

# Enmeshed or Engaged? Program Evaluation Strategies in Cross-System Programs

The National Conference on  
Substance Abuse, Child Welfare  
and the Courts

September 15, 2011

1

**Ruth Huebner, Ph.D.,**  
Child Welfare Researcher,  
Eastern Kentucky University

**Robert Walker, M.S.W., L.C.S.W.,**  
Behavioral Health Researcher,  
University of Kentucky

2

## Disclaimers and Caveats

- **We are evaluation researchers:**
  - Using naturalistic or observational designs
  - We are engaged, but sometimes worry about being enmeshed scientific advisors using a partnership model with practitioners.
  - Both of us are deeply embedded in state government funded projects.

3

- **We want to open a conversation** with you regarding the evaluator role and implementation of program evaluation.
- **Our premise:** The demands of the current services grants exceed traditional role definitions between program and evaluation professionals in terms of:
  - complexity of service delivery;
  - complexity of selection and use of evidence-based practices; and
  - how to represent what services actually do.

4

## Objectives

- To explore emerging roles for program evaluators in projects with inter-agency collaborations.
- To demonstrate how practitioners and administrators can take a more active role in owning the evaluative processes and outcomes.
- To explore strategies of merging provider-collected with administrative data from multiple data sets.
- To understand analysis and model building with complex nested data.

5

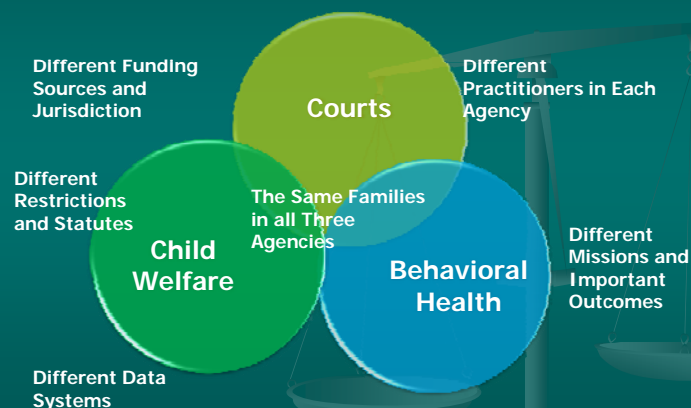
## The Role of the Evaluator in Program Evaluation

**ENGAGED PARTNER,  
SCIENTIFIC ADVISOR,**

**OR; STRICT INDEPENDENCE  
FOR PURE "SCIENCE"?**

6

## The Research Challenge: Can you Control the Environment?



## Research Design Challenges

### Randomized Control Trial

- Discrete research question or problem:
  - Often short term
  - Readily measured and simple, often singular
  - Focus on discrete disorders
  - Controlled conditions
- Discrete Intervention
  - Medication
  - Simple comparison of intervention v. TAU

### Cross Agency Studies

- Complex Problems
  - Substance abuse
  - Multiple co-morbidities
  - Families (whatever they are)
  - Communities
  - Multiple agencies
- Complex Intervention
  - Multiple strategies
  - Many staff and turn over
  - Many individual exceptions
- Complex outcomes

8

## Problems with “Pure” Science

- While clinical trials are the gold standard, they need close scrutiny in application to real world clients.
- They focus on single, discrete disorders and exclude subjects with co-morbidities.
- They also rarely account for subtle cultural or inequality effects among the sample by comparison to ‘real’ clients.
- They generally rely on well trained providers whose skill level may exceed community practice.

9

## Another worry about Evidence-Based Practices

- Many practitioners are unprepared to read or read critically the methodology and analytical sections of research papers.
- Governmental requirements to use evidence-based practices diminishes the *science* supporting *individually applied interventions* in favor of program adoption of a favored practice.
- Dogma? Or science? Can't be both.

10

- Most everyone fails to note that the “change” reported in clinical trials in behavioral health is simply subjects no longer meet all the diagnostic criteria for a disorder.
- The dropping of one DSM criterion can mark the change from disordered to recovered.

