed, the conclusions which are drawn from them, and the judgments delivered upon them."

– Mark Twain



Step 5: Justifying Conclusions

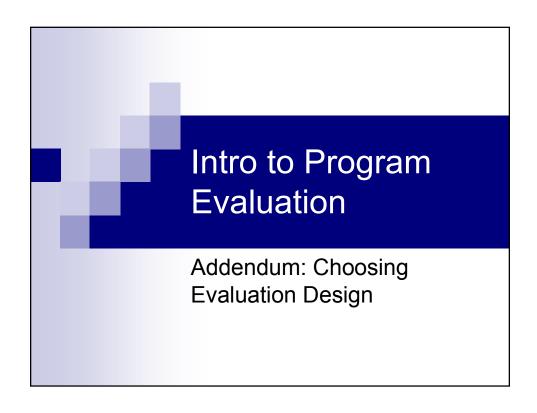
- Analyzing and synthesizing data are key stepe now
- BUT REMEMBER: "Objective data" are interpreted through a prism of stakeholder "values"
- Seeds planted in Step 1 are harvested now. What did we learn in stakeholder engagement that may inform what we analyze and how?

101



Reminder: Some Prisms

- Cost and cost-benefit
- Efficiency of delivery of services
- Health disparities reduction
- Population-based impact, not just impact on those participating in the intervention
- Causal attribution
- "Zero-defects"



Thinking About Cause: Evaluation Design Continuum

| Non | Quasi | |
|--------------|--------------|--------------|
| Experimental | Experimental | Experimental |
| Weakest | Stronger | Strongest |



Requirement

Implications

Experimental and control conditions

Must be at least two groups: One that gets the program, one that does not

Single experimental condition

Must be only one activity or program that distinguishes the experimental and

control conditions

Random assignment to conditions

Participants are just as likely to be assigned to experimental condition as to

the control condition

Pre- and post-program measurements

At a minimum, measures are taken from people in both conditions before the program begins and after it is over

105 105



"Classic" Experimental Design

RE

 O_1

X

 O_2

RC

 O_3

 O_4

Where:

R= Random assignment

E=Experimental group

C=comparison group

O=Observation/Data Collection

Design Continuum:

What's Missing as Move Right →Left? (Why) Does It Matter?

| Non Experimental | Quasi Experimental | Experimental |
|---------------------|-----------------------|------------------|
| E: X O | X O | X O (R) |
| C: | O | O (R) |
| E: O X O | O X O | O X O (R) |
| C: | O O | O O (R) |
| E: C: | 0 0 0 X 0 0 0 | O O O X O O O(R) |



Group Exercise: Choosing Design

- What might an experimental design look like?
 - How close can you come?
 - What do you have to compromise?
 - (Why) does it matter?



Experimental Model as Gold Standard

- But, sometimes "fool's gold"
 - □ Internal validity vs. external validity
 - □ Community interventions
- So
 - □ Sometimes → "Right", but hard to implement
 - □ Sometimes → Easy to implement, but "wrong"

109 109



Beyond the Scientific Research Paradigm

• Complex programs and community initiatives

And since these initiatives are based on multisource and multi-perspective community
collaborations, their goals and core
activities/services are constantly changing and
evolving to meet the needs and priorities of a
variety of community stakeholders. In short, these
initiatives are "unevaluatable" using the dominant
natural science paradigm (Connell, et. al., 1995)



Other Ways to Justify...

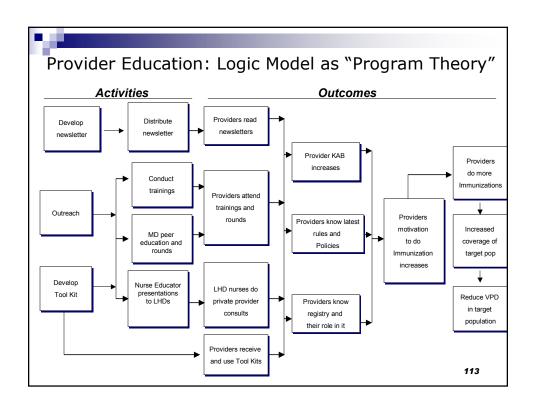
- Proximity in time
- Accounting for/eliminating alternative explanations
- Similar effects observed in similar contexts
- Plausible mechanisms/program theory

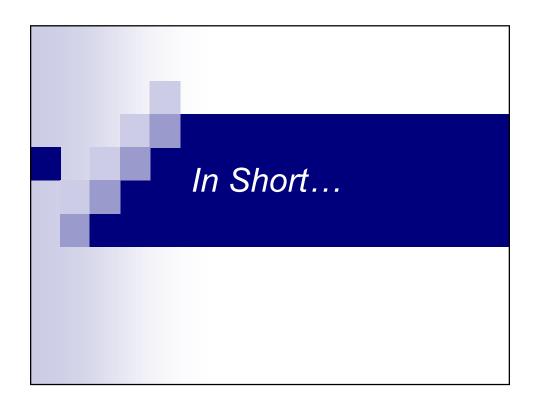
111 111



Provider Ed: "Proving Higher Coverage is "Due to Us"

- Proximity in time
- Accounting for/eliminating alternative explanations
- Similar effects observed in similar contexts
- Plausible mechanisms/program theory







Upfront Small Investment...

- Clarified relationship of activities and outcomes
- Ensured clarity and consensus with stakeholders
- Helped define the right focus for my evaluation
- Framed choices of indicators and data sources

115



Where Next....

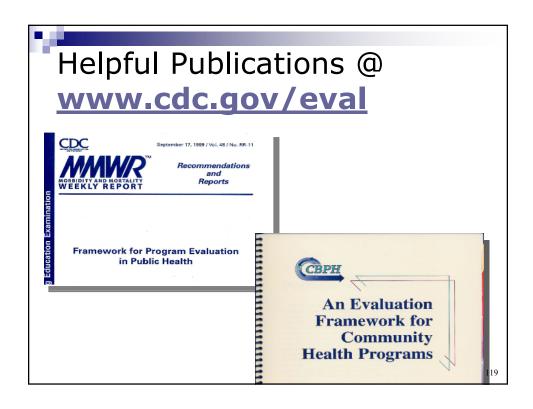
- Finalize indicators and data sources for questions
- Analyze data
- Draw conclusions and results
- Turn results into action

But...

■Better progress on these later steps because of the upfront work on Steps 1-3!!!

117

Intro to Program Evaluation Life Post-Session



Helpful Resources

- NEW! Intro to Program Evaluation for PH Programs—A Self-Study Guide:
 - http://www.cdc.gov/eval/whatsnew.htm
- Logic Model Sites
 - □ Innovation Network:
 - □ http://www.innonet.org/
 - □ W.K. Kellogg Foundation Evaluation Resources: <u>http://www.wkkf.org/programming/overview.aspx?CI</u>

 D=281
 - University of Wisconsin-Extension: http://www.uwex.edu/ces/Imcourse/
- Texts
 - □ Rogers et al. Program Theory in Evaluation. New Directions Series: Jossey-Bass, Fall 2000
 - □ Chen, H. Theory-Driven Evaluations. Sage. 1990 120

Community Tool Box http://ctb.ku.edu

