Supplements to *Evaluation Highlight No. 2*: How are States and evaluators measuring medical homeness in the CHIPRA Quality Demonstration Grant Program?

Evaluation Highlight No. 2 is the second in a series of reports that present descriptive and analytic findings from the national evaluation of the Children's Health Insurance Program Reauthorization Act (CHIPRA) Quality Demonstration Grant Program. In the *Highlight*, we discuss the measurement of medical homeness in selected demonstration projects and describe the development of the Medical Home Index-Revised Short Form (MHI-RSF), an adaptation of the short version of the Medical Home Index (MHI), for use in evaluating the demonstration projects. The full text of the *Highlight* is available on the National Evaluation of the CHIPRA Quality Demonstration Grant Program Web page.

This resource includes five supplements to the *Highlight*. First, we provide background information on two medical home measurement tools. Next, we describe the construction of a cross-state database that we will analyze for the national evaluation. We also outline the methods for collecting the qualitative data that are analyzed in the *Highlight* and present findings from the psychometric assessment of the MHI-RSF. Finally, we explore the MHI-RSF scores in each of six States providing baseline data to the national evaluation team.

1. Background Information on Medical Home Measurement Tools

The MHI is a self-assessment tool developed by the Center for Medical Home Improvement (CMHI), with Federal support from the Maternal and Child Health Bureau (MCHB).[1] The MHI allows practices to gauge their "medical homeness," identify areas for quality improvement, and compare their scores to regional or national benchmarks. The MHI consists of 25 questions that fall into 6 domains of medical homeness: organizational capacity, chronic condition management, care coordination, community outreach, data management, and quality improvement. Each theme is scored 1 to 8, with 1 representing the most basic care and 8 representing the most comprehensive care; scores are totaled and then standardized to a scale of 0-100 for ease of interpretation. The six domains align with the American Academy of Pediatrics (AAP) and the MCHB definition of a medical home. The MHI usually can be completed in less than 2 hours by a physician and a non-physician staff member working together.

The National Committee on Quality Assurance (NCQA) 2008 Physicians Practice Connections-Patient Centered Medical Home (PPC-PCMH) tool was designed and has been used widely for the purpose of recognizing practices that may then receive financial incentives for meeting requirements as medical homes. The tool includes 166 questions pertaining to 9 standards, including access and communication; patient tracking and registry functions; care management; patient self-management support; electronic prescribing; test tracking; referral tracking; performance reporting and improvement; and advanced electronic communication. The Web-based survey is completed by up to four members of a practice, and data and supporting documentation are submitted to NCQA for scoring. Completion of the survey and necessary documentation takes an estimated 40-80 hours.[2] The 2011 NCQA PCMH recognition tool updates the 2008 standards and includes approximately 150 questions pertaining to 6 standards, including promoting access and continuity, planning and managing care, identifying and managing patient populations, providing self-care support, tracking and coordinating care, and measuring and improving performance.[3]

2. Constructing a Cross-State Database

With the goal of performing a cross-state analysis of the impacts of medical home interventions on quality of care for children, the national evaluation team considered the design of each State's demonstration project, the potential for obtaining claims and administrative data from each State, and the State's choice of medical home measurement tool. To build a database that could support a cross-state analysis, the national evaluation team worked with the demonstration States to determine whether they would be able to contribute appropriate claims and administrative data on Medicaid and CHIP children. The team also had a preference for including data from States with a comparison group design. Thus, in early discussions, several States were excluded from the cross-state database for one or more reasons, and ultimately six States (Illinois, Maine, Massachusetts, North Carolina, South Carolina, West Virginia) agreed to provide claims data to the national evaluation team.

A key challenge in constructing a cross-state database pertains to collecting comparable medical home measures across States. Among the States contributing claims data to the analysis, three were using the NCQA tool to measure medical homeness (Illinois, Maine, South Carolina), two were using the MHI (Massachusetts, North Carolina), and one State was using a combination of multiple tools (West Virginia).This led to concerns that, without a consistent method for measuring medical homeness across States and practices, a rigorous cross-state impact analysis would not be feasible. The national evaluation team therefore asked selected States to collect supplemental data on medical homeness using the MHI-RSF. Ultimately, six States (Alaska, Maine, Massachusetts, Oregon, South Carolina, West Virginia) collected the MHI-RSF in their intervention and/or comparison practices to assist the national evaluation team. As a result of these efforts by demonstration States, the national evaluation team's options for cross-state analysis were improved considerably. Table 1 summarizes these data collection details for each State.

