ABDUL LATIF JAMEEL
Poverty Action Lab

TRANSLATING RESEARCH INTO ACTION

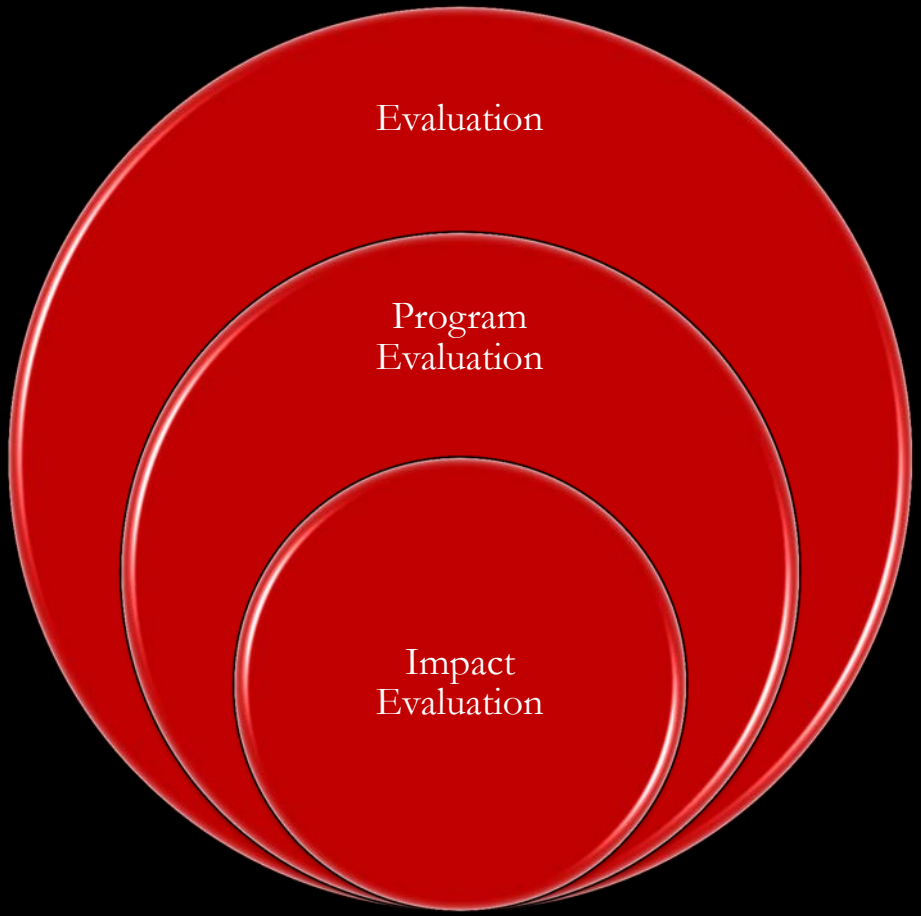
What is Evaluation?

Marc Shotland
J-PAL Global

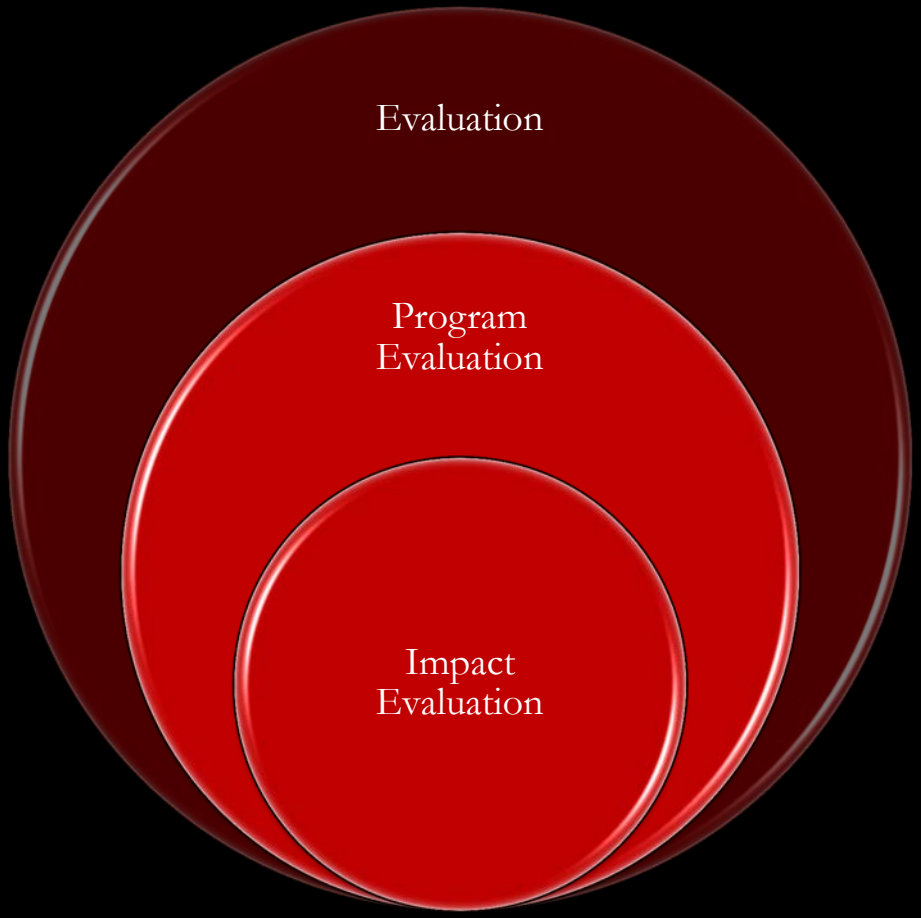
Course Overview

1. What is evaluation?
2. Measuring impacts (outcomes, indicators)
3. Why randomize?
4. How to randomize?
5. Sampling and sample size
6. Threats and Analysis
7. RCT: Start to Finish
8. Cost Effectiveness Analysis and Scaling Up

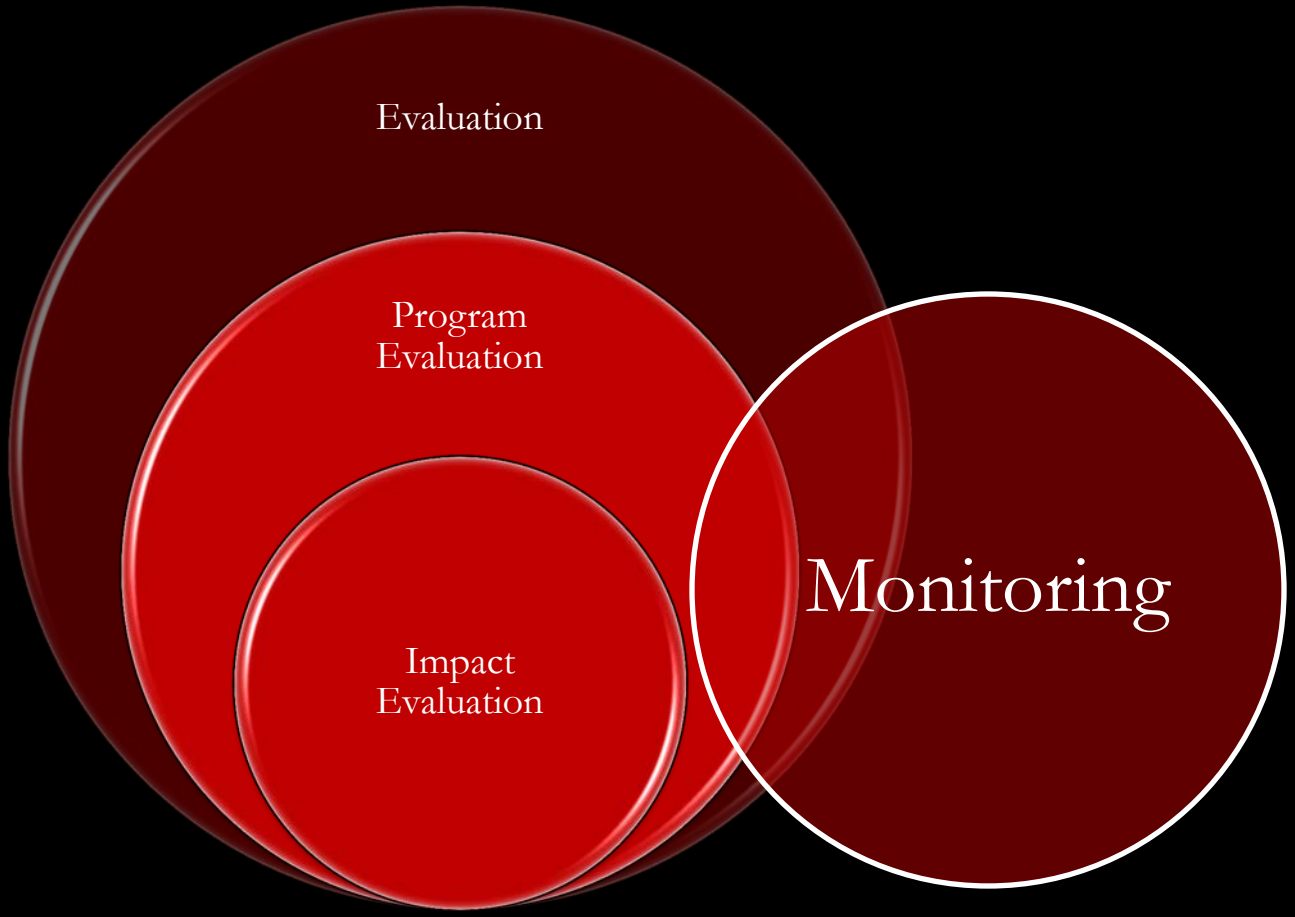
What is Evaluation?



