

# Measuring Benchmarks: Indicators and Tools

Carlos Cano, MD, MPM

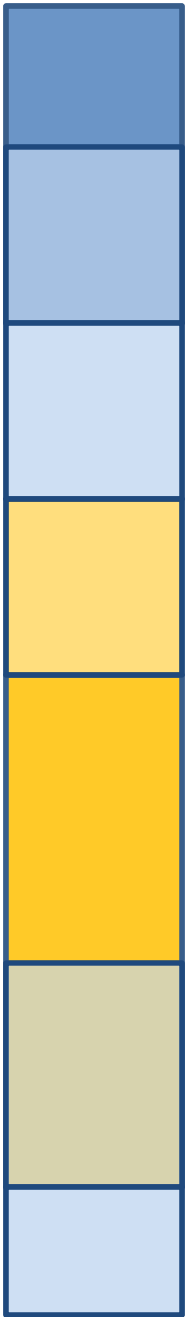
*Health Resources and Services Administration*

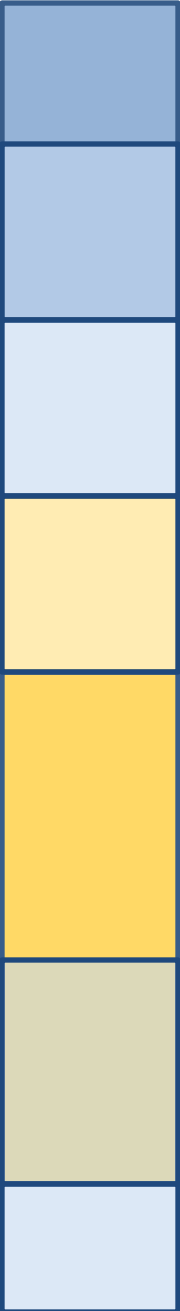
Susan Zaid, MA

*James Bell Associates, Inc.*

Mary Klute, Ph.D.

*Clayton Early Learning Institute*





# DOHVE: Design Options for Maternal, Infant, and Early Childhood Home Visiting Evaluation

- Working with US-DHHS to support the federal Maternal, Infant, and Early Childhood Home Visiting Program
- DOHVE:
  - ✓ Design options for a federal evaluation of evidence-based home visiting programs
  - ✓ Evaluation-related Technical Assistance (TA) for “promising approaches”
  - ✓ TA for grantees’ continuous quality improvement, Management Information Systems (MIS), and benchmarks



# DOHVE Evaluation TA Team

- James Bell Associates (JBA)
- MDRC
- Cincinnati Children's Hospital Medical Center & Every Child Succeeds



# Goals of Session

- Provide an overview of legislative requirements for measuring & reporting benchmarks
- Issues to consider when identifying data sources & instruments to assess participant outcomes
- Using data to guide services and quality improvement efforts – a grantee example: Clayton Early Learning Institute (Denver, CO)



# SELECTING BENCHMARK INDICATORS AND TOOLS



# Benchmarks in Legislation

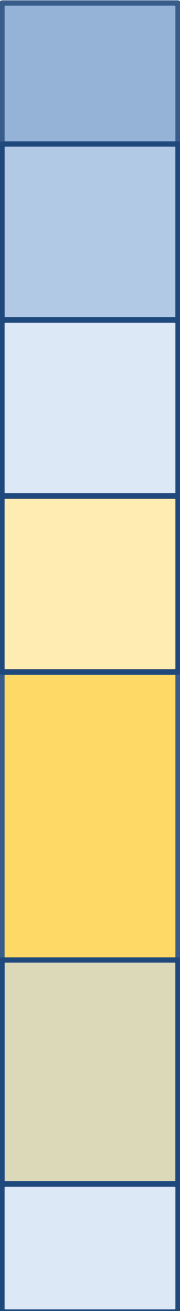
- The grantee “establishes, subject to the approval of the Secretary, quantifiable, measurable benchmarks for demonstrating that the program results in improvements for the eligible families participating in the program in each of the following areas...”