Table 1. Baseline Data Collection in the CHIPRA Quality Demonstration Grant Program Medical Home Initiatives

State	Providing claims data to national evaluation team	NCQA data collection	MHI data collection	Data included in *Evaluation Highlight #2*
Illinois	Yes	NCQA 2011	No	No
Maine	Yes	NCQA 2008	MHI-RSF	No
Massachusetts	Yes	No	MHI/MHI-RSF	Yes
North Carolina	Yes	No	MHI	Yes
South Carolina	Yes	NCQA 2011	MHI-RSF	Yes
West Virginia	CHIP only	NCQA 2011 self-assessment	MHI-RSF	Yes
Alaska	No	NCQA 2011 self-assessment	MHI-RSF	Yes
Florida	No	No	MHI	No
Idaho	No	No	MHI	No
Oregon	No	NCQA 2011 self-assessment	MHI-RSF	Yes
Utah	No	No	MHI	No
Vermont	No	NCQA 2008/2011	No	No

Notes: Massachusetts is collecting the MHI in intervention practices and the MHI-RSF in comparison practices. Idaho and Utah are collecting components of the MHI and the Clinical Microsystem Assessment Tool. Florida and Maine will be providing MHI and MHI-RSF data as they become available.

3. Qualitative Data Collection Methods

The national evaluation team gathered information on the medical home assessment tools chosen by each State and the reasons for their choices through a series of conversations with State officials and other stakeholders during the first year of the demonstration. During the planning phase, presentations were made to the States by the NCQA, developers of the MHI, and the national evaluation team about the different medical home assessment tools and their potential pros and cons for States and evaluation purposes.

In addition to these early discussions, the national evaluation team conducted site visits to demonstration States from March to August 2012, when States were in the early phases of project implementation, approximately 2 years into their 5-year CHIPRA quality demonstration projects. The site visits consisted of semi-structured interviews conducted with purposively selected key informants. Respondents included demonstration staff, practices participating in the demonstration, representatives from consumer advocacy organizations and professional associations, and other key stakeholders. Interviews were conducted in person, during site visits

to demonstration States. The few respondents who were unavailable for interviews on site were interviewed by phone on dates preceding or following that State's site visit. One interviewer and one note-taker participated in each interview. Interviews were audio-recorded with respondents' expressed consent. These interviews provided additional insights into how the States and practices were using the medical home assessment tools during the implementation period and additional information on their perceived strengths and weaknesses.

4. Psychometric Assessment of the Medical Home Index-Revised Short Form

Creating the Medical Home Index-Revised Short Form

The MHI-RSF was developed as a low-burden medical home assessment tool for the national evaluation of the CHIPRA Quality Demonstration Grant Program. The tool modifies the Center for Medical Home Improvement's (CMHI) MHI-Short Version (MHI-SV) by adding four additional items from the full Medical Home Index (MHI). W. Carl Cooley and Jeanne McCallister of CMHI provided input into the modification of the MHI tool to create the MHI-RSF.

The 25 item MHI has been shown to be a reliable and valid tool to assess medical homeness at the practice level and its six domains align with the American Academy of Pediatrics and the Maternal and Child Health Bureau definition of a medical home.[4] CMHI used statistical techniques (factor analysis) on the full MHI to develop the 10 item MHI-SV as an interval measure for practices undergoing medical home transformation.[5]

The MHI-RSF includes all 10 items from CMHI's MHI-SV and adds four items from the full MHI to capture additional components of medical homeness (Table 2). Two of the added items fall under the medical home domain of Data Management, which was not included in the MHI-SV but has become more important in recent years as technology has advanced. The final two new items, Communication/Access and Family Involvement, were prioritized by CMHI as important additions.