Program Evaluation

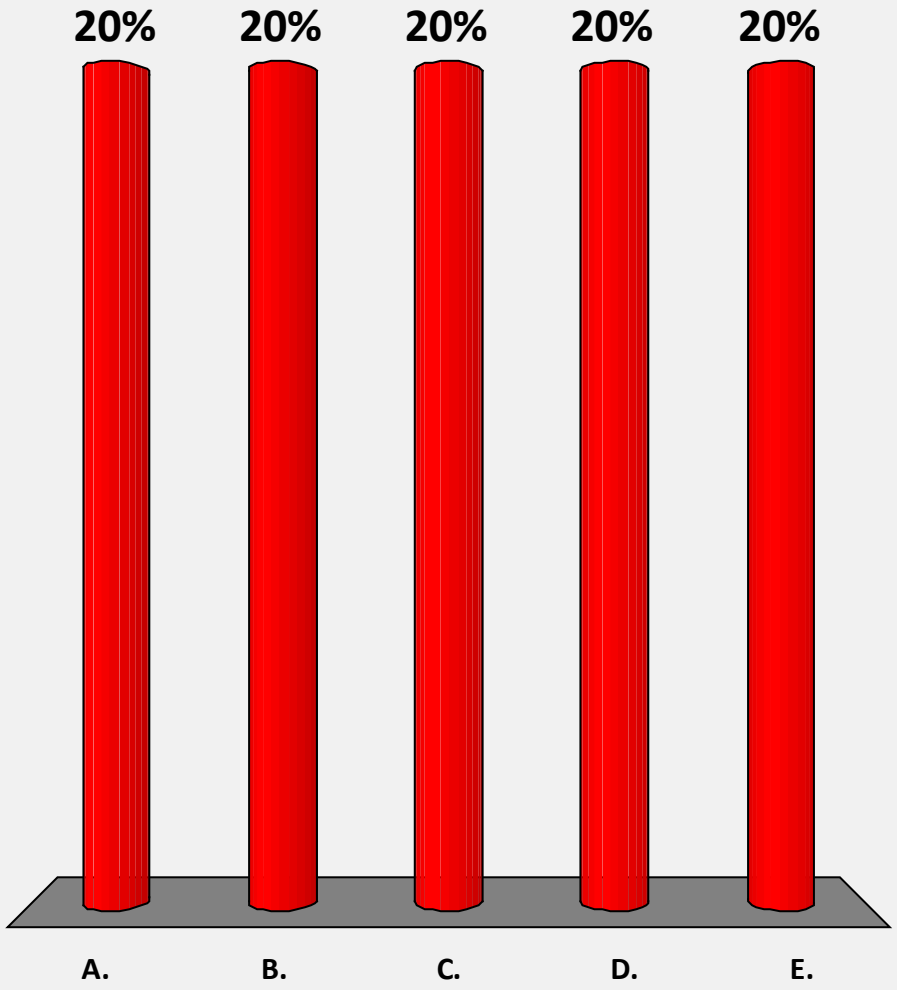


Monitoring and Evaluation

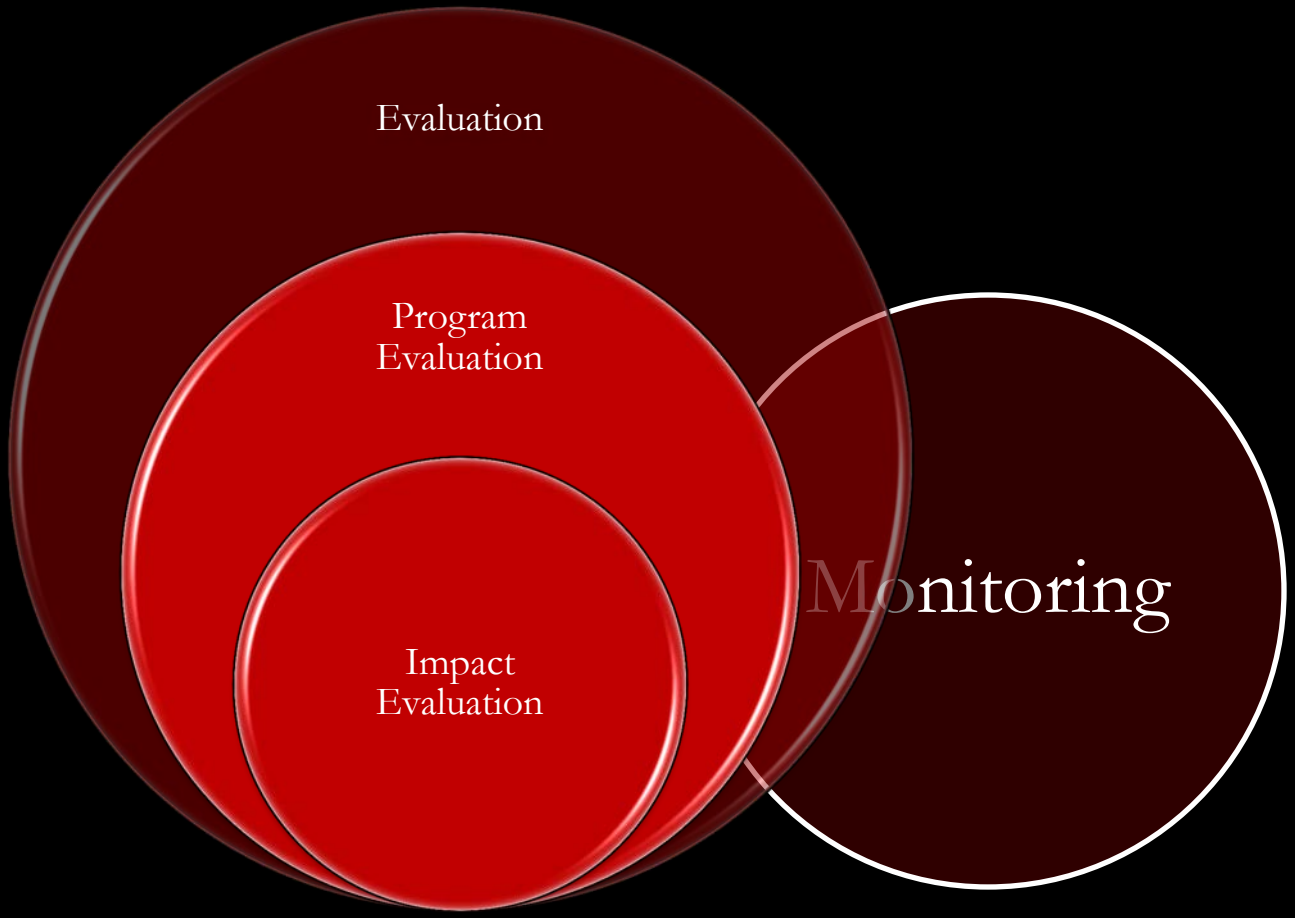


What's the difference between: Monitoring and Evaluation

A. Nothing. They are different words to describe the same activity
B. Monitoring is conducted internally, Evaluation is conducted externally
C. Monitoring is for management, Evaluation is for accountability
D. Don't know
E. Other



