

What Are We Measuring ?

Major Data Types

- Health Status *Indicators*
 - Morbidity & mortality measures
 - Health risk measures
- Health Services Data
 - also: sometimes used as Indicators*
- Population Data
 - Describe populations
 - Denominators for rates
- Qualitative Data



Qualitative Data Collection

- How to collect
 - Focus groups
 - Key informant interviews / diaries
 - Client surveys/ feedback forms
 - Observation



Qualitative Data Collection *continued*

- Knowledge gained can be influenced by context and interaction between researcher and participant
- Often useful to help interpret quantitative findings
- Extremely important for understanding “Why ?” questions



Where do local Quantitative Data come from ?

- **Health Status *Indicators***
 - Births & deaths: Vital Statistics
 - Disease registries: esp. Communicable, Cancer
 - Risk Factor: BRFSS and YRBS
 - Local surveys: NYC Community Health Survey
- **Health Services Data**
 - SPARCS system
 - BRFSS service utilization questions
 - Local data systems (e.g. clinic networks)
- **Population Data**
 - Census Bureau / local planning department

Population-Based Rates

- Why use rates ?
- How to calculate rates

$\frac{\text{\# cases in population}}{\text{\# in population at risk}} \times \text{constant}$
$\frac{\text{\# prostate cancer deaths among Home County residents}}{\text{\# males in Home County population}} \times 100,000$

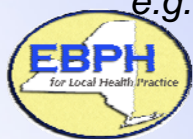
Note similarity to percentages

Population-Based Rates *(continued)*

- Standard practice for different constants
 - Birth risk factor rates are percentages
- Underlying meaning of a rate ?
- Comparing rates of disease between groups or populations
 - Risk-ratio
 - *Incidence of colorectal cancer NYS 1997-2001*
Males: 73.9, Females: 54.1 cases per 100,000
M:F risk ratio = 1.37

Types of Rates

- Crude rates (*aka* Unadjusted)
 - calculation as shown
- Category- specific rates (*aka* Stratified)
e.g., Age-specific, race-specific rates
 - numerator & denominator restricted to subgroup
- Adjusted rates (*aka* Standardized)
e.g., Age-adjusted rates



What data do I need ?

- Community-level data (*e.g., county*)
- Data to place local figures in context
 - How does my county compare to region ?
 - ... to comparable counties ?
 - ... to State ("Rest of State") rates ?
 - ... to US rates ?
 - ... to Healthy People 2010 Objectives ?

Descriptive Epidemiology: Asking Questions from the Data

Be prepared to “Drill-down” in the analysis of health status indicators

- Have the rates of illness changed over time ?
- Is the rate the same in the city and outside the city ?
- How does health status vary among race-ethnicity groups in the community ?
... by SES ? ... by age ?
- How do care patterns compare in my community vs. others ? ... are they changing over time ?
- Is my community’s population getting older ?

Descriptive Epidemiology: Asking Questions from the Data

Start to “Drill-down” in the analysis of health status indicators

- Have the rates of illness changed over time ?
- Is the rate the same in the city and outside the city ?
- How does health status vary among race-ethnicity groups in the community ?
... by SES ? ... by age ?
- How do care patterns compare in my community vs. others ? ... are they changing over time ?
- Is my community’s population getting older ?