11

- The clinical trial outcomes almost never consider actual functionality such as improved overall well-being, improved employment, relationship satisfaction, social interactions, etc.
- They tend to focus on very simple, concrete outcomes with an assumption that not meeting DSM criteria mean overall improvement.
- They cannot be tried with all key ethnicities or cultural groups due to funding limitations.

12

## Challenges of Random Assignment

- Seen as the 'gold standard' of program evaluation but is increasingly difficult when dealing with people in naturalistic settings. Discussion by Guo & Fraser (2010)
- Randomization may not exclude hidden selection bias and is influenced by processes such as the Hawthorne effect and potential biases in assignment to conditions
- Average TX effects mask individual results that may stem from multiple and important factors.

*Propensity Score Analysis*. Sage Publications

13

## Positive Outcomes Fade Over Time

- Lehrer (2010). *The truth wears off*. New Yorker, 12/13.
  - Even studies of drug effects when replicated over time tend to show diminished results. This is surprising, given that drugs are supposedly highly consistent. Other results also diminish.
- Positive results are not published.
- Studies with animals and humans are highly susceptible to many sorts of perception and selection biases.

14

## A Variety of Methods

- GAO Report (2009). *Program Evaluation A Variety of Methods Can Help Identify Effective Interventions*
- Requiring randomization as the sole proof of intervention effectiveness will exclude many potentially effective and worthwhile practices.
- Random assignment is not practical or even possible for some studies – particularly when court and legal proceedings are involved.

15

## Programs Change Over Time

- Klein (July 18, 2011). *Head Start Doesn't Work*. Time Magazine
  - Proven in the 1960's to improve school performance.
  - Taken to scale
- Head Start Study by DHHS found that the effects were minimal and vanished by 1<sup>st</sup> grade.
- Regression toward the mean or mediocrity.

16

## Randomized Control Trials of Parachutes

- Did you know that there have never been ANY RTCs of parachute use as a prevention of death when falling from heights?
- The parachute companies are making huge profits with only observational evidence.
- Smith & Pell (2008). British Medical Journal



17

## The Current Reality of Program Evaluation

- The data for program evaluation must come from practitioners who are often unsophisticated about scientific data collection methods, and anxious or uninterested in evaluation results.
- Approaches like random assignment are generally viewed with resistance. Research clarity may fall victim to routine clinical habits.

18

## Emerging Role?

### Traditional

External  
Expert  
Practitioners are Dependent  
Independent Judgment  
Report to Prove

### Scientific Advisor

Internal  
Coach  
Practitioners are Partners  
Collaboration  
Reports to Improve

19

## Scientific Advisory Roles

- What is meant by 'scientific advisor?'
- We think there is need for someone to be a bridge between practice and science – both the science of selecting/implementing practices and the science of examining outcomes.
- Examples:
  - Guiding providers about a range of practices for the given client population;
  - Suggesting instruments to measure key clinical or behavioral characteristics for both clinical and research needs;
  - Suggesting ways to improve fidelity or at least describable service characteristics.

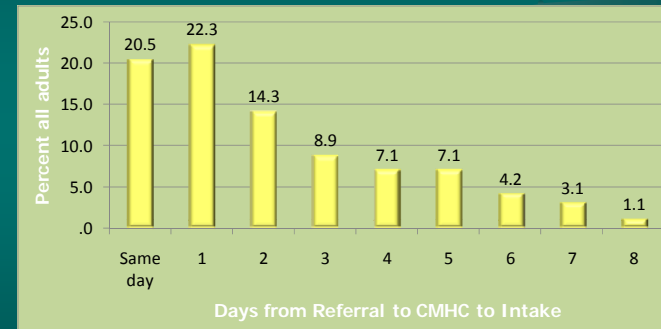
20

## Continuing Advisor Role

- Check for adherence or fidelity to program procedures and serve as a catalyst for growth and increased potency for intervention. Make adjustments to program; test the reality of program assertions.
- Improve program accountability for achieving results rather than completing process or producing outputs.
- Examine the impact of strategies on results to infuse this idea: "Does what we do make a difference in what we want to achieve?"
- Guide decisions on continuing programs and showing the value – costs/benefits.