# Data Collection on 6 Benchmark Areas

- Maternal and newborn health
- Child injuries; child abuse, neglect, or maltreatment; emergency department visits
- School readiness and achievement
- Crime or domestic violence
- Family economic self-sufficiency
- Coordination and referrals for other community resources and supports

# Framework for Reporting

- 
- Report on:
    - All benchmark areas
    - All constructs under each benchmark area
    - Individual vs. community-level data
  - Flexibility:
    - Grantees choose measures
    - Define improvement/targets
  - Templates for annual reporting:
    - Will be made available later



# What is Improvement?

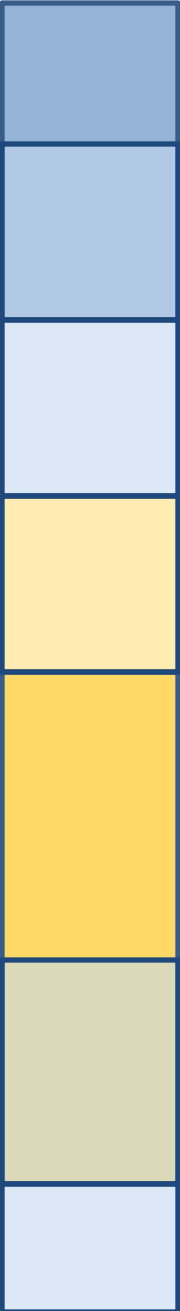
- 
- Real but feasible
    - Defined by each grantee
    - Half of the constructs
  - Benchmarking is distinct from evaluation
    - Performance monitoring
  - Intent
    - Avoid unreasonable burden
    - Invitation to take ownership of activity
      - Programmatic purpose: continuous quality improvement (CQI)
    - Joint effort
      - Opportunity for state leadership



# State Plan Must Include:

- Proposed measures including reliability/ validity of measure
- Proposed definition of improvement
- Proposed data collection and analysis plan
  - Justification of appropriateness of the measures
  - Who will be measured
  - How often the measures will be collected
  - Discussion of training for data collection and analysis
  - Discussion of how data will be analyzed
  - Sampling plan (if proposed)
  - Discussion of how benchmarks will be used in CQI plan
  - Discussion of privacy, data safety and monitoring

# An Opportunity for State Leadership

- 
- Long-term horizon
    - Experimentation
    - Scenarios for the program in 5-10 years
  - Towards a “system of systems” for early childhood (EC)
    - HV models form a coherent program
      - Centralized intake or common assessment tool
      - Core measures across models
    - Integration of HV with other EC programs
      - Governance: e.g., Early Childhood Advisory Council and Early Childhood Comprehensive Systems grants
      - Place-based initiatives

# Joint Federal-State Effort

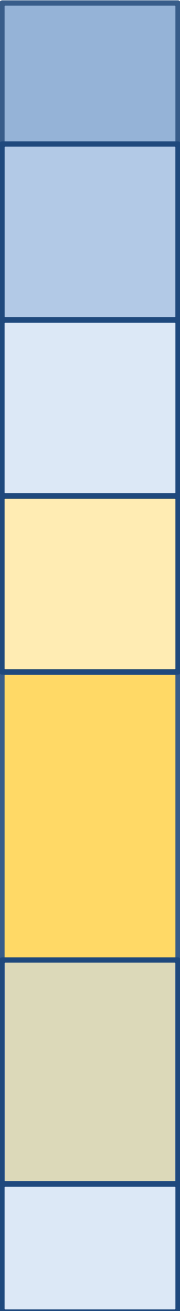
- 
- Federal-State partnership
    - Updated state plan
      - Dialogue during review process
  - Technical assistance
    - Multifaceted, ongoing
      - Webinar, compendium of measures, individualized
  - Facilitate State-to-state partnerships
    - Peer learning networks



# Benchmark Requirements in Summary

- All States/territories must include a plan for collecting benchmark data in their Updated State Plan
- Plan should include information about each construct and the measure selected to assess each construct for each benchmark area
- States must propose a measure with a description including reliability/validity of measure along with analysis plan
- Constructs must be measureable and must be assessed at different points in time to measure change

# Data Collection Methods

- 
- Use multiple data collection methods
    - Self report by client
      - Survey
      - Interview
      - Focus groups
    - Home visitor collection
      - Direct observation
    - Abstraction of administrative data



# Data Collection Sources

- To determine the best data source, consider:
  - *What source is likely to provide the most accurate information?*
  - *What source is the least costly or time consuming?*
  - *Will collecting information from a particular source pose an excessive burden on that person?*

(Program Manager's Guide to Evaluation, 2010)