Questions can be endless-- need to organize

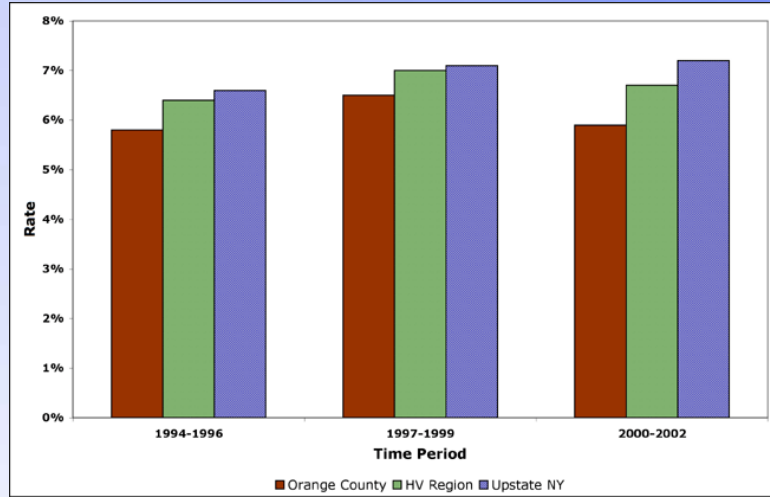
Develop an Analysis Plan: Characterize Disease Patterns

- **By Person** (demographic groups)
 - Groupings defined by age, gender, race-ethnicity, educational attainment, socio-economic status
- **By Place**
 - Nested levels: National, state, county, minor civil divisions (cities & towns), census tract
 - ZIP codes— convenient but problematic
 - Broad categories— *e.g.*, urban vs. rural
- **By Time**
 - Long- or short-term trends
 - Cohort effects

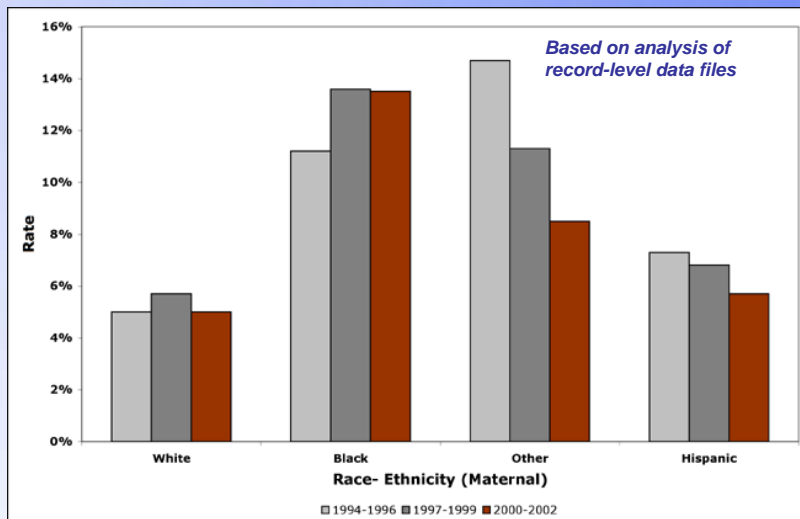
Identifying Health Disparities in the Community

- Identify racial and ethnic disparities in health status indicators
- In smaller communities, typically cannot stratify by socioeconomic status
- Use published literature as guide
(see MMWR. 2005;54:1-3)
- Using standard race-ethnicity groupings for comparison w/NYS and US rates
 - Typically:* Black, non-Hispanic;
White, non-Hispanic;
Asian, non-Hispanic;
Hispanic

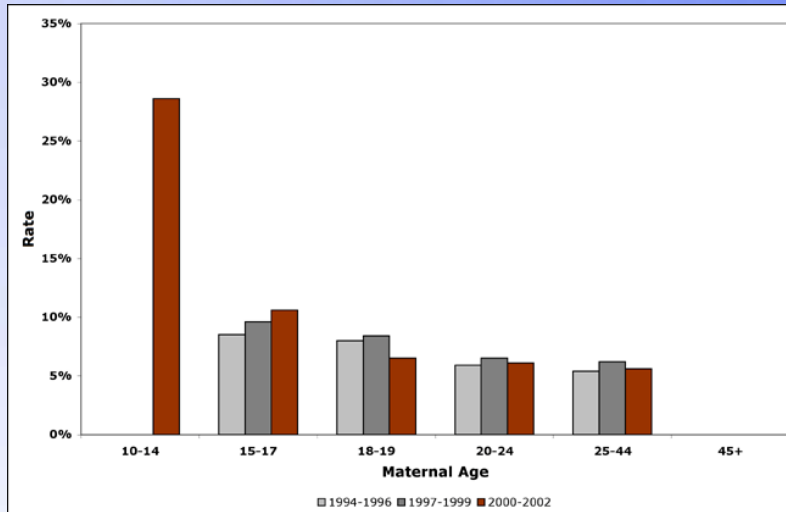
Low Birthweight Births 1994 through 2002 Orange County, Hudson Valley, and Upstate NY



