Moving Beyond CHIPRA Reporting:  
The Creation of an Assessment Tool to Analyze Opportunities for Improvement

Background

As part of the Children’s Health Insurance Program Reauthorization Act (CHIPRA) Quality Demonstration Grant, Florida and Illinois are jointly working on a number of initiatives to achieve the overarching goal of the grant – to improve child health and child health outcomes. In accordance with one federal category of grant activities, Florida and Illinois have been experimenting with and evaluating the use of a federally selected Core Set of Children’s Health Care Quality Measures for Medicaid and CHIP.

Florida and Illinois have devoted substantial time and resources to calculating and reporting on the Core Set measures in each Federal Fiscal Year (FFY) of the grant, from 2010 through 2014. This has been a significant success, with both states realizing improvements and efficiencies in measurement processes and being able to expand the scope of their data reporting with each reporting cycle. This expansion in scope has included more than double the number of measures reported in both states, as well as the inclusion of a larger portion of the publicly funded population within each measure. Both Florida and Illinois have advanced to being able to report 25 of the 26 Core Set measures, including all of the measures in the original Core Set that are currently required for reporting and two of the three new core measures released in January 2013 and added to the Core Set.¹

Even while accomplishing this significant achievement in using the Core Set measures, the states began to look beyond reporting the measures to the next step. By 2012, when measure calculation and reporting processes were well established, the states started to ask what should be done with the data that was now readily available. How should the states sort through these rates and determine where quality improvement was needed? First assessments of the data showed that many areas were below national averages. Could the states tackle all of these areas? And just as relevant, should they? The states recognized the need for a way to determine focus and direction. Which of these areas should be targeted for quality improvement to ensure that the states were making the best possible use of limited state resources? As a result of these early deliberations, Florida and Illinois sought to develop an objective method for assessing the Core Set rates to determine which measures should be targeted for quality improvement.

Creation of the Assessment Tool

The tool’s primary purpose is to assist Florida and Illinois in determining how to allocate limited state resources toward the opportunities where performance improvement efforts could have the greatest impact. The tool should use objective strategies for determining where the greatest opportunities are.

The states sought to combine three key dimensions in the assessment of measures in a way that was easily understood and interpretable.

The three key assessment dimensions

- Measurement year performance compared to national benchmarks.
- Measurement year performance compared to prior year performance.
- Measure amenability to quality improvement.

Building upon available quality assessment tool examples, the states developed a matrix that combines and organizes these dimensions into a single diagram to allow the states to readily assess where the greatest opportunities for performance improvement are across the many measures of child quality.

¹ Neither state reported the Behavioral Health Risk Assessment for Pregnant Women measure.
Key Assessment Dimensions

Measurement Year Performance Compared to National Benchmarks

The first key dimension is comparison to national benchmarks. While CHIPRA Core Set benchmarks do exist for some measures, HEDIS Medicaid HMO benchmarks were chosen for their reputation as an industry standard and because they are derived from a larger data set than are CHIPRA Core Set benchmarks. For non-HEDIS measures, national benchmarks reported in the Secretary’s CHIPRA report were used.² Measures were placed into columns based on whether their performance was above average (defined as the 75th percentile or higher), below average (defined as the 25th percentile or lower) or average (between the 25th and 75th percentiles). For some measures, no national benchmarks were available for comparison. These measures were reported in the “average” column.

These columns displaying performance on each measure in relation to national benchmarks form the horizontal axis of the assessment tool.

Performance assessment tool, displaying the first dimension of assessment

<table>
<thead>
<tr>
<th>Below Average</th>
<th>Average</th>
<th>Above Average</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Measurement Year Performance Compared to Prior Year Performance

The second dimension is comparison to prior year performance. Utilizing z-tests of statistical significance, the rates were compared to the prior measurement period to determine if rates had significantly changed. For most measures, improvements and declines are defined as differences that are statistically significant at the p<.05 level. The measures were placed into rows based on whether they had significantly improved, declined or remained the same. For some measures, changes from the previous year could not be assessed – either because different measure specifications were used in the previous year and the rates were not considered comparable, or rate calculation methodology included a combination of hybrid and administrative data and denominators were not available for statistical comparison. These measures appear in the “no change” row of the matrix.

The rows displaying performance on each measure in relation to prior year performance form the vertical axis of the assessment tool (see table). Organizing measures along these two axes representing the first two assessment dimensions created a readily interpretable infographic of measures most in need of improvement. Measures in the “above average/improving” area of the infographic show measures least in need of improvement; measures in the “below average/declining” area of the infographic show measures most in need of improvement. Measures in other areas show a more moderate need of improvement, e.g., measures that are “above average/declining”, or “below average/improving.” Viewing these two dimensions in tandem gives a clear picture of which measures need improvement. Adding the third dimension, the states were able to take the assessment to the next step – of the measures that need improvement, which are most able to change?

