



---

## Supporting Evidence-Based Home Visiting to Prevent Child Maltreatment

---

### Evaluating Outcomes Across Evidence-Based Home Visiting Models: Shared Expectations and Measurement

January 25, 2012

26<sup>th</sup> Annual San Diego International Conference on Child and Family Maltreatment

Debra A. Strong, Kim Boller  
Mathematica Policy Research



# Measuring Outcomes

---

- **Measuring outcomes is an important element in replicating/scaling up evidence-based programs**
- **For example, the federal MIECHV home visiting program requires states to report 6 benchmarks, including:**
  - Child health
  - Maternal health
  - Child development and school readiness
  - Positive parenting practices
  - Reductions in child maltreatment



# Measuring Outcomes for EBHV

---

- **EBHV asked: “Do home visiting programs improve family and child outcomes when implemented in the ‘real world’ and supported by investments in infrastructure?”**
- **Each grantee was required to conduct its own outcome evaluation**
  - Grantees chose different program models and served different target groups
  - So the cross-site evaluation was planned as a ‘systematic review’ of grantee evaluations



# **We Developed Shared Outcomes and Measures**

---

- **We identified outcomes and measures in collaboration with grantees and their evaluators**
- **We provided technical assistance on**
  - Strengthening evaluation designs
  - Implementing measures
  - Collecting data
- **Although the cross-site review of evidence will no longer take place, useful lessons are emerging from the EBHV experience**



# Focus of This Presentation

---

- **What process and criteria were used to choose outcomes and shared measures?**
- **What outcomes and measures were selected, and why?**
- **What information and support has been provided to implement the measures?**
- **How have the evaluations progressed?**
- **What are the implications for MIECHV and other evidence-based initiatives?**



# The Process

---

- **The cross-site team met with EBHV peer learning network (PLN) members**
- **The group identified potential outcome domains, constructs, and measures**
- **Group members did research and recommended a ‘common core’ set of constructs and measures**
- **Mathematica reviewed other constructs and measures proposed by individual grantees for their evaluations**



# The Selection Criteria

---

- **Constructs:**
  - Common interest across multiple grantees
  - Captures what was being attempted (and measured) across multiple program models
  - “Practicality” based on evaluation budgets and expertise
- **Measures:**
  - Validity, reliability, psychometric properties
  - Successful use in other large-scale research
  - Appropriateness for varied cultural, racial, ethnic, and linguistic backgrounds
  - Availability



# The Recommended Core Constructs

---

- **Parent outcomes**
  - Substance use
  - Depression
  - Use of harsh discipline
- **Child health, development, and well-being**
  - Immunizations
  - Emergency room visits
  - Hospital overnights
  - Developmental progress
  - Behavior problems
  - Child maltreatment



# Selected Measures: Parents

---

- **Substance use**

- Michigan Alcohol Screening Test (MAST) or
- Drug Abuse Screening Test (DAST)

- **Depression**

- Center for Epidemiological Studies Depression Short Form (CES-D SF)

*Both measures are free, easy to score, with good reliability and validity*

- **Harsh discipline**

- Spanking in past month (yes/no) and frequency if 'yes'



# Child Health

---

- **Immunizations**
  - Question on immunization status (scaled from “completely up to date” to “never received any”)
- **Emergency room visits**
  - Three questions on whether child ever taken to emergency room; if ‘yes’ how often, and how many visits were due to accident or injury
- **Hospital overnights**
  - Four questions on whether child stayed overnight in hospital since birth, and if ‘yes’ how many times for at least one night and number of nights, and how many were due to accident or injury



# Child Development

---

- **Ages and States Questionnaires (ASQ-3)**
  - 30 parent-reported questions
  - Covers ages 40 months to 60 months
  - Good reliability and validity, reasonable cost, quick scoring
- **Denver II**
  - 125-item direct assessment of child
  - Covers birth through 6 years
  - Administration and scoring require training
  - Reasonable reliability and validity; reasonable cost



# Child Social-Emotional Development

---

- **Behavior Problems Index (free)**
  - 28 parent-reported items
  - Covers ages 3 to 17 years
- **Child Behavior Checklists**
  - 99 parent-reported items
  - Covers ages 1.5 to 18 years
  - Reasonable cost
- **Brief Infant-Toddler Social Emotional Assessment**
  - 42 parent-reported items
  - Covers ages 12 months to 26 months
  - Reasonable cost