21

## E.g. Program Strategy: Quick Access to the First Treatment



90% go from Referral to Intake in 8 days

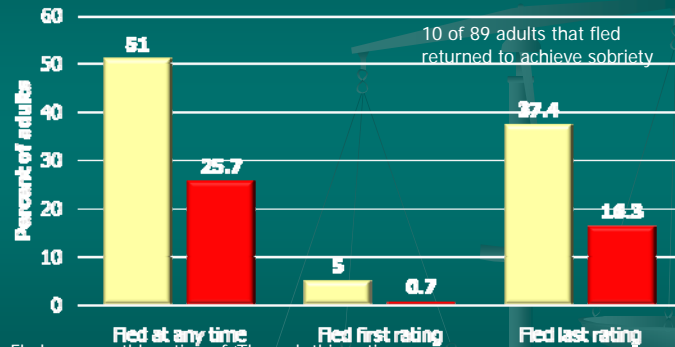
May 2011

22

## Adults that Flee

(closed case opened >=6months; N=251)

■ Male ■ Female



Fled = a monthly rating of 'Through this entire month, unable to locate this adult at any time'

March 3, 2011

23

10 of 89 adults that fled returned to achieve sobriety

## Discussion

How have you struggled with your role?

What do you think your role should be?

Is there room for variation in this among different programs and evaluators?

24

## Engaging Practitioners In Evaluation Research

Overcoming Terror Among  
Practitioners and Learning to  
Learn Together

25

## Program Evaluation and Practice: Goals and Fears

- **Articulate and then monitor clear outcome expectations** (“but we will know if it is working and it is way too hard to show it!”)
- **Improve program delivery and outcomes** (“but the program is based on solid documented principles and EBPs, so we know what we are doing, and this will be so boring and take so long!”).

26

## Program Evaluation and Practice: Goals and Fears

- **Show effects of an innovative program** (“but what if we find out it isn’t working – what then?”)
- **Demonstrate benefits in terms of a contract, funding agency, public, customer, or government** (“but the results are statistically significant at the .05 level!”) (Research myth)

27

## Things to Remember in this Partnership

- Evaluators may carve out various roles in publicly funded programs, but...
  - It is likely that the data belong to the program, not the evaluator;
  - Some contracts call for program roles in publications based on program data;
  - Negotiating terms for *all* communications to the public about what can and cannot be said about the results is critical.

28

## Engaging Practitioners

Practitioners own the data *and* the results unless otherwise agreed.

- *If you want providers onboard when you land, make sure they are onboard when you take off.*
- *Ask them "How will you know if it is working?"*
- *Teach them about the uselessness of mere anecdotes.*

29

## Engaging Practitioners

Evaluation can be a force for change and improvement.

- *Program first rather than evaluation first.*

Evaluation is a process not a report.

- *Explore options and ideas, not jump to conclusions.*
- *Work together to understand program processes and outcomes*

30

## Engaging Practitioners

Providers can tell the story too.

- *Present findings in graphs, diagrams that they understand.*

Overcoming terror and learning to learn together.

- *Present snippets of data at every meeting and generate reports?*

Ask "What does this mean for families?"

- *Translate results in the numbers of children or families.*

31

## Discussion

How have you engaged practitioners?

What challenges do you face?

Contractual or interpersonal barriers?

32



## Nested and Integrated Data

The Challenges of Cross System Data Complexity

33

## Making Data Work Together

*An Idea whose time has come!*

- Provider-Collected data
- Census Data by County
- Child Welfare Data
  - Child based
  - Family based
- Court Data

*An Idea that is very hard to implement!*

34

## Provider-Collected Data

- Be prepared for the limitations that go with these data.
- Reliability and validity need to be described carefully.
- Provider data in electronic form may be part of MIS financials and may not capture salient clinical variables as needed.
- Extraction of data from records is extraordinarily complex and compromised by reliability problems.