Table 2. Domains and Topics on the MHI and MHI-RSF

MHI Domains	MHI Topics	Topic is on MHI-RSF
1. Organizational capacity	1.1 Mission of the practice	
	1.2 Communication/access	X*
	1.3 Access to medical records	
	1.4 Office environment	
	1.5 Family feedback	X
	1.6 Cultural competence	X
	1.7 Staff education	
2. Chronic condition management	2.1 Identification of CSHCN	X
	2.2 Care continuity	X
	2.3 Continuity across settings	
	2.4 Cooperative management with specialists	X
	2.5 Supporting transition to adult services	X
	2.6 Family support	
3. Care coordination	3.1 Role definition	X
	3.2 Family involvement	X*
	3.3 Child and family education	
	3.4 Assessment of needs/plans of care	X
	3.5 Resource information and referrals	
	3.6 Advocacy	
4. Community outreach	4.1 Community assessment of needs of CSHCN	X
	4.2 Community outreach to agencies and schools	
5. Data management	5.1 Electronic data support	X*
	5.2 Data retrieval capacity	X*
6. Quality improvement	6.1 Quality standards (structures)	X
	6.2 Quality activities (processes)	

Notes: CSHCN= children with special health care needs
* indicates that the topic was not on the original MHI-SV

Like the full MHI, each item on the MHI-RSF is scored 1 to 8, with 1 representing the most basic care and 8 representing the most comprehensive care. Scores are totaled and then standardized to a scale of 0-100 for ease of interpretation. In addition, overall and domain specific means (range: 1-8) are calculated.

The resulting MHI-RSF strikes a balance between comprehensively representing the domains of medical homeness as defined by the American Academy of Pediatrics (AAP) and the Maternal Child Health Bureau (MCHB), capturing all six domains in the full MHI, and being low burden for intervention and comparison practices participating in the demonstration.

Psychometric Analysis of the MHI-RSF

The MHI-RSF is an adaptation of two previously validated medical home assessment tools. To further assess the scientific properties of the MHI-RSF, we conducted descriptive and psychometric analyses on baseline data from 104 pediatric and other child-servicing practices participating as intervention or comparison practices in the CHIPRA Quality Demonstration Grant Program in six States. The evaluation team:

- Compared scores from the MHI-RSF and full MHI among the same practices.
- Calculated rank-order correlations between the MHI-RSF scores and MHI scores among the same practices.
- Calculated the internal reliability of the MHI-RSF.
- Analyzed performance on the MHI-RSF by several practice characteristics thought to be associated with medical homeness, providing evidence of the tool's validity.

MHI-RSF versus MHI Scores

At baseline, 33 practices shared data on the full MHI and 71 practices on the MHI-RSF; counts include both intervention and comparison practices. Among the practices that provided data for the full MHI, we compared the scores from the 25 MHI items to the subset of 14 items included in the MHI-RSF. Table 3 shows descriptive statistics for the overall standardized total score (standardized to 100 points for both the MHI and MHI-RSF), the overall mean score (range: 1-8), and six domain mean scores (range: 1-8) for the 25 MHI items versus the subset of 14 MHI-RSF items. Figure 1 depicts the mean scores graphically.

Table 3. Comparison Between Scores on the 25-Item MHI and the Subset of 14 Items in the MHI-RSF, Among 33 Practices that Completed the MHI

	n	Mean (SD)	Minimum	Median	Maximum
MHI- full set of 25 items					
Overall standardized total score†	33	54.90 (12.3)	26.50	52.00	88.00
Overall mean score	33	4.39 (1.0)	2.12	4.16	7.04
Domain mean scores					
Organizational capacity	33	4.39 (0.9)	3.00	4.29	6.57
Chronic condition management	33	4.67 (1.0)	2.67	4.33	7.33
Care coordination	33	4.33 (1.3)	1.33	4.00	7.33
Community outreach	33	4.02 (1.6)	1.50	4.00	8.00
Data management	32	4.81 (1.5)	1.00	4.50	8.00
Quality improvement	31	3.65 (1.7)	1.00	3.50	8.00
MHI – subset of 14 MHI-RSF Items					
Overall standardized total score†	33	56.61 (12.5)	25.89	56.25	89.29
Overall mean score	33	4.53 (1.0)	2.07	4.50	7.14
Domain mean scores					
Organizational capacity	33	4.79 (1.0)	3.00	4.67	8.00
Chronic condition management	33	4.56 (1.0)	2.50	4.50	7.00
Care coordination	33	4.65 (1.5)	1.00	4.33	7.67
Community outreach	33	3.76 (2.0)	1.00	3.00	8.00
Data management	32	4.81 (1.5)	1.00	4.50	8.00
Quality improvement	31	3.48 (1.9)	1.00	3.00	8.00