Program Evaluation



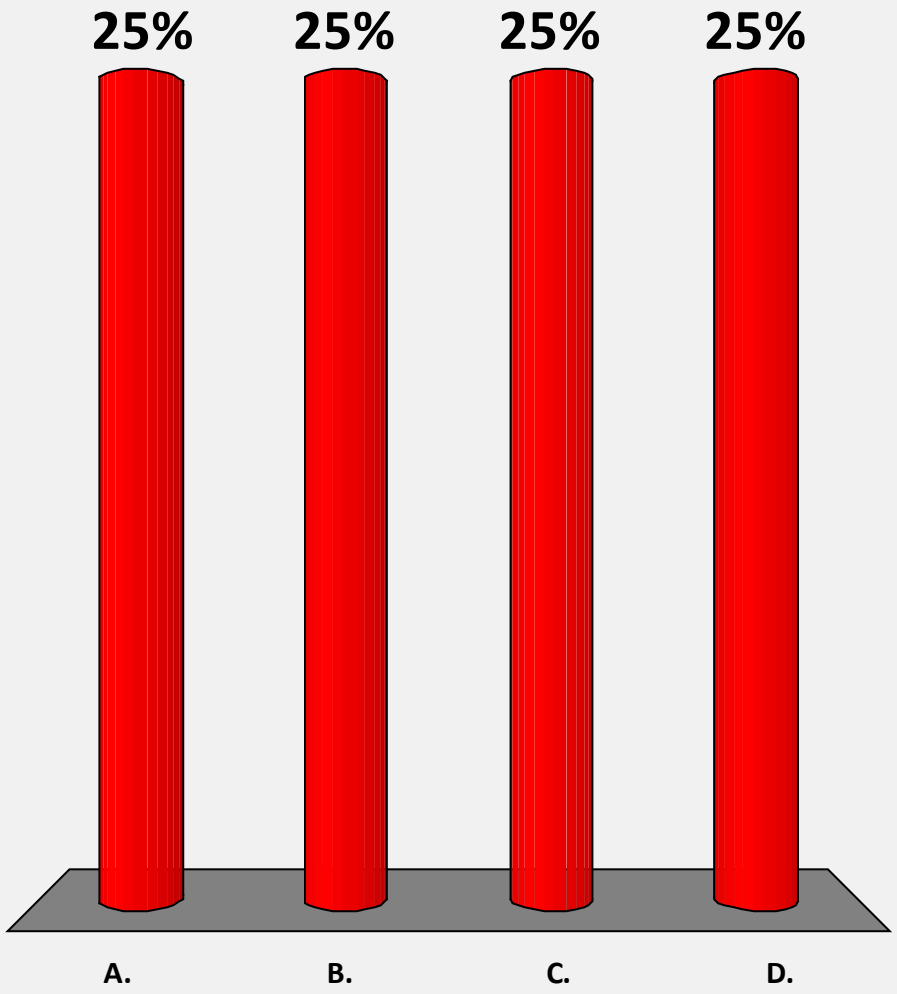
Components of Program Evaluation

- Needs Assessment
- Program Theory Assessment
- Process Evaluation
- Impact Evaluation
- Cost Effectiveness

- What is the problem?
- How, in theory, does the program fix the problem?
- Does the program work as planned?
- Were its goals achieved? The magnitude?
- Given magnitude and cost, how does it compare to alternatives?

Evaluation should usually be conducted:

A. Externally and independent from the implementers of the program being evaluated
B. Externally and closely integrated with program implementers
C. Internally
D. Don't know

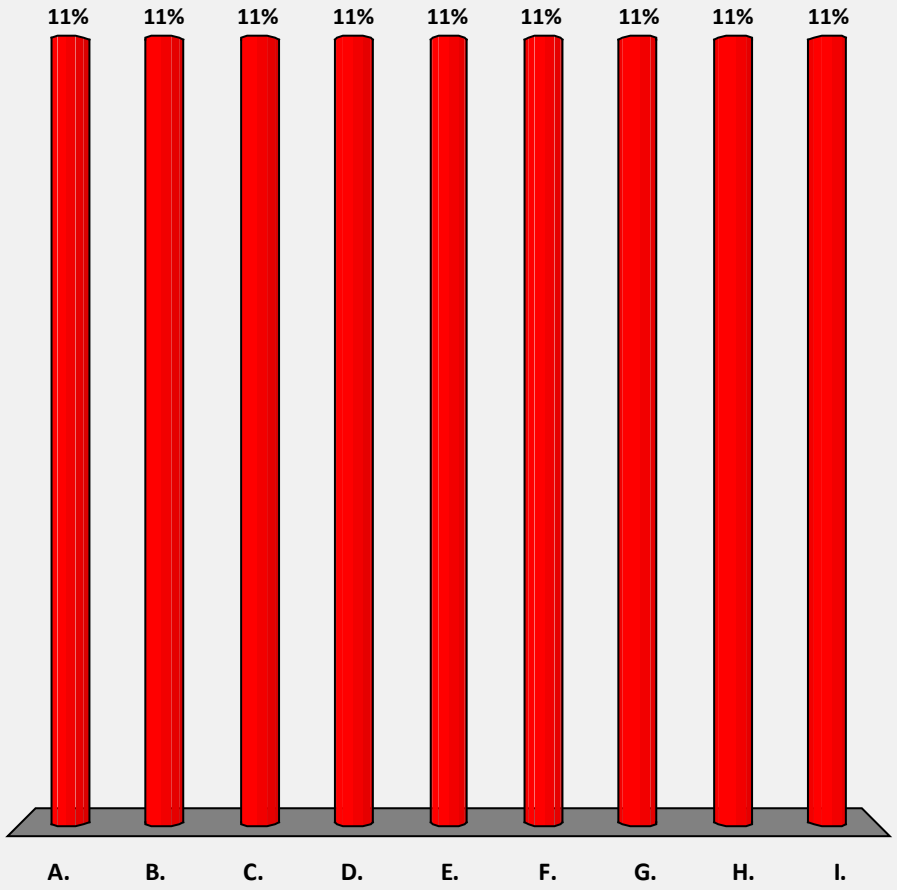


Who is this evaluation for?

- Academics
- Donors
  - Their Constituents
- Politicians / policymakers
- Technocrats
- Implementers
- Proponents, Skeptics
- Beneficiaries

Who is your ***most important audience*** for evaluation?

A. Agency leadership
B. Donor Politicians / policymakers
C. Donor Constituents
D. Academics
E. Indonesian Politicians / policymakers
F. Technocrats
G. Implementers
H. Proponents, Skeptics
I. Beneficiaries



Does Aid Work?

Aid Optimists

"I have identified the specific investments that are needed [to end poverty]; found ways to plan and implement them; [and] shown that they can be affordable."

Jeffrey Sachs
End of Poverty

Aid Pessimists

"After $2.3 trillion over 5 decades, why are the desperate needs of the world's poor still so tragically unmet?

Isn't it finally time for an end to the impunity of foreign aid?"

Bill Easterly
The White Man's Burden

Where does J-PAL fall in this debate?

- Somewhere in the middle
- We believe aid *can* help
- Too often it does not
- We don't know when it does and when it doesn't
- So we have a lot to learn about *how* Aid can help

How can impact evaluation help us?

- Surprisingly little hard evidence on what works
- Can do more with given budget with better evidence
- If people knew money was going to programs that worked, could help increase pot for anti-poverty programs
- Instead of asking "do aid/development programs work?" should be asking:
  - Which programs work best, why and when?
  - Which concepts work, why and when?
  - How can we scale up what works?
- Add to our body of evidence
  - part of a well-thought out evaluation (research) strategy

Programs and their Evaluations: where do we start?

