Recommended Initial Core Set of Children's Healthcare Quality Measures for Voluntary Use by Medicaid and CHIP Programs

A Report to the Agency for Healthcare Research and Quality National Advisory Council on Healthcare Research and Quality

by

The Agency for Healthcare Research and Quality National Advisory Council Subcommittee on Children's Healthcare Quality Measures for Medicaid and CHIP Programs

October 21, 2009

Executive Summary

Title IV of the Children's Health Insurance Program Reauthorization Act (CHIPRA; Public Law 111-3) required the Secretary of the U.S. Department of Health and Human Services (HHS) to identify and post for public comment by January 1, 2010, an initial, recommended core set of children's health care quality measures for voluntary use by Medicaid and Children's Health Insurance Programs (CHIP), health insurance issuers and managed care entities that enter into contracts with such programs, and providers of items and services under such programs.

This report presents a brief summary of the processes used to identify an initial core set of children's healthcare quality measures for recommendation to the AHRQ National Advisory Council on Healthcare Research and Quality, the measures recommended, and next steps to be taken by AHRQ and CMS. **Table 1** below summarizes the recommended core measure set.

The initial core measure set includes one or more measures for almost all of the health care topics and criteria specified in the legislation. Quality measures are recommended for services to prevent disease and promote health, and to treat and manage a spectrum of acute and chronic conditions experienced by children, including physical, mental, and dental disorders. The measure set includes measures designed to assess family experiences of care and availability of services. Measures address services provided across the age continuum and in both the outpatient and inpatient settings. All but four of the measures are supported by evidence for a relatively high level of validity. The validity ratings for the others are supported by substantial professional consensus.

There were, however, a number of legislative topics for which currently available, valid, and feasible measures could not be identified, and some legislative criteria that could not be met. These include measures of the "most integrated health care delivery settings," more valid measures of availability of services, and importantly, a core measure of duration of enrollment and coverage for use in quality reporting and surveillance. In addition, neither the recommended measure set as currently specified, nor the body of measures in use by Medicaid, CHIP and others, currently meet the CHIPRA goals of identifying disparities by race and special health care needs status or measuring and improving quality across all enrollees in Medicaid and CHIP programs. Thus, additional work is needed to develop measures and specifications to meet these challenges, and to provide technical assistance to the Medicaid and CHIP programs and the plans and providers on whom they rely to deliver high quality care.

Table 1.

SNAC-Recommended Initial Core Set of Children's Healthcare Quality Measures, SNAC Priority Score, Rank by SNAC Priority Score, and, By Legislative Measurement Category, September 17-18, 2009

Internal AHRQ/SNAC Control #	LEGISLATIVE TOPIC AREA/ Subtopic/Brief Measure Label	SNAC Priority Score*	Ranking by SNAC priority score*
	PREVENTION AND HEALTH PROMOTION		
	Prenatal/Perinatal		
PHP-2	Frequency of ongoing prenatal care (NCQA measure)	54	2
PHP-1	Timeliness of prenatal care (NCQA measure)	43	9
PHP-26A	% of live births weighing less than 2,500 grams	43	10
PHP-38	Cesarean Rate for Low-risk First Birth Women	28	16
	Immunizations		
PHP-5	Immunizations for 2 year-olds (NCQA measure)	63	1
PHP-6	Adolescent immunization (NCQA revised for 2010)	45	7
	Screening		
PHP-19A	Body Mass Index (BMI) documentation 2 - 18 year olds (NCQA measure)	52	4
PHP-33	Rates of screening using standardized screening tools for potential delays in social and emotional development	43	11
PHP-12	Chlamydia screening 16-20 year-old females (NCQA measure)	30	14
	Well-Child Care		
PHP-9	Well-Child Visits (WCV)-three NCQA measures: 1) WCVs in the First 15 months of life; 2) WCVs in the third, fourth, fifth and sixth years of life; 3) Adolescent WCV	50	5
	Dental		
PHP-43	Total eligibles receiving preventive dental services (EPSDT measure Line 12B)	49	6
	MANAGEMENT OF ACUTE CONDITIONS		
	Upper Respiratory -- Appropriate Use of Antibiotics		
AC-2	Pharyngitis - appropriate testing (NCQA measure)	20	18
AC-17	Otitis Media with Effusion - avoidance of inappropriate use of systemic antimicrobials	20	18
	Dental		
AC-3	Total EPSDT eligibles who received dental treatment services (EPSDT CMS Form 416 Line 12C)	26	17