# Data Collection Sources

- You must decide what data source would best capture how you define the construct
- Take your population needs/characteristics and other contextual factors into consideration when selecting data source (measuring emergency room visits – self-report vs. medical records)
- Self-report may be easier to collect but less reliable; medical records may be more accurate but involve more time and bureaucratic channels



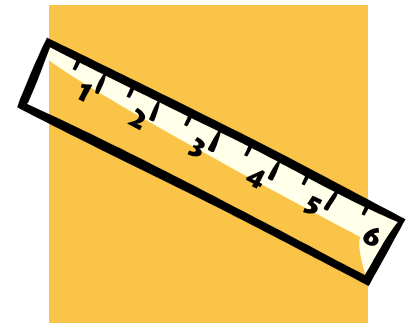


# Defining Constructs: Using Quantifiable Objectives

- **Resource:** *Healthy People 2020* at <http://www.healthypeople.gov/2020/topicobjectives2020>
  - Provides examples of how to clearly define constructs
- **Construct:** Prenatal Care
- **Objective:** Increase the **rate** of pregnant women served by the program who receive prenatal care in **the first trimester** by **10%** from **year 1** to the **3-year** benchmark reporting period
- **Possible measurement sources:** interview, self-report surveys, administrative records

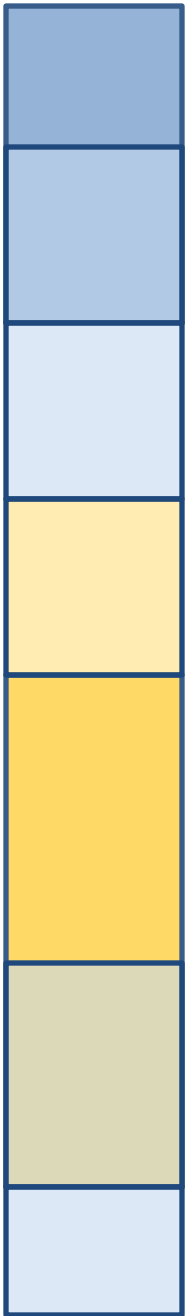
# Considerations in Choosing Specific Measurement Tools

- Standardization of measure
- Training requirements
- Reliability
- Validity
- Sensitivity
- Specificity
- Cost: both in dollars and time
- Utility of scores for staff
- Appropriateness to population

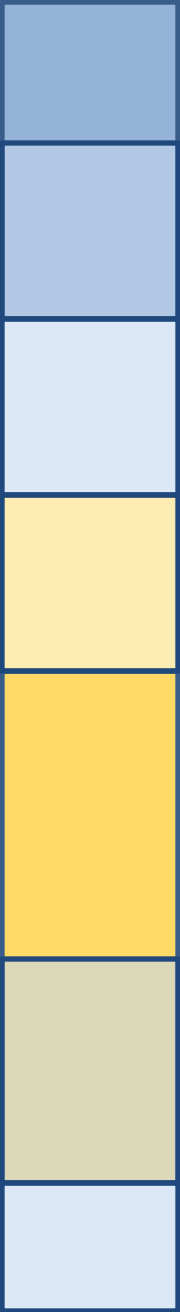


# Standardization of Measurement Tools

- What is a standardized measure?
  - Designed in such a way that the questions, conditions for administering, scoring procedures, and interpretations are consistent
  - Administered in same manner to all subjects all the time
  - Has an established protocol for scoring and interpretation of the results
  - Has been used with a large population, so that we know how children/parents typically score
- One advantage of standardized measures is the ease in interpreting the scores



# Standardization of Measurement Tools

- 
- If you are using a standardized measure, it is important to consider the normed group your population will be compared to and if the findings are generalizable to the population in your program
  - Often inadequate for describing subgroups; may not be appropriate for all populations
  - Does using a standardized measure guarantee good results?
    - A standardized test administered under non-standard conditions – results can be meaningless. Example: timed test administered under untimed conditions; 1 year follow-up completed at 16 months.