Intervention

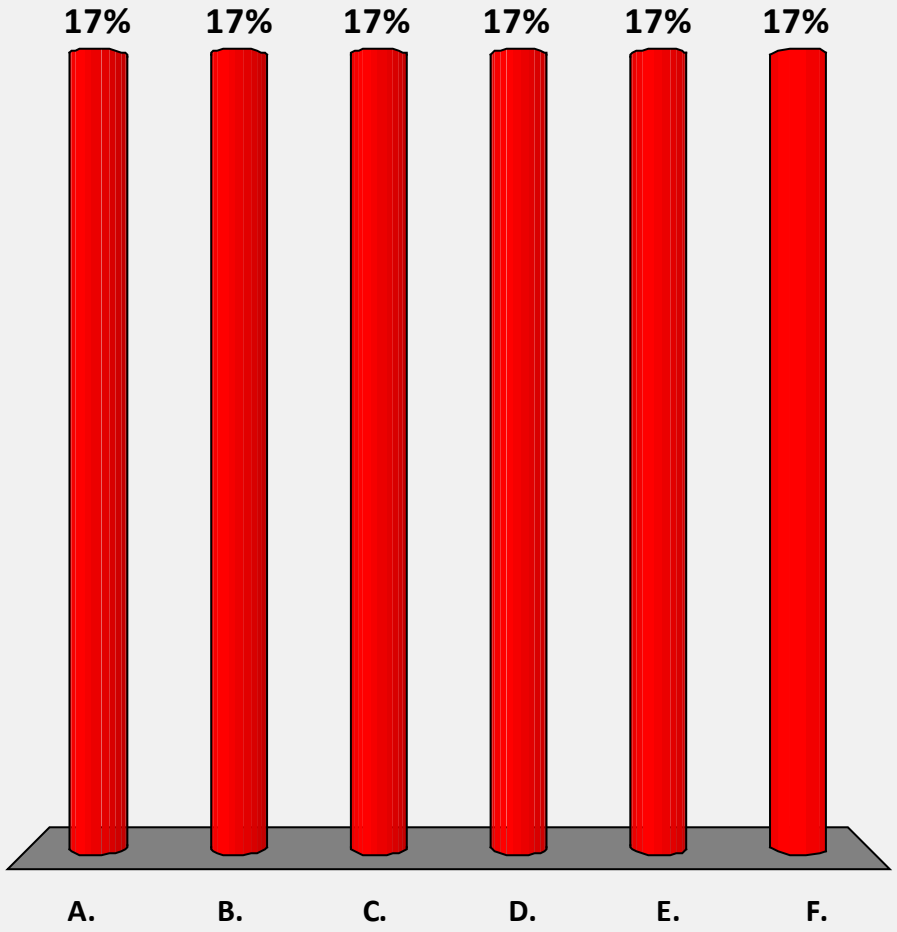
- Start with a problem
- Verify that the problem actually exists
- Generate a theory of why the problem exists
- Design the program
- Think about whether the solution is cost effective

Program Evaluation

- Start with a question
- Verify the question hasn't been answered
- State a hypothesis
- Design the evaluation
- Determine whether the value of the answer is worth the cost of the evaluation

What do you think is the most cost-effective way to reduce diarrhea?

A. Develop piped water infrastructure
B. Improve existing water sources
C. Increase supply of and demand for chlorine
D. Education on sanitation and health
E. Improved cooking stoves for boiling water
F. Improve sanitation infrastructure



Identifying the problem

NEEDS ASSESSMENT

The Need

- Nearly 2 million children die each year from diarrhea
- 20% all child deaths (under 5 years old) are from diarrhea

The Likely Problem

- 13% of world population lacks access to "improved water sources"

The Goal

- MDG: “reduce by half the proportion of people without access to sustainable drinking water”

The Solution(s)

Really the Problem?

- *Quantity* of water is a better determinant of health than *quality* of water (Curtis et al, 2000)
- Water quality helps little without hygiene (Esrey, 1996)
  - 42% live without a toilet at home
- Nearly 2.6 billion people lack any improved sanitation facilities (WHO)
- People are more willing to pay for convenient water than clean water
- Chlorine is very cheap,
  - In Zambia, $0.18 per month for a family of six
  - In Kenya, $0.30 per month
- Yet less than 10% of households purchase treatment
- 25% of households reported boiling their drinking water the prior day

Kremer, Michael, Amrita Ahuja and Alex Peterson Zwane. "Providing Safe Water: Evidence from Randomized Evaluations" Discussion Paper 2010--23, Cambridge, Mass.: Harvard Environmental Economics Program, September, 2010.

Alternative Solution(s)?

Devising a Solution

- What is the theory behind your solution?
- How does that map to your theory of the problem?

Blueprint for Change

PROGRAM THEORY ASSESSMENT

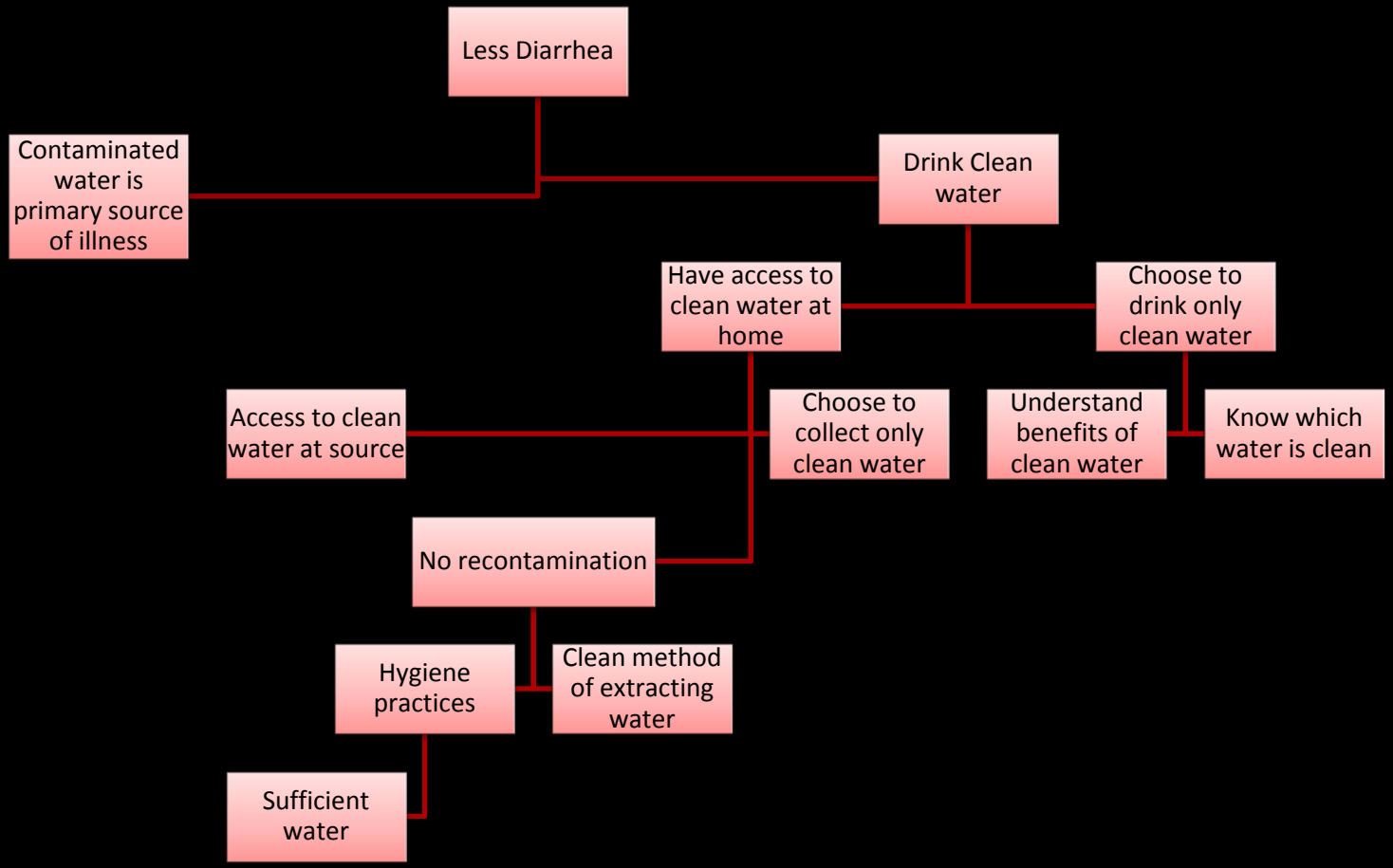
Program Theory Assessment

- Logical Framework (LogFrame, LFA)
- Theory of Change
- Results Framework
- Outcome Mapping

- Causal chain
- Causal model
- Cause map
- Impact pathways
- Intervention theory
- Intervention framework
- Intervention logic
- Investment logic
- Logic model
- Outcomes chain
- Outcomes hierarchy
- Outcome line
- Program logic
- Program theory
- Programme theory
- Results chain
- Theory-based evaluation
- Theory-driven evaluation
- Theory-of-action

Source: Patricia Rogers

Theory of Change



Log Frame

	Objectives Hierarchy	Indicators	Sources of Verification	Assumptions / Threats
Impact (Goal/ Overall objective)	Lower rates of diarrhea	Rates of diarrhea	Household survey	Waterborne disease is primary cause of diarrhea
Outcome (Project Objective)	Households drink cleaner water	(Δ in) drinking water source; E. coli CFU/100ml	Household survey, water quality test at home storage	Shift away from dirty sources. No recontamination
Outputs	Source water is cleaner; Families collect cleaner water	E. coli CFU/100ml;	Water quality test at source	continued maintenance, knowledge of maintenance practices
Inputs (Activities)	Source protection is built	Protection is present, functional	Source visits/ surveys	Sufficient materials, funding, manpower

Source: Roduner, Schlappi (2008) Logical Framework Approach and Outcome Mapping, A constructive Attempt of Synthesis

Program Theory Assessment

- How will the program address the needs put forth in your needs assessment?
  - What are the prerequisites to meet the needs?
  - How and why are those requirements currently lacking or failing?
  - How does the program intend to target or circumvent shortcomings?
  - What services will be offered?