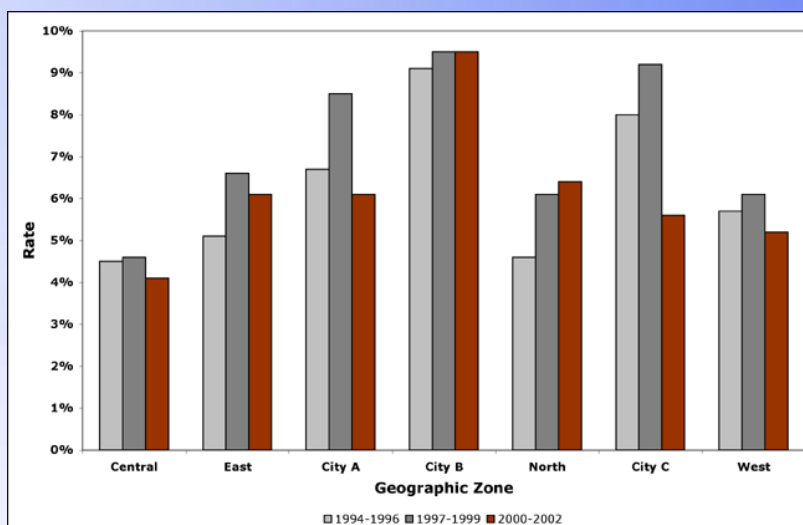
Low Birthweight Births 1994 through 2002 Orange County Race-Ethnicity Comparison



Low Birthweight Births 1994 through 2002 Orange County Maternal Age Group Comparison

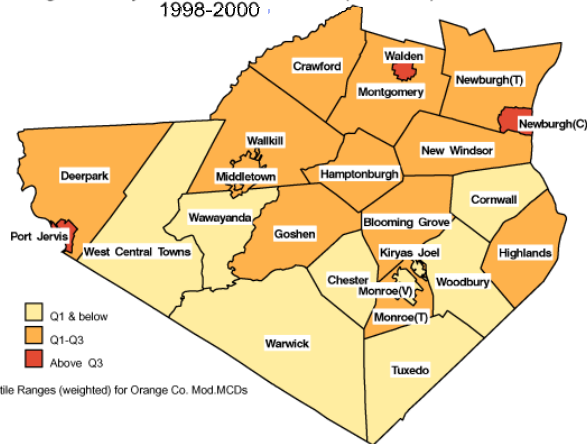


Low Birthweight Births, 1994 through 2002 Orange County Geographic Zones



Displaying Geographic Data in Map Format

Incidence of Late / No Prenatal Care Births
Orange County Minor Civil Divisions (Modified)
1998-2000



Confidence Intervals

- 95% confidence interval: defines a region in which the true mean will fall 95% of the time
“Confidence limits” are the two end-points of this range
- For samples > 100, multiply
standard deviation * 1.96
Upper 95% C.L. = average + (1.96 * s)
Lower 95% C.L. = average - (1.96 * s)
- *For confidence intervals for rates,*
refer to “Silent Partner” extract

Is my rate statistically different from yours ?

- Simple description of the data is a shaky foundation for programmatic decisions
- Need to employ statistical “hypothesis tests”
 - common examples: t-test, chi-squared test, analysis of variance, regression
- “Null hypothesis” is that two rates do not differ
- “ $p < 0.05$ ” statement: less than 5% probability that a difference at least this large occurs by chance

But we're not "going there" !