35

## Provider Data

- Don't assume that terms mean what everyone says or thinks they mean.
- Your evaluator use of terms may vary greatly from provider use.
  - Even 'simple' terms can be a source of confusion – what is 'admission', 'case management', - how is discharge defined, etc.

36

## Census Data

- Be sure to look into other secondary sources of census data and analyses of these data.
- University of Wisconsin has county-level data for every county in the U.S. with many other variables added to the census facts.

37

## Child Welfare Data

- Check to see how these data are configured and how these tables can or cannot relate to other state data sets.
- How does the child protective service data system preserve identities across varying family configurations?
  - i.e., one child who now has a new stepparent and new step siblings, but the other child of the original family does not and is separated.

38

## Court Data

- Again, check carefully for formatting, how identifiers work across docketed cases.
- Do the data allow for following persons or just cases?
- Are these data already matched or matchable to police or corrections data?

39

## Matching

- Programs serving families with maltreatment histories often must match:
  - Child welfare data (NCANDS)
  - Court data
  - Police and arrest data
  - Behavioral health provider data
  - Treatment Episode Data Set (TEDS)
  - Vital statistics data (birth, death)
  - Research-specific data

40

## Constraints

- Various confidentiality and data security provisions often limit client identifier uses.
- HIPAA rules can interfere with the use of identifiers.
- State IT support for doing anything other than the quotidian is generally tepid.
- Many state data sets are in antique formats with out-of-date data dictionaries, etc.
- “Confidentiality” is also a way for agencies to avoid the work involved in doing something with the data tables.

41

## Tips for Matching

- Whenever possible, collect data only once.
- Try to reduce duplication not only for program efficiency, but two measures of the ‘same thing’ often are not the ‘same thing.’
- Be sure to develop clear anchor dates for needs matching.
  - Most state and agency datasets contain multiple year data – watch for contamination from previous year data.

42

- Check carefully for agreement of variable definitions across datasets.

- E.g., many datasets will include a field for ‘substance use’.
  - Was this a lifetime measure, past 30 day, past 12 month, past 90 days?
  - Is it specific to type of drug or alcohol?
  - Does the field incorporate or signify multiple drugs?
- Decisions about which substance use variable to use may depend on which dataset has the most robust measure.

43

## What is a Service?

- How are services defined by the different systems?
- What events are recorded? Are events things that happen to the case or are they what an agency does? Both?
- How up-to-date are the data you are using?

44

## Provider Opinion Data

- What are the variations in opinions that color the client data set?
  - E.g., around client compliance coding?
  - Progress?
  - Termination status?
- Are there provider biases that are affecting these measures?

45

## What is Admission? Discharge?

- What constitutes case opening in each of the datasets?
- What constitutes case closing? How do you handle "continued generally?"
- What definitions must you use for your reporting to the RPG and for the local evaluation?
  - Are these the same?

46

## What to do with Conflicting Data

- How are you going to handle conflicts between data sets?
  - E.g., the child welfare dataset shows the client employed.
  - The provider shows unemployed, disabled.
  - The court data shows client stole from wife's employer.
- Are your rules for RPG the same as for the local evaluation?

47

## Talk through your Queries

- The development of grounded queries may take several runs and trials.
- There may be many preliminary steps to prepare a dataset before matching begins.
- For example, the algorithm for obtaining birth events data in Kentucky takes 13 steps just to prepare the table and identify the potential sample.