Performance assessment tool, combining the first two dimensions of assessment

Measure Amenability to Quality Improvement

Aside from the first two dimensions of comparative performance, a number of relevant factors impact whether a measure is amenable to quality improvement – in other words, whether quality improvement efforts are likely to have an impact on measurement performance and the quality of care. Florida and Illinois evaluated a number of different factors that were relevant to whether the Core Set

² The Department of Health and Human Services Annual Report on the Quality of Care for Children in Medicaid and CHIP. (Published yearly).
measurements were amenable to change in their states. The factors determined most relevant are described below.

- **Contractual mandate**: The measure is contractually mandated to be reported by the state’s managed care organizations. Medicaid managed care plans are obligated to report on, and show strong performance in, a variety of measures. Plans have an interest in improving measures that fall within this scope. Measures that are not within the scope of mandated reporting lack a health plan requirement for high performance. As most Florida and Illinois Medicaid recipients are enrolled in managed care plans, buy-in and support by managed care plans is critical to the success of any performance improvement efforts.

- **Performance Improvement Project mandate**: There is an existing mandate that managed care plans develop a Performance Improvement Project for the measure. Several measures are selected each year by each state for federally-required Performance Improvement Projects. As such, these measures are highly amenable to performance improvement, since improvement efforts focused on these measures will coincide with existing performance improvement efforts.

- **High-Impact**: The measure impacts a high percentage of children in the state. Measures with larger populations represent greater opportunity to improve care for the states’ overall populations, so Florida and Illinois determined that measures impacting 50,000 or more children were more amenable to improvement efforts.

- **Understandability**: The measure is easily interpreted. Measures that are more straightforward better allow plans, providers and advocates to readily understand, and buy in to, what needs to be improved.

- **Administrative burden**: Assessed as three distinct factors: data collection is easy, data collection is low cost, and data collection and processing are not comparatively time consuming. These three factors of amenability to improvement decrease the burden of regular remeasurement, making frequent reassessments to check the progress of performance improvement more feasible.

- **Availability of improvement tools**: Known methods of achieving better outcomes are available. The existence of evidence-based solutions that are readily available to show stakeholders how to improve performance greatly increases a measure’s amenability to improvement.

Florida and Illinois used these eight factors to determine how amenable the Core Set measures were to improvement. Each Core Set measure was analyzed to determine how many of these eight amenability factors were present. The states defined measures that were “highly amenable to performance improvement” as measures in which the majority (i.e., at least five) of the factors were present. The measures that were determined to be highly amenable to quality improvement efforts were flagged (denoted in bold font) in the infographic, creating a single infographic that displays all three key assessment dimensions. In the table below, Measures B, C, F and G were denoted as most amenable to improvement.

**Interpreting the Assessment Tool**

The single infographic allows for a simple visual assessment of measures in need of and amenable to improvement. Bolded measures in the red, bottom left corner of the matrix (i.e., those that are below average and declining) are the most in need of and amenable to improvement. Measures that meet these requirements are flagged for further analysis for probable performance improvement opportunities. Other red and yellow categories may also be considered for potential quality improvement, depending on the scope of resources available to invest in improvement efforts. Measures in the green, upper right corner of the matrix that are above average and demonstrating improved performance have the least need for quality improvement.

Performance assessment tool, displaying the measures most in need of and amenable to improvement
Implications of the Assessment Tool

The assessment tool has fulfilled its primary purpose in assisting Florida and Illinois in determining how to allocate resources toward the opportunities where performance improvement efforts could have the greatest impact. Using the tool, each state was able to identify measures that were further targeted for performance improvement. Since its inception, the states have created a new assessment tool annually to assess the current year’s rates, identify new opportunities for improvement, and assess progress on measures previously identified. The states will be able to utilize the tool subsequent to the CHIPRA grant to continue to assess the measures that they report. Illinois has incorporated the assessment tool into its state processes for measure assessment and quality improvement going forward.

Secondarily, this tool could be useful not only to Florida and Illinois in assessing the CHIPRA Child Core Set measures, but would have broader utility too – for other performance improvement stakeholders and other measurement sets. The broader utility of the tool resides in its easy adaptability to other assessment purposes. The concept of a single, easily interpretable infographic that concurrently displays several key dimensions of assessing improvement opportunity is something that can be adapted to fit a myriad of purposes. In its current form, other states, governing bodies, health plans and providers could adopt the tool for their own assessment. Further tailoring can also easily occur to accommodate other measurement sets, other benchmarks and performance standards, and other determinations and factors of amenability to improvement. The sustainment of this tool in Florida and Illinois and its spread to other purposes more broadly will be a significant enduring accomplishment of the CHIPRA Quality Demonstration Grant.

This document was developed under grant CFDA 93.767 from the U.S. Department of Health and Human Services, Centers for Medicare & Medicaid Services. However, these contents do not necessarily represent the policy of the U.S. Department of Health and Human Services, and you should not assume endorsement by the Federal Government.

For further information, contact Jodi Bitterman at jbitterman@healthmanagement.com.