# Training Requirements

- Some measurement tools have very specific training requirements before staff can administer the measure
- Some measurement tools do not
  - In-house training
  - Create internal data-collection protocol
- Consider periodic booster trainings – “refreshers”



# Reliability of Measurement Tools

- Is the instrument reliable?
  - A reliable measure is both consistent and stable at measuring a construct
  - The results are repeatable
  - Example: child development screening tool



# Validity of Measurement Tools

- Is the instrument valid?
  - A valid measure is one that measures the concept it was intended to measure.
  - Did you answer the question you intended to answer?
  - Example:

**Objective:** To increase the social-emotional development of children birth to 5 years served by the program from baseline to 1 year post enrollment

**Measure:** A scale that assesses physical development (reliable, but not valid)

# Reliability and Validity

– What's the difference?



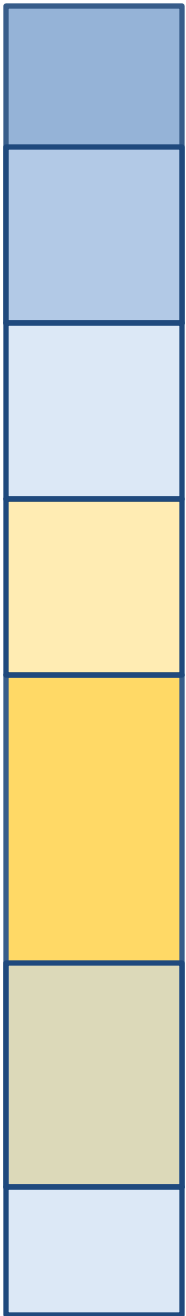
by Experiment-Resources.com

Shuttleworth & Martyn (2008). Validity and Reliability. Retrieved 2/18/11 from Experiment Resources: <http://www.experiment-resources.com/validity-and-reliability.html>.



# Sensitivity & Specificity of Measurement Tools

- Sensitivity of the instrument
  - The degree to which an instrument correctly identifies those individuals who have a specific condition
- Specificity of the instrument
  - The degree to which an instrument correctly “screens out” those individuals who do not have a specific condition



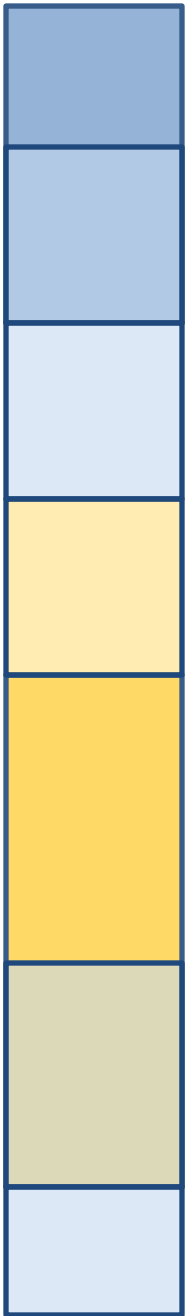


# Costs Associated with Measurement Tools

- What are the financial costs?
  - Purchase of the instrument
  - Training people to use it
  - Data collectors
    - Internal staff activities or external staff/evaluator
- What other costs are associated with using the tool?
  - How much time will be spent implementing and using this tool
  - Can it replace another, less optimal tool

# Utility of Scores for Staff

- Integrating benchmark data with program CQI
- Ideally, benchmark data should not be collected only to assess the program and include findings in a report
- Data should be made available and reviewed ongoing to provide feedback at all staff levels as to how effective the program is with each family (i.e., what areas need improvement, where have changes been made, etc.)



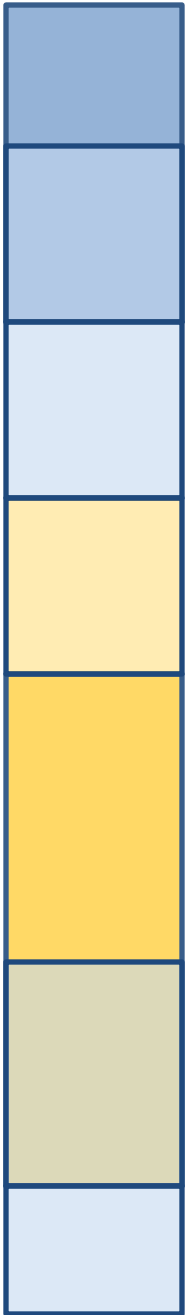
# Appropriateness of Measurement Tools

- Is the tool appropriate for the children/family participating in your program?
  - Culturally appropriate
  - Developmentally/age appropriate
  - Language
  - Literacy levels



# Appropriateness of Measurement Tools

- Considering the appropriateness of a tool for your target population can arguably be the most important factor to consider when selecting a measure
- Example – even if a measure is reliable and valid, if it is at a literacy level higher than that of most participants, how much confidence would you have in the results?