Making the program work

PROCESS EVALUATION

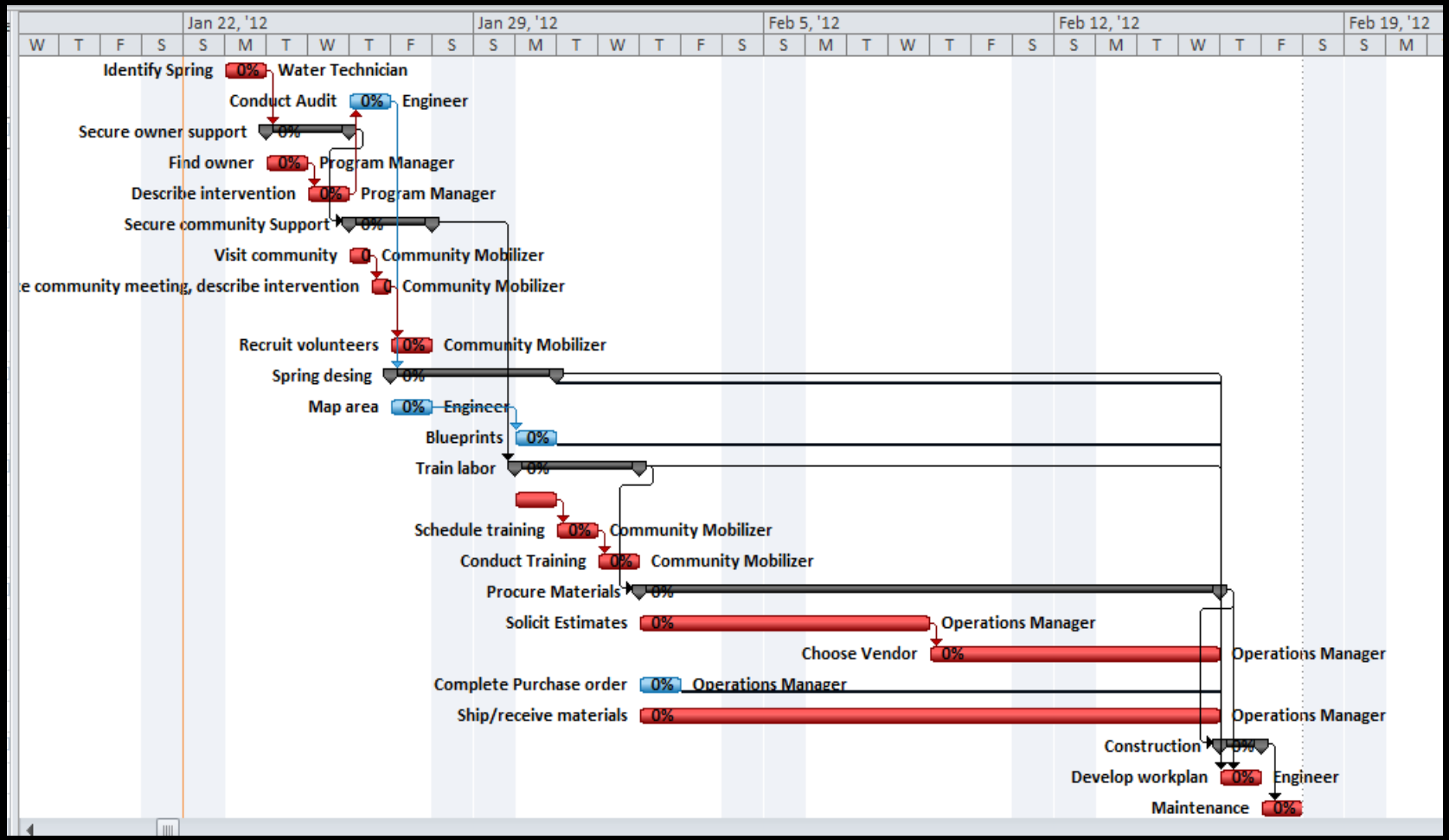
Process Evaluation

- Supply Side
  - Logistics
  - Management
- Demand Side
  - Assumption of knowledge, preferences
  - Assumptions of response

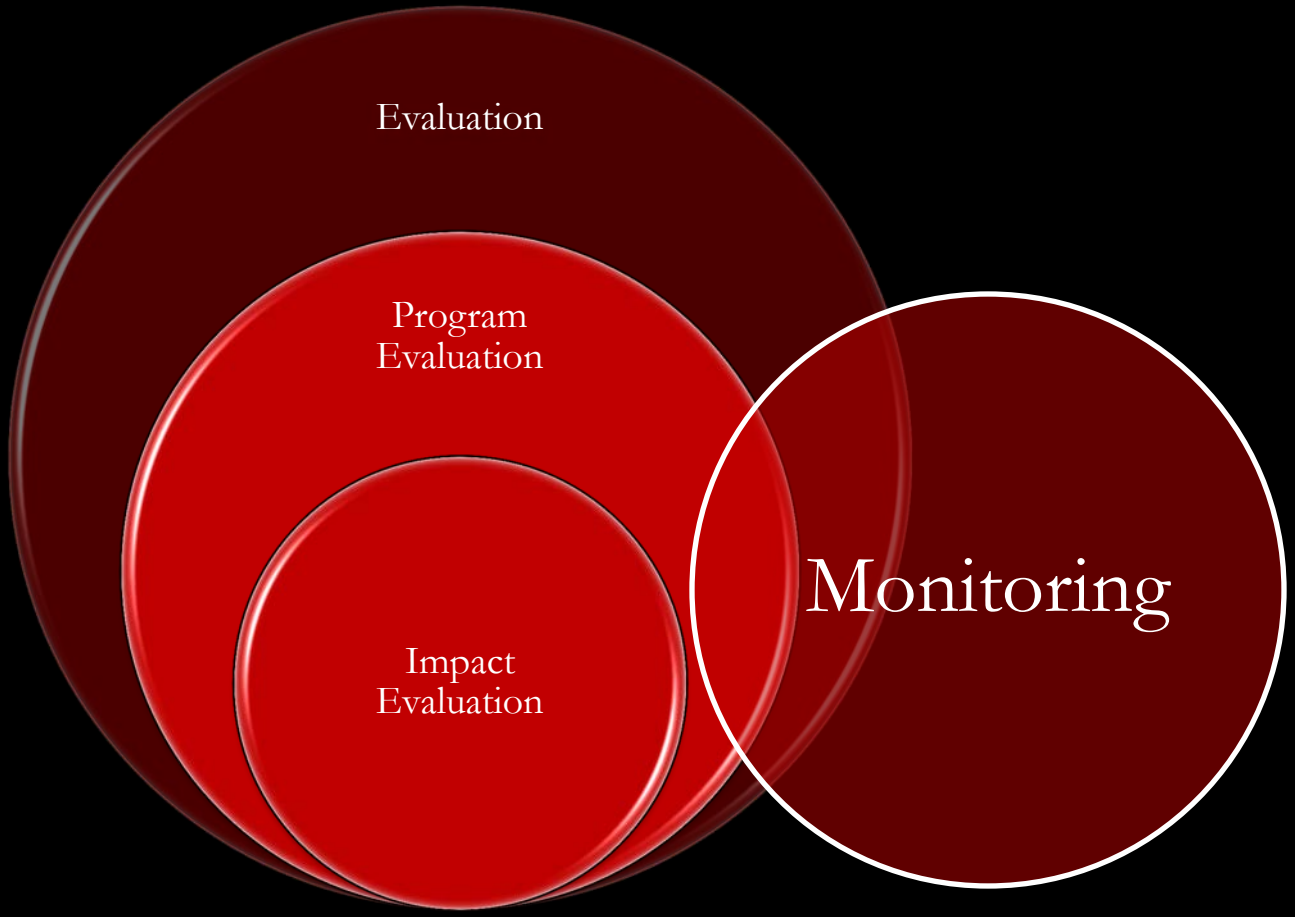
Process Evaluation: Logistics

- Construction
  - Construct spring protection
  - Installing fencing
  - Installing drainage
- Maintenance
  - Patch concrete
  - Clean catchment area
  - Clear drainage ditches

Process Evaluation: Supply Logistics



Monitoring and Evaluation



Process Evaluation: Demand-side

- Do households collect water from improved source?
- Does storage become re-contaminated?
- Do people drink from “clean” water?