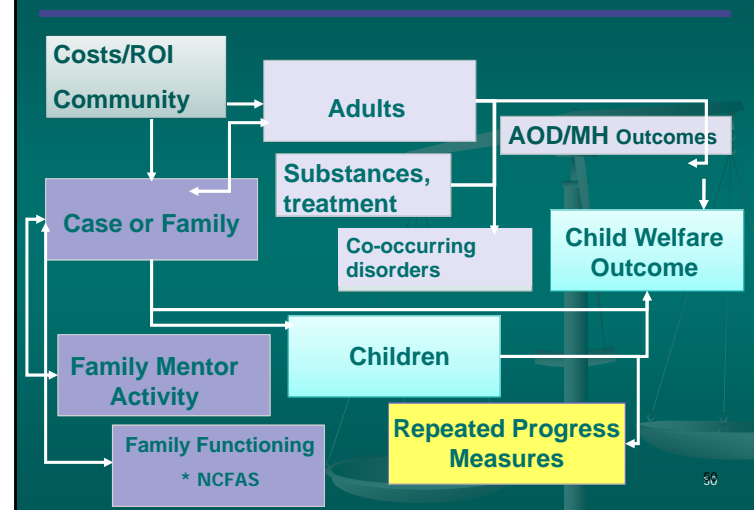
48

## Bin Bag or Managed Data Tables?

- Some state datasets simply admit anything submitted as long as field definitions are correct.
- What data cleaning is done by the state IT staff? What procedures are used for this?
- Could data cleaning effect what you are looking for?
- Consultation is critical to avoid mistaken understanding of tables and variable definitions.

49

## Nested Data from Several Sources



50

## Our Challenges in Using These Data

- Timing and sequence of events: When is recurrence of child abuse *really* recurrence?
- Cutoff points for any strategy: What is 'quick' access to treatment?
- Messy data or incomplete data: How to avoid bias in correcting, discarding?

51

## Steps in Analysis for Paired Datasets

- Descriptive statistics
- Code for group membership
  - Naturally occurring groups or propensity scores
  - Groups based on sequence of events
  - Groups based on low or high dose
- Comparative statistics control for case risk, age of the child, or other confounds as needed.

52

- Map into logic models
- Building models – logistic regression, survival analysis, SEM to examine multiple relationships.
- Translate back into 'What does this mean for families, for children, for communities?'
- Present in graph or picture.

53

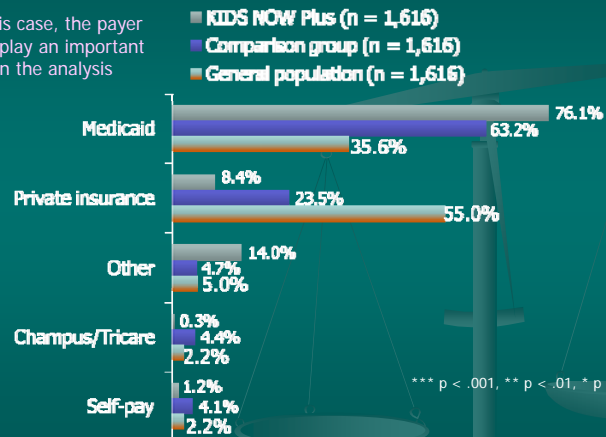
## Keep Vigilant for 'Hidden' Confounds

- In examining birth outcomes for a sample of pregnant women with substance use disorders, examining other demographics may be important.
- Payer source is of growing concern as part of health outcomes.

54

## Source of Pay for Birth of Baby\*\*\*

In this case, the payer may play an important role in the analysis



## Discussion

What challenges do you face when merging and analyzing data?

What are your unresolved issues and challenges?

56

## Thank You and Contact Information

---

**Ruth A. Huebner, PhD**  
Child Welfare Researcher  
Kentucky Department for Community  
Based Services  
[rutha.huebner@ky.gov](mailto:rutha.huebner@ky.gov)

**Robert Walker, M.S.W., L.C.S.W., M.Div.**  
University of Kentucky  
333 Waller Avenue, Suite 480  
Lexington, KY 40504  
(859) 257-6623  
[Robert.Walker@uky.edu](mailto:Robert.Walker@uky.edu)  
[jrwalk0@email.uky.edu](mailto:jrwalk0@email.uky.edu)