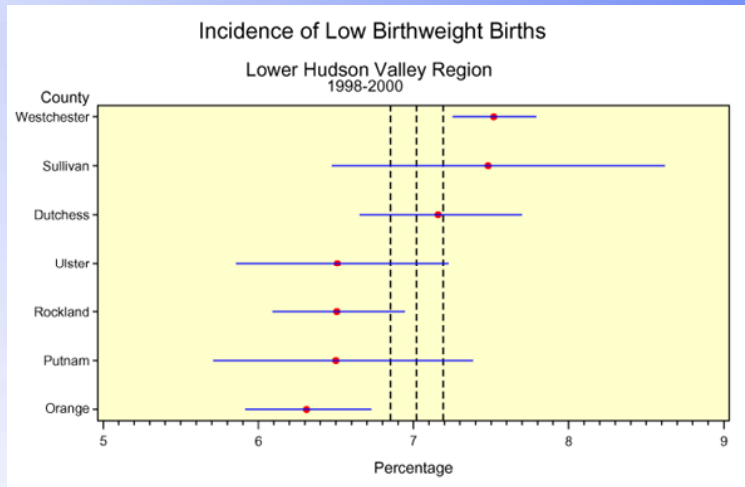
Use Confidence Intervals for Simple Hypothesis Test

“Quick and dirty” hypothesis test:

- If confidence intervals around rates for two populations **overlap**, they are **not statistically different**.
- If they **do not overlap**, then the rates **are statistically different**



Graphical Example



Dealing with Small Number Problems in Epidemiology

- In descriptive analyses, rarely can increase sample size
- How small is **too** small ?
Rule of thumb:
Rates based on fewer than 20 events (or 30) are unstable and should not be reported
<http://www.health.state.ny.us/diseases/chronic/ratesmall.htm>
- **Solution:**
 - Combine data across time &/or space
 - e.g., three year moving averages in time-series data
 - e.g., E-BRFSS combines smaller counties

Concluding thoughts

- Check your analysis carefully **before** going public !
- Organize your data analysis
- Drill-down analytically to characterize health status as far as the data will allow
- Statistical hypothesis tests should be used for important comparisons
- Be cognizant of small number problems in analyzing subsets of data

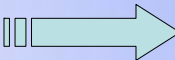
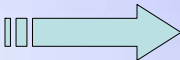


About Research...

- Research is an organized quest for knowledge
- It tests a specified hypothesis with a pre-planned study design
- Evaluation of public health interventions is applied research



Two Major Categories of Applied Epidemiological Research

- Research to identify risks associated with health-related conditions
 Type 1 evidence
- Research to evaluate the effectiveness of public health interventions
 Type 2 evidence

Threats to Validity

- Three Major Types of Bias
 - Selection bias
 - Differences from external population (external)
 - Differences between groups (internal)
 - Information bias *e.g. recall bias*
 - Confounding biases (many)
 - Secular trends
 - Contamination effects
 - Compliance issues
- Biases vary among different study designs and affect their relative strengths

Experimental and Quasi-Experimental Designs

- Type II Studies
- Measure response to an experimental “treatment” (intervention)
- Assignment to treatment vs control group may be randomized (= experimental) or not (= quasi-experimental)
- Response to treatment measured at some point in future

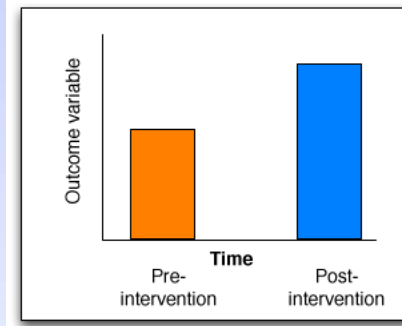
Quasi-Experimental Designs

- Often used in evaluation of interventions
- Assignment to treatment group is arbitrary or purposeful — **not** randomized
- Two general categories
 - Pre-test/ post-test comparison *without* control group (time series)
 - Pre-test / post-test comparison *with* control group (non-randomized trial)