# Compendium of Measurement Resources

- The DOHVE TA team is compiling a compendium of measurement resources relevant to home visiting
- Not exhaustive
- Can serve as a tool to use along with other available resources
- Lists measures by domains and subdomains
- Check the DOHVE website for new resources:  
[http://www.mdrc.org/project\\_12\\_104.html](http://www.mdrc.org/project_12_104.html)



# Creating a Research-Program Partnership: Lessons Learned from the Bounce Learning Network Implementation Study

Mary Maguire Klute





# About Educare

- Educare Schools serve high-risk children, birth to age 5, and their families
- All Educare Schools are part of the Bounce Learning Network
- Provide high-quality center-based care using blended public and private funding
- Educare model has 12 core features ([www.educareschools.org](http://www.educareschools.org))
  - One of them has to do with using research-based strategies





# Bounce Learning Network

## Implementation Study

- Twice yearly standardized assessments of children
- Birthday-related assessments for 2 and 3 year olds
- Annual parent interviews
- Annual classroom observations
- Annual staff surveys
- Exit interviews with parents of kindergarten-bound children





# Lessons Learned

- **Minimize burden**
  - Incorporate study procedures into existing structures as much as possible
  - If possible collect data that can serve multiple purposes





# Lessons Learned

- **Maximize the benefit** for the people most affected by the data collection. This really helped us move staff from tolerating the study to being *invested* in the study.
  - Share data with staff in a way that helps them do their work
  - Schedule data sharing at a time that makes sense for staff



# Lessons Learned

- **Negotiate a data-sharing schedule** and be open to modifying it continually
  - Who needs to know?
  - What do they need to know?
  - How often do they need to know it?
  - Who can share the data?



# Lessons Learned

- **Make use of all types of information.**
  - How do benchmark results fit with other sources of information (formal and informal)?
  - Avoid conversations about which source of information is right and which is wrong.
  - Make examining program data alongside evaluation data part of the data sharing plan.



# Lessons Learned

- **Focus on the journey instead of the endpoint.**
  - View pitfalls as opportunities not failures.
  - Reflect often on how far you've come.





# Recent Webinars

- Building a culture of quality in home visiting—  
January 13, 2011
- Designing and Using an Effective Data  
Management System: Components and  
Considerations- February 24, 2011
- All webinar slides and recorded sessions are  
available at:  
[http://www.mdrc.org/project\\_12\\_104.html](http://www.mdrc.org/project_12_104.html)



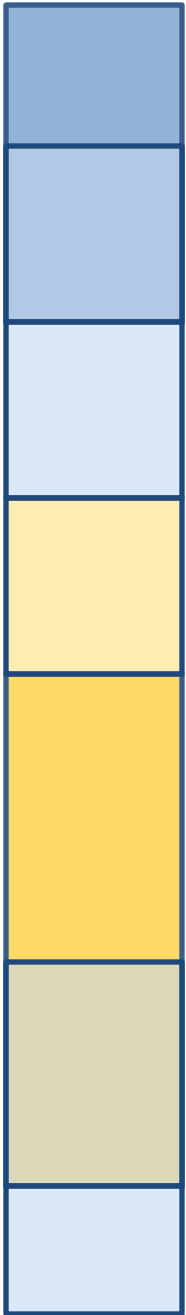
# Next Steps

Stay tuned for additional webinars, individualized TA, and other information from the DOHVE Evaluation TA team on:

- ✓ Compendium of Measures
- ✓ Written Brief on Selecting Measurement Tools
- ✓ Measurement tool for tracking information on referrals and coordination
- ✓ Webinar – April 14th (3:00–4:30 ET)  
*Rigorous Evaluations with Small Sample Sizes*
- ✓ And more...



# Questions and Comments





# For more information...

Carlos Cano, Health Resources and Services Administration  
*ccano@hrsa.gov*

Lauren Supplee, Administration for Children and Families  
*lauren.supplee@acf.hhs.gov*

Susan Zaid, James Bell Associates  
*szaid@jbassoc.com*

Mary Klute, Clayton Early Learning Institute  
*mklute@claytonearlylearning.org*