† standardized to a 100-point scale

Figure 1. Comparison Between Scores on the 25-Item MHI and the Subset of 14 Items in the MHI-RSF, Among 33 Practices that Completed the MHI

Overall and Domain Specific Mean Scores, 25 MHI Items vs Subset of 14 MHI-RSF Items
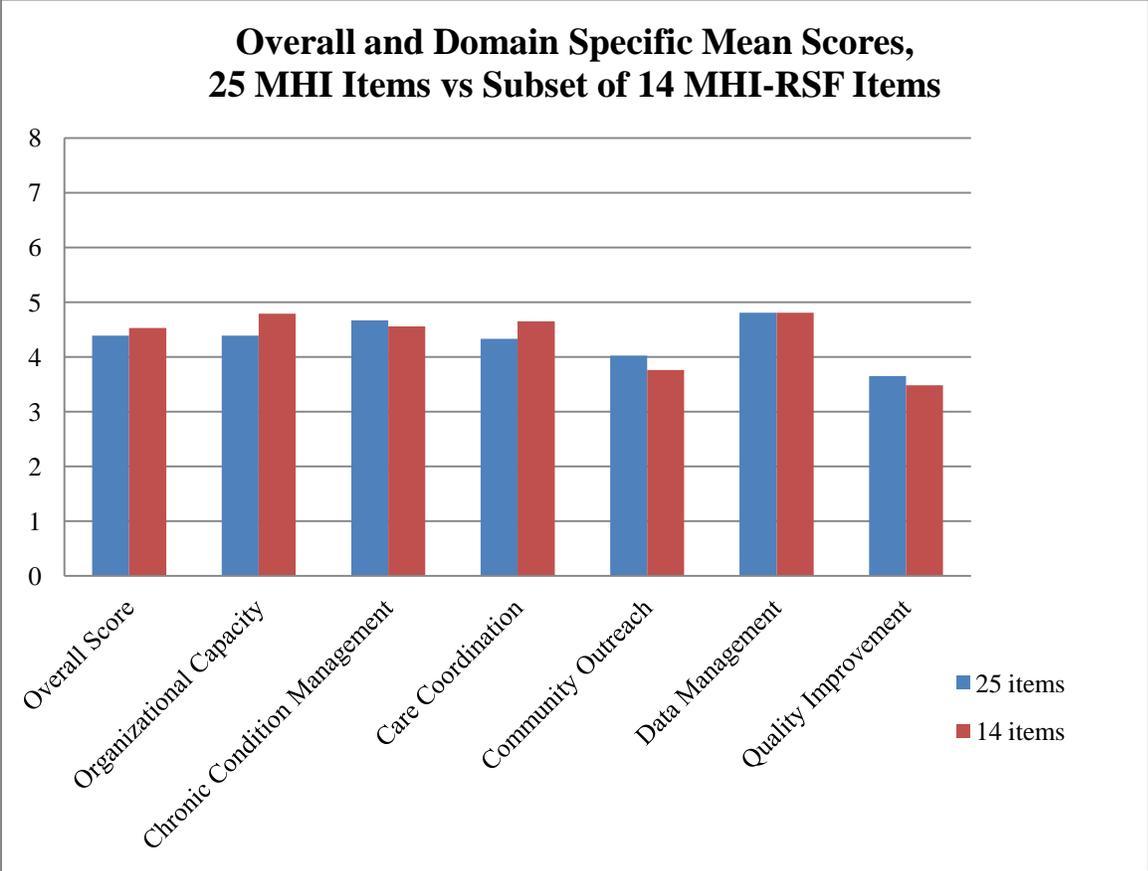
To assess whether the 33 practices rank similarly in medical homeness scores when measured by the MHI-RSF and MHI, we calculated Spearman rank-order correlation coefficients (Table 4). These correlation coefficients can range from 0 to 1. Correlations closer to 1 indicate that practices rank similarly when measured on the two tools. That is, practices that rank highest on the MHI also rank highest on the MHI-RSF. For the overall standardized total, overall mean, and domain-specific means, we observed high correlation coefficients ranging from 0.81 to 1.00. All coefficients were statistically significant. This analysis provides evidence that the MHI-RSF and MHI tools rank practices similarly on medical homeness.