Process was okay, so….

- What happened to diarrhea?

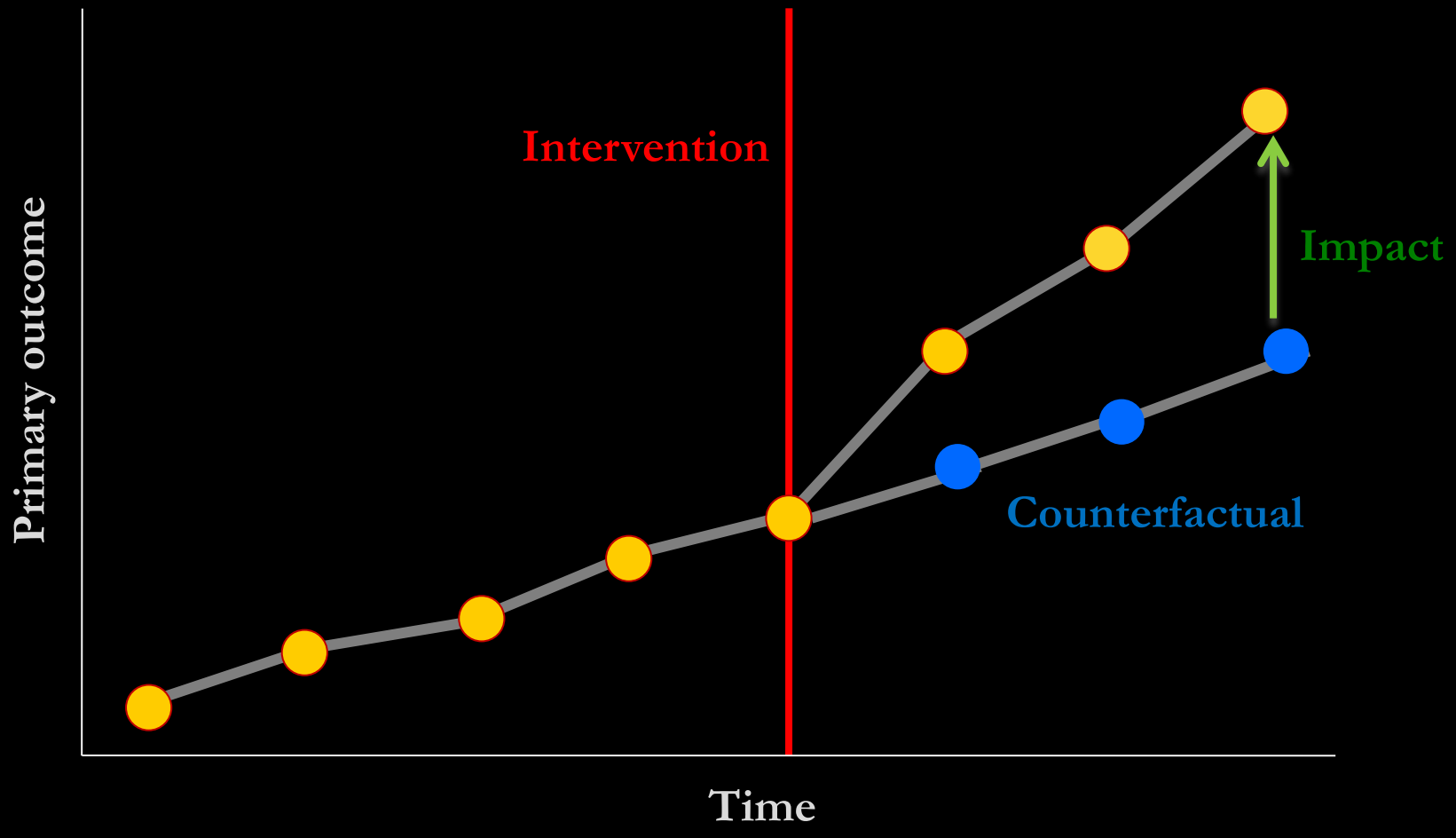
Measuring how well it worked

IMPACT EVALUATION

Did we achieve our goals?

- Primary outcome (impact): did spring protection reduce diarrhea?
- Also distributional questions: what was the impact for households with good v. bad sanitation practices?

What is Impact?



How to measure impact?

- What would have happened in the absence of the program?

- Take the difference between

what happened (with the program) …and
\- what would have happened (without the program)
= IMPACT of the program

Constructing the Counterfactual

- Counterfactual is often constructed by selecting a group not affected by the program

- Randomized:
  - Use random assignment of the program to create a control group which mimics the counterfactual.

- Non-randomized:
  - Argue that a certain excluded group mimics the counterfactual.

How impact differs from process?

- When we answer a process question, we need to describe what happened.

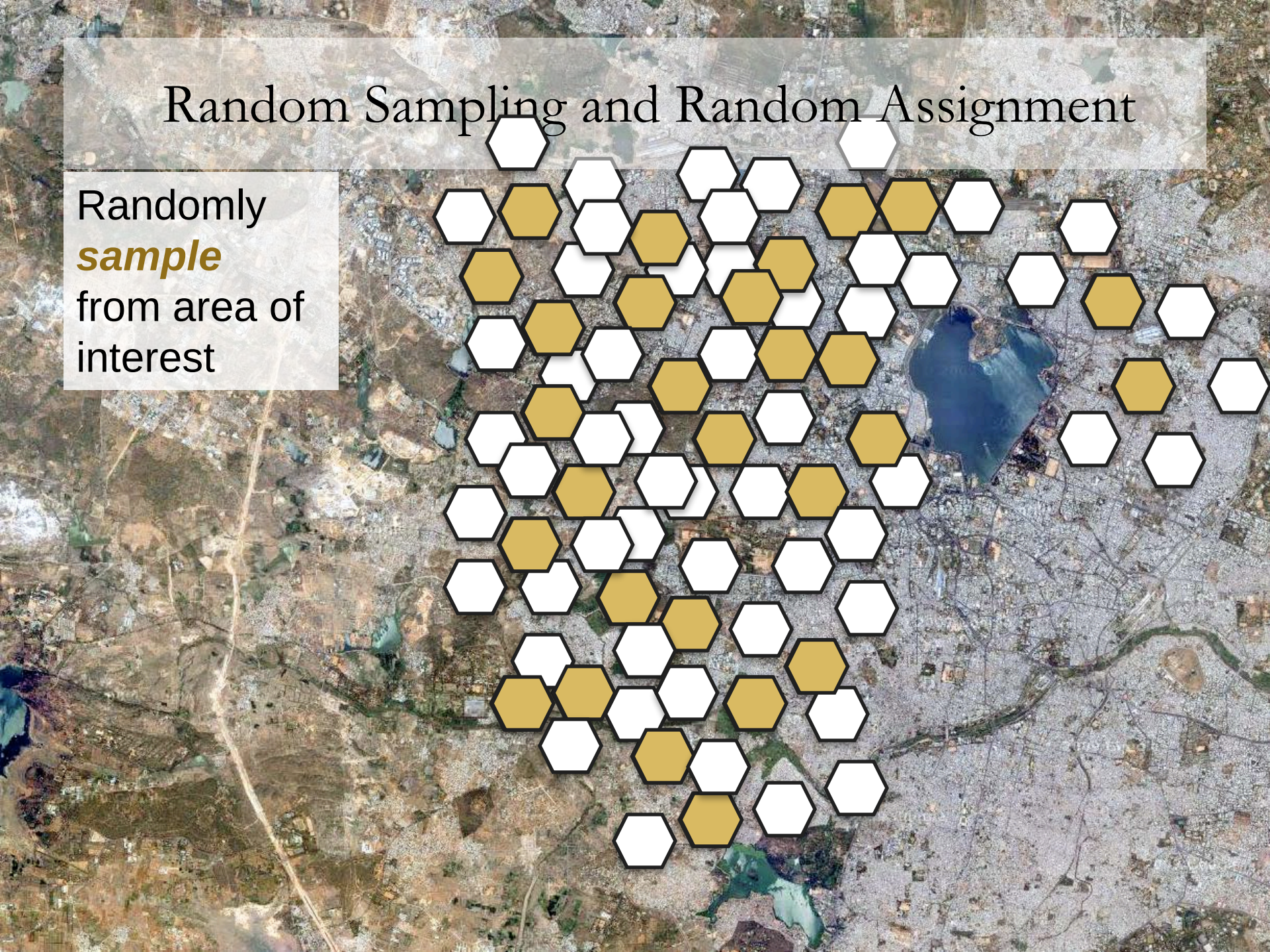
- When we answer an impact question, we need to compare what happened to what would have happened without the program