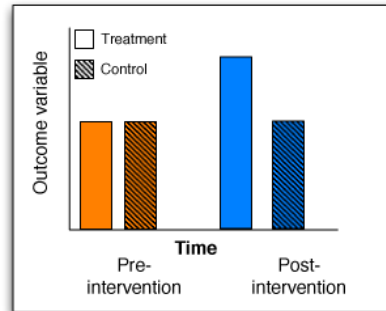


Quasi-Experimental Designs

- No control group
- With control group



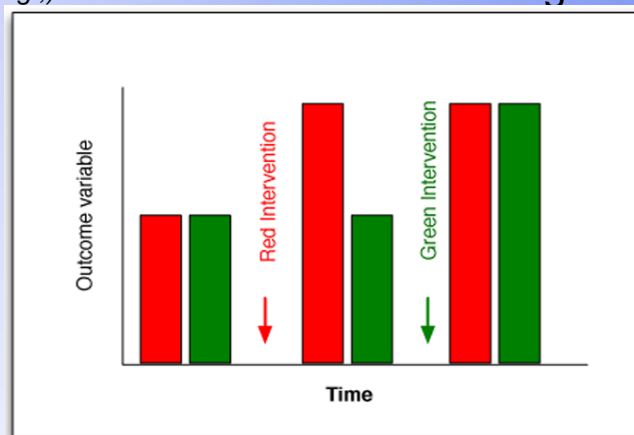
But what would have happened in the absence of the intervention ? (are there any secular trends ?)



But what if the Control Group is different from the Treated Group ?

Quasi-Experimental Designs Can be made Stronger

- (e.g.,) Phased Intervention Design



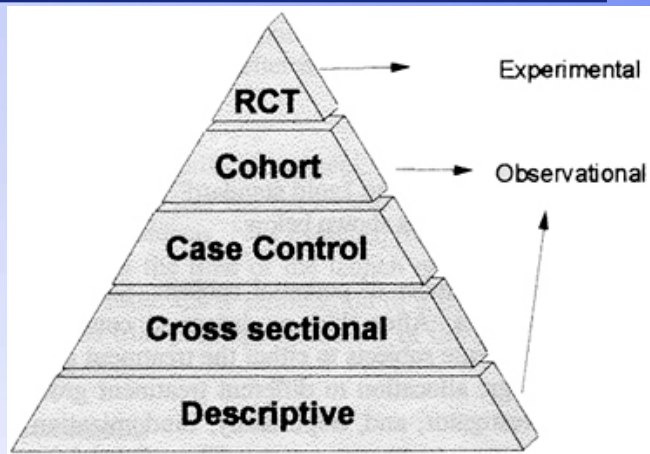
Experimental Designs

- Two types
 - **Group-randomized trials:** *Entire populations* (communities) are randomized to treatment / control groups
 - **Clinical trials:** *Individual participants* are randomized to treatment / control groups
- Randomized trials are the 'gold standard' for showing causal associations
- Ethical issues raised by randomization
- Often not feasible for evaluating public health interventions

General Hierarchy of Research Design Quality

- Randomized clinical and/or community trial
- Cohort studies
- Case-control studies
- Cross-sectional studies (prevalence studies)
- Descriptive ecological (correlational) studies
- Case series/case reports
- Individual evidence (personal experience/ expert opinion)

Traditional Pyramid of Study Design Strength



from: <http://www.sunmed.org/Clinical.html>

Meta-Analysis

- Systematic comparison of multiple studies that examined similar interventions
 - Studies may have different outcome measures or methodologies
- Statistical analysis of pooled data across studies
Each study provides a single data point
(often the odds-ratio of intervention:control)
- Widely used in Community Guide & other evidence-based reviews

The Intervention Decision Matrix

Intervention	Option#1	Option#2	Option#3	Option#4
Effectiveness				
Feasibility				
Affordability				
Political Acceptability				
Social and Political Will				
Unintended Consequences				
Final Priority				

What is Evaluation?