Table 4. Spearman Rank-Order Correlations Between MHI-RSF Scores and MHI Scores (n= 104 practices)

	Spearman Correlation Coefficient
Overall standardized total	0.95*
Overall mean score	0.95*
Domain mean scores	
Organizational capacity	0.81*
Chronic condition management	0.94*
Care coordination	0.89*
Community outreach	0.87*
Data management	1.00*†
Quality improvement	0.88*

*$p<0.05$
† The items in the data management domain are the same for the MHI-RSF and the MHI, resulting in a perfect correlation.

Reliability

We assessed the internal reliability of the MHI-RSF among 104 intervention and comparison practices that provided baseline data on either the MHI-RSF or MHI.[6] We calculated the raw score Cronbach's alpha, a coefficient ranging from 0 to 1 that is a function of the degree of correlation among items in the tool. Higher scores indicate greater correlation among items. A Cronbach's alpha above 0.7 is considered a marker of a good scale.[7] The Cronbach's alpha for the MHI-RSF was 0.89 (Table 5). All 14 items showed moderate to strong item-to-total correlations, ranging from $r= 0.47$ to 0.72 (not shown) indicating that they are measuring different dimensions of the same construct. Inter-item correlations ranged from $r= 0.14$ to 0.73 and were statistically significant for nearly all (82 of 91) of the item pairings (not shown). Deletion of any one item did not increase the Cronbach's alpha above 0.89, suggesting that no items are adding error and no items should be cut (not shown).

Table 5. Internal Reliability of the MHI-RSF (n= 104 practices)

	Number of Items	Cronbach's Alpha
Overall internal reliability	14	0.89
Domain internal reliability		
Organizational capacity	3	0.61
Chronic condition management	4	0.70
Care coordination	3	0.83
Community outreach	1	n/a
Data management	2	0.73
Quality improvement	1	n/a

Additionally, among the 33 practices that completed the full MHI, we compared the internal reliability of the 25 MHI items (Cronbach's alpha= 0.93) to the subset of 14 MHI-RSF items (Cronbach's alpha= 0.88). Because the Cronbach's alpha is in part determined by the number of items in a scale, we expect that the internal reliability will decrease simply by reducing the number of items. The Spearman Brown Prophecy Formula can predict the Cronbach's alpha that would be expected based on subtracting a specified number of items. The Spearman Brown Prophecy Formula predicted that by cutting the 25-item MHI to the 14-item MHI-RSF, we would expect the internal reliability to drop from 0.93 to 0.89 in this subset of practices. We observed an MHI-RSF internal reliability (Cronbach's alpha= 0.88) very close to the prediction, suggesting that the difference in the reliability of the MHI-RSF and MHI is related only to the decrease in items.

In the full sample of 104 practices, we also calculated the internal reliability of the domain scores on the MHI-RSF (Table 5). Two of the MHI-RSF domains (Community Outreach and Quality Improvement) have only one item, so internal reliability is not applicable. The Cronbach's alphas of three domains (Chronic Condition Management, Care Coordination, and Data Management) were 0.7 or greater, indicating acceptable reliability; however, the internal reliability of the Organizational Capacity domain was low (Cronbach's alpha= 0.61). The three items in the Organizational Capacity domain had fair to moderate inter-item correlations (range: r= 0.24 to 0.45; all p<0.05), indicating the low internal reliability may be driven by the small number of items in the domain rather than by poor inter-item correlation.

We did not test inter-rater reliability of the MHI-RSF. See Cooley et al (2003) for information on inter-rater reliability of the full MHI.[3]

Validity

While we did not design a validity test prior to the baseline analysis of the MHI-RSF, we were able to examine indicators of validity by calculating the MHI-RSF Overall Standardized Total Score stratified by practice characteristics that are expected to be related to medical homeness. Information on these additional practice characteristics were collected on the MHI-RSF. Among

all 104 practices, the following characteristics are statistically associated with having higher MHI-RSF overall total scores: being involved in other medical home or quality improvement initiatives; having a care coordinator present in the practice; being knowledgeable about and regularly applying the concepts of the AAP medical home definition; and being knowledgeable about and regularly applying the concepts of the MCHB elements of family-centered care (Table 6). This analysis provides evidence of known-group validity of the MHI-RSF, indicating that the MHI-RSF is measuring the concept it was designed to assess.