The "gold standard" for Impact Evaluation

RANDOMIZED EVALUATION

Random Sampling and Random Assignment

Randomly ***sample*** from area of interest

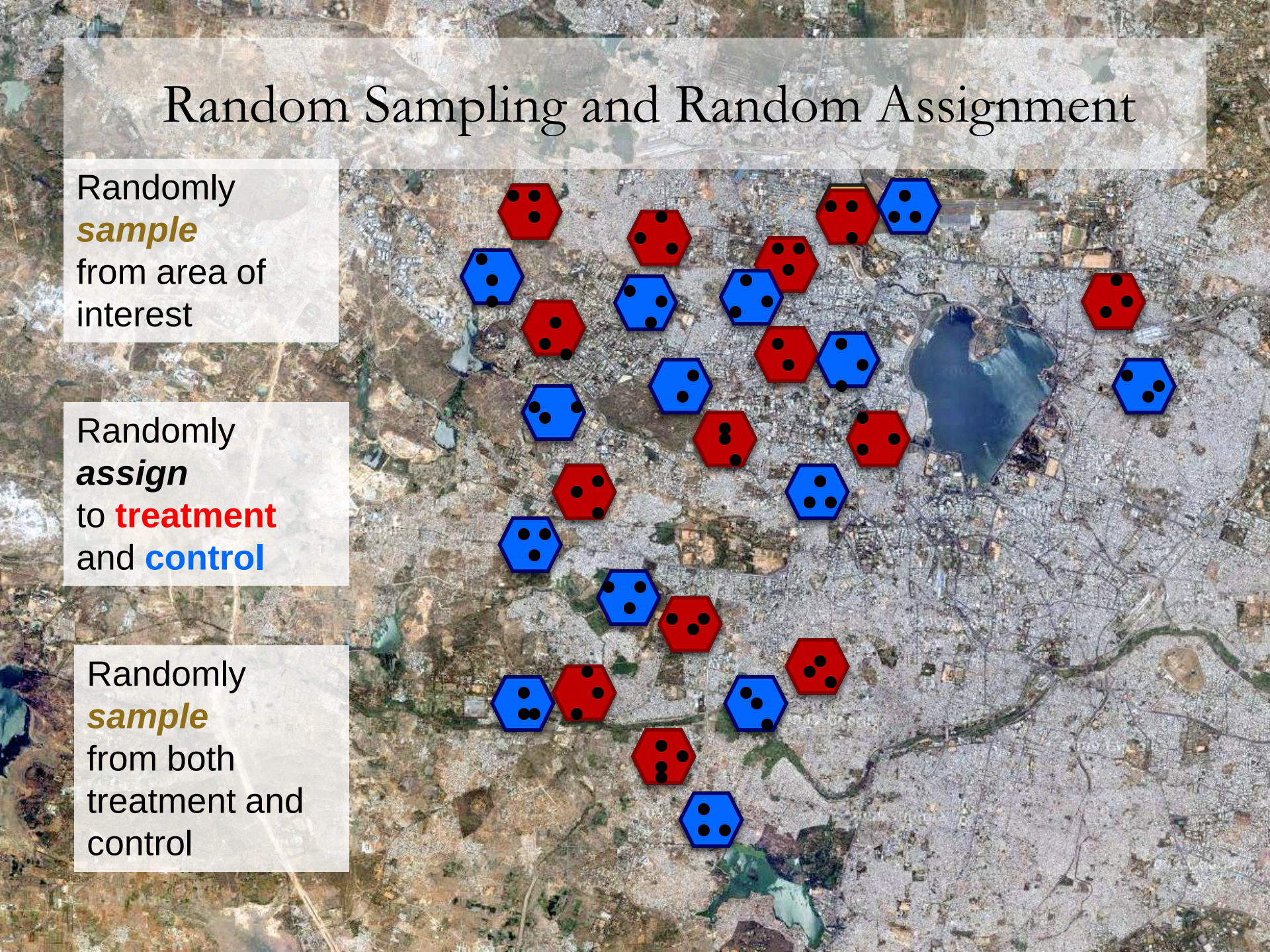


Random Sampling and Random Assignment

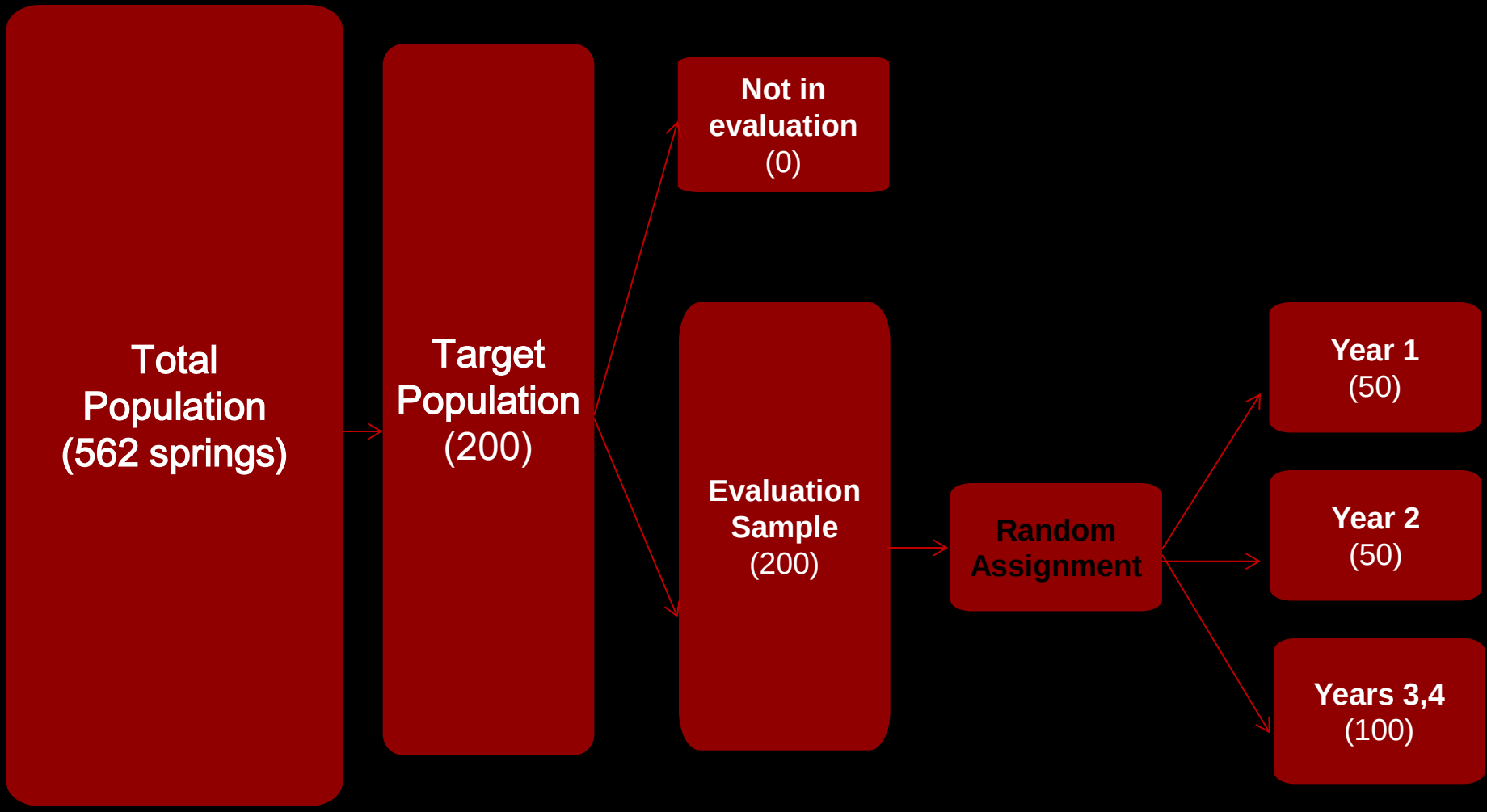
Randomly ***sample*** from area of interest

Randomly ***assign*** to **treatment** and **control**

Randomly ***sample*** from both treatment and control



Spring Cleaning Sample



Impact

- 66% reduction in source water e coli concentration
- 24% reduction in household E coli concentration
- 25% reduction in incidence of diarrhea

Making Policy from Evidence

Intervention	Impact on Diarrhea
Spring protection (Kenya)	25% reduction in diarrhea incidence for ages 0-3