- “A process that attempts to determine as systematically and objectively as possible the relevance, effectiveness, and impact of activities in light of their objectives.”

From: Last JM. *A Dictionary of Epidemiology. Third Edition.* New York: Oxford Press; 1995.

- Complex and diverse field. Ranges from
 - simple monitoring of program activities ... to...
 - intervention trials with complex experimental designs



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, that ***there was no punishment more severe than eternally futile labor....***”



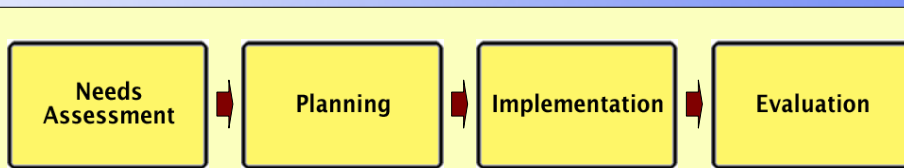
The Myth of Sisyphus

– MMWR

Framework for Program Evaluation in Public Health

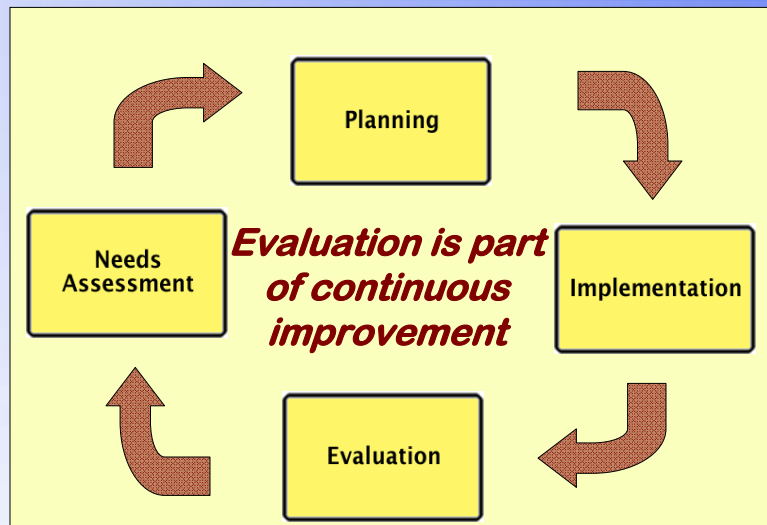
Evaluation as Part of the Program Planning Process

Historically portrayed as part of a linear process:



Evaluation is Pass / Fail

Contemporary Program Planning Process



Evaluation Questions

- Is the program working?
 - Is the intervention being delivered as intended?
 - What aspects of the program are working well or poorly?
 - What can be done *now* to improve the program?
- Did the intervention/program work?
 - For whom?
 - Under what conditions?
 - Were the benefits worth the cost?
 - What program components were most effective?

Evaluation Polarities

Monitoring	↔	Evaluation
Process	↔	Outcome/ Impact
Formative	↔	Summative
Qualitative	↔	Quantitative



Levels of Summative Evaluation

- **Reaction** - Did they like it ? *(especially training)*
- **Impact** - Did it have measurable effects ?
 - Knowledge gained
 - Beliefs changed
 - Behavior changed*(usually) in persons participating in the intervention*
- **Outcome** - Did it affect the population ?
(i.e., changes in population-based rates ?)