Table 6. MHI-RSF Standardized Total Scores by Practice Characteristics Likely to be Associated with Medical Homeness

Practice Characteristic	n	Standardized Total Score Mean (*sd*)
Involvement in other medical home or quality improvement initiatives*		
No	30	46.88 (12.8)
Yes	52	55.77 (12.0)
Unknown	1	60.71 (--)
Care coordinator present*		
No	68	50.44 (11.9)
Yes	20	62.50 (13.6)
Familiarity with the AAP medical home definition*		
No knowledge of the concepts	5	46.61 (15.6)
Some knowledge/not applied	15	50.00 (11.3)
Knowledgeable/concept sometimes applied in practice	50	50.99 (11.9)
Knowledgeable/concepts regularly applied in practice	30	59.63 (12.6)
Unknown	4	58.04 (11.6)
Familiarity with MCHB elements of family-centered care*		
No knowledge of the concepts	21	50.81 (11.0)
Some knowledge/not applied	23	45.96 (10.0)
Knowledgeable/concept sometimes applied in practice	39	54.84 (13.1)
Knowledgeable/concepts regularly applied in practice	16	63.08 (11.7)
Unknown	5	56.25 (10.8)

Notes: *$p<0.05$; not all practices answered all questions on the survey, so the total number of observations across questions may vary.
AAP = American Academy of Pediatrics; MCHB = Maternal and Child Health Bureau

5. State-by-State Analysis of the MHI-RSF

In addition to the descriptive analysis presented in *Evaluation Highlight No. 2*, we explored the MHI-RSF scores by State (Table 7). The methods for selecting demonstration practices and collecting data varied across States. Each State sought to include practices that varied along key dimensions, such as size, ownership, and geographic location. Despite these efforts, the practices selected in a particular State may not reflect the mix of practices in the State as a whole. Moreover, some States have a very small number of demonstration practices. Together, these limitations indicate that the results presented below should not be interpreted as representative of a State as a whole, and comparisons across States should be interpreted with caution.

The mean standardized total score among intervention practices ranged from 51.5 in Alaska to 61.8 in North Carolina. Practice scores varied widely within each State, as indicated by the minimum and maximum total scores. While States do vary in baseline scores, the results suggest that intervention practices in all States have considerable opportunities to benefit from the PCMH interventions implemented through the CHIPRA demonstration grants.

Table 7. Distribution of MHI-RSF Standardized Total Scores Across Intervention Practices in Six States

State	Number of Intervention Practices	Mean (SD)	Minimum	Median	Maximum
Alaska	3	51.5 (14.5)	40.2	46.4	67.9
Massachusetts	13	53.0 (8.6)	40.2	50.9	72.3
North Carolina	12	61.8 (13.6)	46.4	58.5	89.3
Oregon	8	54.1 (6.1)	44.2	53.6	66.1
South Carolina	17	58.6 (12.6)	43.8	55.4	91.1
West Virginia	10	52.1 (10.7)	32.1	53.1	67.0

Note: Overall standardized total scores are standardized to a scale of 1-100. SD is standard deviation.

Endpoints

[1] Additional information can be found at:
http://www.medicalhomeimprovement.org/knowledge/practices.html.

[2] Burton R, Devers K, Berenson R. Patient-centered medical home recognition tools: A comparison of ten surveys' content and operational details (2nd edition). Baltimore, MD: U.S. Centers for Medicare & Medicaid Services, 2012.

[3] Additional information available at http://www.ncqa.org/tabid/631/Default.aspx. A self-assessment version of the tool can be found at http://www.pcdc.org/resources/patient-centered-medical-home/pcdc-pcmh/pcdc-pcmh-resources/PCDC-PCMH/ncqa-2011-medical-home.html.

[4] Cooley WC, McAllister JW, Sherrieb K, Clark RE. The Medical Home Index: Development and validation of a new practice-level measure of implementation of the medical home model. Ambul Pediatr 2003; 3(4):173-180.; American Academy of Pediatrics, Medical Home Initiatives for Children with Special Needs Project Advisory Committee. The medical home. Pediatrics 2002; 110:184-186.

[5] Center for Medical Home Improvement. Medical Home Index- FAQ. Available at http://www.medicalhomeimprovement.org/pdf/FAQ_Measurementrev11-09.pdf. Accessed 11/1/12.

[6] For practices that completed the full MHI, we used only the subset of 14 items from the MHI-RSF.

[7] Nunally JC, Bernstein IH. Psychometric Theory (3rd edition). New York: McGraw-Hill; 1994.