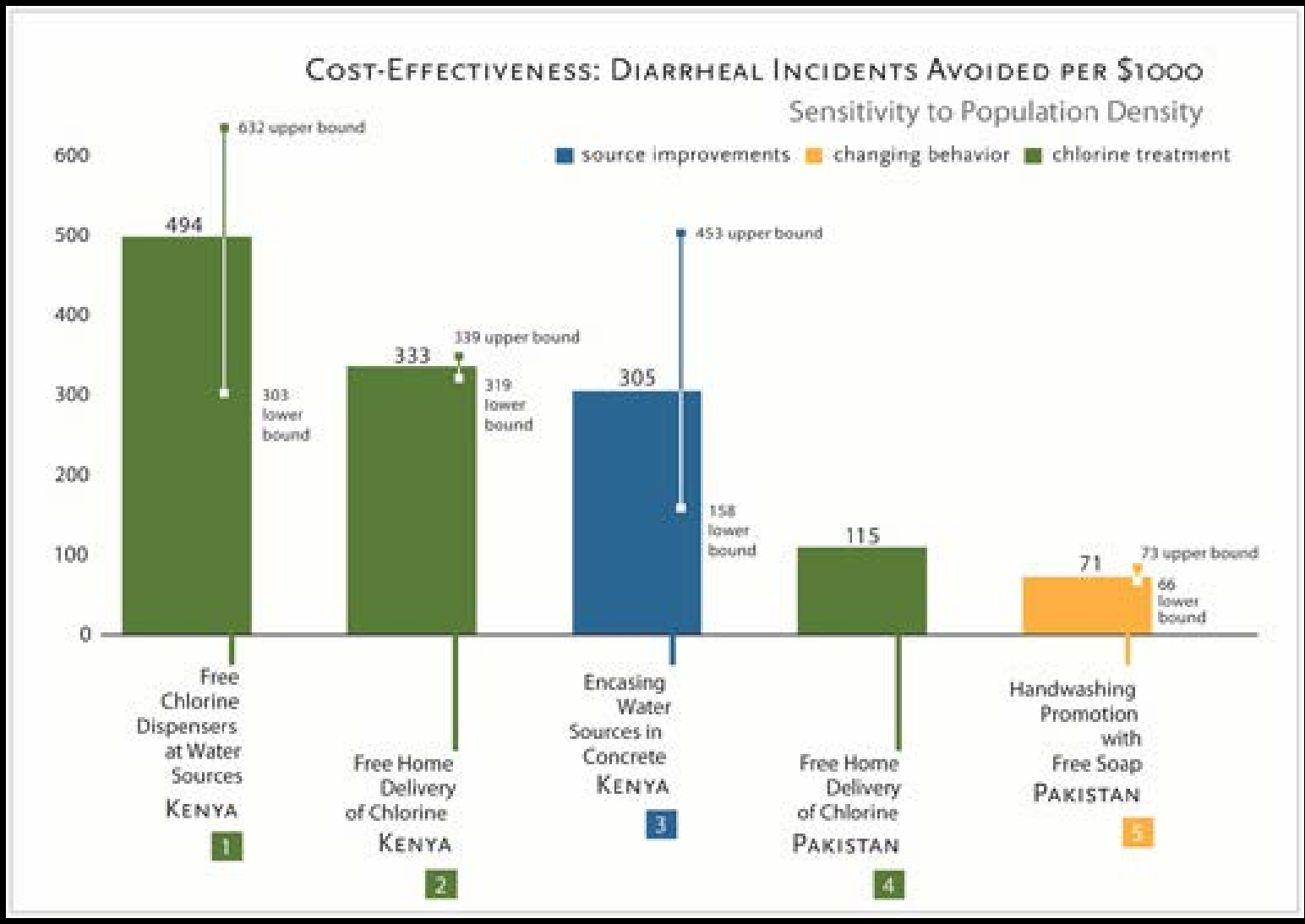
Making Policy from Evidence

Intervention	Impact on Diarrhea
Spring protection (Kenya)	25% reduction in diarrhea incidence for ages 0-3
Source chlorine dispensers (Kenya)	20-40% reduction in diarrhea
Home chlorine distribution (Kenya)	20-40% reduction in diarrhea
Hand-washing (Pakistan)	53% drop in diarrhea incidence for children under 15 years old
Piped water in (Urban Morocco)	0.27 fewer days of diarrhea per child per week

Evidence-Based Policymaking

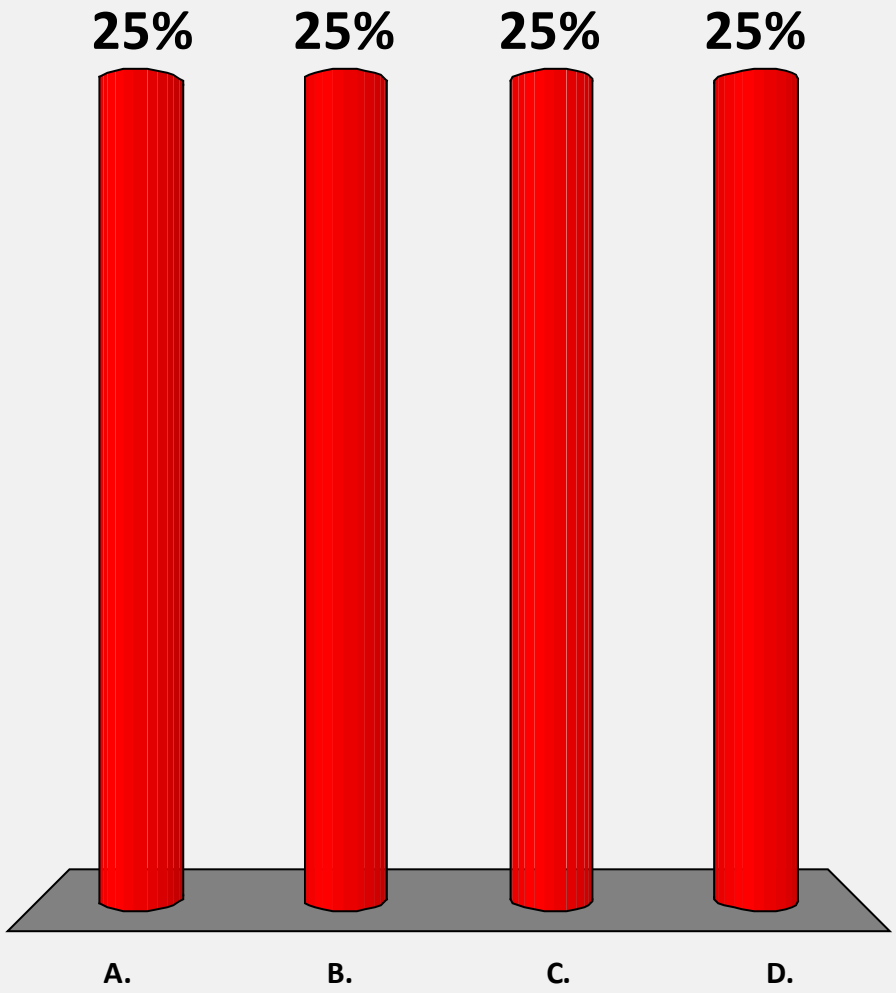
COST-EFFECTIVENESS ANALYSIS

Cost-Effectiveness Diagram



When is a good time to do a randomized evaluation?

A. After the program has begun and you are not expanding it elsewhere
B. When a positive impact has been proven using rigorous methodology
C. When you are rolling out a program with the intension of taking it to scale
D. When a program is on a very small scale e.g one village with treatment and one without



When to do a randomized evaluation?

- When there is an important question you want/need to know the answer to
- Timing--not too early and not too late
- Program is representative not gold plated
  - Or tests an basic concept you need tested
- Time, expertise, and money to do it right
- Develop an evaluation plan to prioritize

When NOT to do an RE

- When the program is premature and still requires considerable "tinkering" to work well
- When the project is on too small a scale to randomize into two "representative groups"
- If a positive impact has been proven using rigorous methodology and resources are sufficient to cover everyone
- After the program has already begun and you are not expanding elsewhere

Developing an evaluation strategy

- Start with a question
- Verify the question hasn't been answered
- State a hypothesis
- Design the evaluation
- Determine whether the value of the answer is worth the cost of the evaluation

- With key questions answered from impact evaluations, process evaluation can give your overall impact
- A few high quality impact studies are worth more than many poor quality ones

- If you ask the right question, you're more likely to care

Components of Program Evaluation

- Needs Assessment
- Program Theory Assessment
- Process Evaluation
- Impact Evaluation
- Cost Effectiveness

- What is the problem?
- How, in theory, does the program fix the problem?
- Does the program work as planned?
- Were its goals achieved? The magnitude?
- Given magnitude and cost, how does it compare to alternatives?