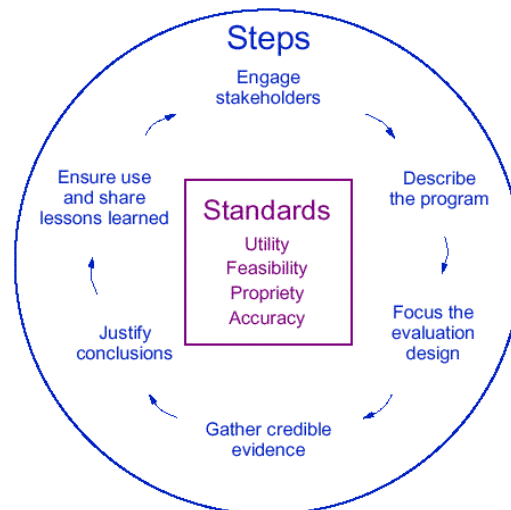
Note: Terminology is not standardized; many evaluators use the terms "impact" and "outcome" evaluation interchangeably.

Framework for Program Evaluation

- 2 year process by CDC
 - designed as a framework for ongoing, practical program evaluation
 - can be integrated with routine program operations
 - involves input from program staff, community members, other stakeholders, not just evaluation experts
- Involves 6 basic **steps** and 4 broad evaluation **standards**

CDC Framework

FIGURE 1. Recommended framework for program evaluation



Standards for Effective Evaluation

At each step, best option is the one which maximizes:

- **Utility:** Choices which best serve information needs of intended users
- **Feasibility:** Choices which are most consistent with logistics and resources
- **Propriety:** Choices most consistent with law, ethics, and due regard for the welfare of those involved and affected
- **Accuracy:** Choices which best reveal and convey technically accurate information

Step 1: Involve Stakeholders

- **Pre-Evaluation:** Early identification of disagreements in...
 - Definition of the problem
 - Priority activities
 - Priority outcomes
 - What constitutes “proof” of success
- **Post-Evaluation:** Get their help with..
 - Credibility of findings
 - Access to key players
 - Follow-up
 - Dissemination of results

Step 2: Describe the Program *(Using Logic Model)*

- Clarity for YOU and/or clarity between you and stakeholders on:
 - What are activities
 - What are intended effects
 - What is the sequence/order of intended effects
 - Which activities are to produce which effects
- Helps in making decisions on where to focus “outcome” evaluation (effects) and “process” evaluation (activities)

Step 3: Focusing the Design *What Question Is Being Asked?*

- What intervention was actually delivered?
- Were impacts and outcomes achieved?
- Was the intervention responsible for the impacts and outcomes?



Step 4: Gathering Evidence

Some Standards for Valuing Evidence

I won't consider this project to have been successful unless.....

- I can attribute any changes to the project
(Often desired; but rarely attained)
- The project reduces disparities
- The project leaves a “legacy”
- The project can be sustained long-term



Step 4: Gathering Evidence

Choosing Data Collection Methods

- How far out the chain of outcomes do you need to measure ?
-What are your “*Intrinsically valued outcomes*” ?
- Influences on choice of methods
 - Time / Cost
 - “Hawthorne effect”
 - Ethics & Sensitivity of the issue
 - Validity & Reliability
- Usually trade-off of accuracy and feasibility

Step 5: Justifying Conclusions

Claims About Intervention Effectiveness

- Performance against a comparison/control group
- Time sequence of changes
- Plausible mechanisms (or pathways toward change)
- Accounting for alternative explanations
- Similar effects observed in similar contexts



Step 6: Ensure Use & Lessons Learned

- Think about dissemination strategies
 - How to reach key stakeholder groups
 - How the message may differ in emphasis among them
- Commitment to respond to findings



Working with an Evaluator

- Functional partnerships between researchers and service providers
 - Earlier incorporation of evaluation design
 - Service providers knowledge of
 - Client/community needs
 - Important outcomes to measure
 - Community factors that influence outcome
- Dynamics of collaboration
 - Setting up equal partnerships
 - Impact of ethnicity and culture on research
 - Exit issues
- 10-15% of program budget is reasonable cost