*All three  
have good  
validity  
and  
reliability*



# Child Maltreatment

---

- **Number of child abuse/neglect reports**
- **Involvement in the child welfare system**
- **Number of foster care placements**

*All to be collected through administrative data*



# Cross-Site Support for Outcome Evaluations

---

- **Provided sample consent language**
- **Obtained and shared select instruments**
- **Conducted training on:**
  - Obtaining consent
  - Conducting random assignment
  - Collecting data
  - Reducing sample attrition
- **Created analytic tools**
  - For power calculations
  - For assessing data quality
- **Gave one-on-one assistance to evaluators**



# Evaluation Design

---

- **We also reviewed grantees' proposed evaluation designs**
  - Examined whether proposed designs could address the research questions grantees proposed
  - Worked with grantees and their evaluators to strengthen designs where possible
- **Descriptive studies were most appropriate in many cases**
  - Helped formulate realistic expectations for what could be learned from non-experimental designs



# Sample EBHV Grantee Outcome Evaluations Under Way

State	Model	Design
CO	SafeCare	RCT* for parents on probation compared to normal probation services
NY	PAT	RCT* testing enhanced model vs. PAT “as usual”
	NFP	Comparison group design testing enhanced model vs. NFP “as usual”
OH	HFA	RCT* testing enhanced model vs. HFA “as usual”
OK	SafeCare	Regression discontinuity + RCT of cultural and linguistic adaptations to SafeCare for Latino families
		Randomized factorial design to compare two service approaches for SafeCare

\* Randomized controlled trial



# There Are Evaluation Challenges

---

- **Grant structure, and changes in funding certainty and source, have**
  - Increased the difficulty of implementing as planned some local outcome evaluations
  - Led to cancellation of the outcomes domain of the cross-site evaluation
- **In some sites, enrollment has been unexpectedly slow**
- **Both positive experiences and challenges suggest implications for other initiatives**



# Implications

---

- **Interest in and support for evidence-based interventions is growing at all levels**
  - Policymakers and providers express support
  - Understanding what it really means takes time
- **Large initiatives provide opportunities to test home visiting with new populations and to hone our understanding of implementation**
- **Those delivering different models can agree on a common set of outcomes if given the opportunity to develop consensus**



# Participatory Approaches have Benefits

---

- **Shared measurement is a middle way between independent and mandated evaluation designs**
  - “Helped us learn how to do a more rigorous evaluation”
  - “Improved my understanding of specific measures and analytic issues”
  - “Opened communication among local evaluators”
    - Providing expertise to one another
    - Sharing data and dissemination
- **Can help establish an ongoing community of research**
  - For example: Early Head Start Research Consortium



# But There Are also Costs

- Participation requires time and resources
- Working out the “ladder of participation” (Arnstein 1969) can be frustrating

Control	<b>Power</b>
Delegated Power	
Partnership	
Placation	<b>Tokenism</b>
Consultation	
Informing	
Therapy	<b>Nonparticipation</b>
Manipulation	



# (Not Everyone Wants to Participate 😊)

Control	<b>Power</b>
Delegated Power	
Partnership	
Placation	<b>Tokenism</b>
Consultation	
Informing	
Therapy	<b>Nonparticipation</b>
Manipulation	
<b>GO AWAY DON'T BOTHER ME!</b>	



# Incentives for Rigorous Evaluation Vary

---

- **Defining “evidence” is itself controversial**
- **Rigorous evaluations have the biggest payoff but can also be the most burdensome**
- **Rigor may be the first thing to go when budgets are cut**
- **States are anxious to collect required benchmarks and performance measures, and with assistance and time can do so!**



## For More Information:

---

- Debra Strong [dstrong@mathematica-mpr.com](mailto:dstrong@mathematica-mpr.com)
- Kim Boller [kboller@mathematica-mpr.com](mailto:kboller@mathematica-mpr.com)
- Melissa Brodowski [melissa.brodowski@acf.hhs.gov](mailto:melissa.brodowski@acf.hhs.gov)
- Publications:
  - <http://www.mathematica-mpr.com/EarlyChildhood/evidencebasedhomevisiting.asp>
  - <http://www.supportingebhv.org/crossite>
- Arnstein, Sherry R. "A Ladder of Citizen Participation." *Journal of the American Institute of Planners* 35 (4), July 1969, 216-224