2011 STAM – CCPRC Meeting
Workshop E3
November 17, 2011, 2:00-3:30 pm

QRIS Validation: A Focus on Program Standards

Description

This session started with presentation of a set of new concepts concerning the validation of program quality standards within QRIS. A summary of findings from a multi-site case study was presented that describes various approaches to the validation of program standards. Finally, a closer look at one State's efforts at validating a QRIS was presented. A facilitated discussion provided opportunities for others to share their ideas and efforts at QRIS validation. Critical issues for discussion included the following: What is the "right time" to implement a validation study? What are the differences between validation and evaluation studies? How much child care program enrollment is "enough" in order to validate components of a State-wide or region-wide QRIS? What are some of the benefits and challenges of including the study of child and or family level outcomes in validating a QRIS?

Presenters

Gail Zellman, RAND Corporation
Michel Lahti, University of Southern Maine
Karen Ruprecht, Purdue University

Scribe

Rebecca Starr, Child Trends

1. **Documents in Session Folder**
   - "Validation of Quality Rating and Improvement Systems for Early Care and Education and School-Age Care," Gail Zellman and Rick Fiene
   - "Validating Standards in Child Care QRIS: Multi-State Case Study," Michel Lahti, Carolyn Langill, Teri Sabol, Tabitha Isner, Rebecca Starr, Kathryn Tout
   - "QRIS Validation: Findings from Indiana's Paths to QUALITY Evaluation Study," Jim Elicker, Carolyn Langill, Karen Ruprecht, Joellen Lewsader, and Treshawn Anderson

2. **Summary of Presentations**
   - **Summary of Presentation #:** Gail Zellman
     - Gail discussed the validation of QRIS for ECE and school-age care. Validation is coming to the forefront in a way it hasn't before and is now part of Race to the Top-Early Learning Challenge. A brief is being finalized about this topic.
     - What is validation?
       - An ongoing process that assesses the degree to which evidence and theory support conclusions from assessments conducted in a specified context.
       - For QRIS, validation studies assess whether rating components and summary ratings can be relied on as accurate indicators of program quality.

- Validation also addresses whether key system components are operating as expected, a more systemic look at validation. Are the things that are going on consistent with the underlying logic model?
  - How does validation differ from evaluation?
    - Validity: The focus is on measurement tools: Do they measure what they say they measure? Goal is to improve measures and other system features.
    - Evaluation: The focus is on whether a specific intervention is effective in reaching objectives. Some QRIS evaluations may also address validity questions, for example, the finding that programs improve quality suggests that QRIS is working.
  - Why should policymakers care about validation? Can lead to increased support for QRIS (if higher-rated programs are better); helps to effectively deploy limited resources; guides appropriate use of quality improvement resources; and helps avoid legal challenges. In high-stakes situations, providers may question ratings.
  - What does it mean to validate a QRIS? It's a complex iterative process that relies on multiple sources of evidence; and expert judgments of the degree to which measures capture key components.
  - Psychometric properties of measures used to assess quality: assess whether component measures and ratings perform as claimed and expected by theory (i.e., ECERS-R).
  - Assess outputs of the rating process: distribution of ratings in relationship to other quality measures; are 4-star programs providing higher care than 3-star programs? Do distributions vary by program type? Are cut scores and combining rules producing appropriate distributions? Do higher rated programs produce better learning outcomes for children?
  - Approaching validation with a plan: develop a plan before implementation; it may help with design stage; some data can be collected as part of ratings or other QRIS activities; include multiple approaches to validation.
  - The brief will provide guidance to developing the plan for validation, keeping in mind: approach, timing, cost issues, and getting by (i.e., can rely on other states' efforts in examining underlying concepts).

- **Summary of Presentation #2:** Michel Lahti
  - Michel discussed a multi-state case study that describes four State efforts in validation (Indiana, Minnesota, Maine, and Virginia).
  - The method used was to review the Compendium to identify how each State evaluates their QRIS and identify common standard areas across states; apply the four component analytic framework for cross-state analysis (concepts, measures, outputs/scores, and outcomes); validate results with State evaluators and staff; and report on ongoing threats to assessing validity.
  - This study looked at how States examine concepts of quality, examined measures used to assess quality, examined outputs or scores, and relationships between tier levels and child outcomes. Challenges were identified in each of these areas.
  - Within the four components for cross-state analysis, threats to validity were identified:

- ▪ Concepts of quality: Stakeholder lack of buy-in to definition of concepts and/or concepts defined in ways that are hard to measure.
  - ▪ Measures used to assess quality: Different data sources; little control over how administrative data is collected (QRIS a secondary use); and significant costs associated with some types of data collection.
  - ▪ Outputs or Scores: Validating to a global measure of program quality has benefits and limitations that need to be considered in design; standards are multi-faceted and can change over time; and program enrollment effects ability to assess multiple tiers of a QRIS.
  - ▪ Ratings related to appropriate outcomes: Significant costs and challenges to data collection and analytical strategies; and limited current research to guide design.

- **Summary of Presentation #3:** Karen Ruprecht
  - o Karen presented findings from Indiana's Paths to QUALITY (PTQ) Evaluation. Indiana has a voluntary system that was started in 2007. It uses a building-block approach with a rating system based on a checklist. Incentives to advancement include higher reimbursement rates for CCDF.
  - o Participation rates are high: 2,100 providers, 85% participation rate for centers.
  - o Study 1 examined concept validity and asked: are the proposed standards likely to result in higher quality care? Improve child outcomes? The researchers concluded that if proposed standards were implemented with fidelity, they should expect higher quality and child outcomes.
  - o Study 2 involved field validation: Were there differences in observed quality using the ERS from level 1 to level 4 providers? Were children in higher rated programs developing more optimally? They found that level 4 programs rated significantly higher on the ERS than those at level 1.
  - o Relating to child outcomes: outcomes were not related to levels; children from all income and education levels had access to higher-rated providers; and no strong, consistent associations were observed between the PTQ level and child outcomes.
  - o Moving forward they plan to: examine the validity of the rating system with respect to standards; look at the reliability of raters; examine ERS quality results and PTQ standards; and make recommendations for revising standards.

3. **Summary of Discussion with Presenters and Participants**
   - Questions and Comments:
     - o Need to be careful about comparing ECERS and FCCERS.
     - o Is the purpose to look at weightings, or perhaps how they hang together across levels? Gail, pilot work in California involved a correlation between ERS and CLASS.
     - o How does this compare to Q-DOT? Can't make direct comparison, but maybe that's why we aren't seeing the relations to child outcomes (Ivelisse Martinez-Beck). None of the programs are scoring above the 5 points threshold.
     - o Shannon Rudisill: The framework is helpful; have seen people defining validity in different ways. This can act as a clarifying agent; defined terms will help in carrying conversations forward. Planning validity is like having program management plan. There is a sequential nature to it.

- o Do you look at reliability? Should be included in framework. Reliability of measures, total rating score, etc.
- o Is it a problem to use ERS for validation when it is a part of the rating? What other measures could we use for validation? ERS is what is in the literature. There aren't other comparable tools out there. CLASS is becoming more common—that is good.
- o ERS is one small piece of rating. How do other aspects play into this? Gail: When measures haven't been validated, we need to collect data on the distribution of scores across levels. In Qualistar, there was no variation in the family partnerships measure.
- o What is being used to validate ratings of school-age care? School-Age Care Environmental Rating Scale and the Youth Program Quality Assessment.