Internal AHRQ/SNAC Control #	LEGISLATIVE TOPIC AREA/ Subtopic/Brief Measure Label	SNAC Priority Score*	Ranking by SNAC priority score*
	ED		
AC-10	Emergency Department Utilization - Average number of emergency room visits per member per reporting period	54	2
	Inpatient		
AC-4	Pediatric catheter associated blood stream infection rates (PICU and NICU)	26	17
	MANAGEMENT OF CHRONIC CONDITIONS		
	Asthma		
CC-19	Annual number of asthma patients (> 1 year-old) with > 1 asthma related ER visit (S/AL Medicaid Program)	53	3
	ADHD		
CC-2	Follow-up care for children prescribed attention-deficit/hyperactivity disorder (ADHD) (Medication Continuation and Maintenance Phase – NCQA measure)	39	12
	Mental Health		
CC-29	Child and adolescent Major Depressive Disorder (MDD) - suicide risk assessment	36	13
CC-5	Follow up after hospitalization for mental illness (NCQA measure)	33	14
	Diabetes		
CC-13	Annual hemoglobin A1C testing (all children and adolescents diagnosed with diabetes)	36	13
	FAMILY EXPERIENCES OF CARE		
FEC-5 &1	HEDIS CAHPS 4.0 including supplements for children with chronic conditions and Medicaid Plans	44	8
FEC-6	Use of Clinician & Group primary care CAHPS survey for practitioners participating in Medicaid and CHIP (CAHPS family of measures)	28	16
	AVAILABILITY OF SERVICES		
AS-2 (new label)	Annual dental visit (NCQA measure)	36	13
AS-1	Access to primary care practitioners, by age and total	28	16

*** SNAC Voting and Calculation of Priority Scores:** After deliberations and voting on day one of the September meeting, 31 measures remained under consideration. On day two, there were three rounds of voting where SNAC members could vote for their top 20 measures out of the 31 that remained. In round one, SNAC members voted for their top 10 measures; in round two their next 5 measures; and in round three their last 5 measures respectively. Measures voted for in the first round received 3 points per vote, measures voted for in the second round received 2 points per vote, and measures voted for in the third round received 1 point per vote. The Priority Score represents the total points assigned to that measure by SNAC members after three rounds of voting.

Background

The Legislation

Title IV (Section 401(a)) of the Children's Health Insurance Program Reauthorization Act (CHIPRA; Public Law 111-3; February 3, 2009) amended Section 1139 of Title XI (42 U.S.C. 1301 et seq.) by adding a new section 1139A on Child Health Quality Measures. Section 1139A called for the Secretary of the U.S. Department of Health and Human Services (HHS) to "identify and publish for general comment an initial, recommended core set of child health quality measures for use by State programs administered under titles XIX and XXI, health insurance issuers and managed care entities that enter into contracts with such programs, and providers of items and services under such programs."

The legislation called for identification of "existing quality of care measures for children that are in use under public and privately sponsored health care coverage arrangements, or that are part of reporting systems that measure both the presence and duration of health insurance coverage over time."

Further, measures were asked to be identified for the following topics, although others could be included: duration of enrollment and coverage; preventive and health promotion services; treatment and management for acute and chronic conditions in children; family experiences of care, most integrated health care settings; and availability of services. CHIPRA also calls for evidence-based measures and measures that can identify disparities in health care quality by race and ethnicity, socioeconomic status, and special health care need. The measures are to be published no later than January 1, 2010.

AHRQ/CMS Partnership

In response to this legislative directive, the Agency for Healthcare Research and Quality (AHRQ) and Centers for Medicare and Medicaid Services (CMS) signed a Memorandum of Understanding in April 2009 giving AHRQ leadership responsibilities for identifying the initial core set, working in very close partnership with CMS. CMS has the authority for implementation of all CHIPRA provisions.

Methods

The initial core set of Children's Healthcare Quality Measures for Voluntary use by Medicaid and CHIP Programs was developed using a transparent and evidence-informed process, informed by broad input from multiple stakeholders. Key components included multiple opportunities for public comment including a CMS-led listening session for Medicaid and CHIP officials; an AHRQ NAC Subcommittee that contributed expertise on validity, feasibility, and importance of measures in use; and supportive background work by AHRQ, CMS, and members of the CHIPRA Federal Quality Workgroup.

Creation of the AHRQ National Advisory Council on Healthcare Research and Quality Subcommittee on Children's Healthcare Quality Measures for Medicaid and CHIP Programs (SNAC)

In May 2009, the AHRQ Director approved a Charter creating the Agency for Healthcare Research and Quality's National Advisory Council for Healthcare Research and Quality (AHRQ NAC) Subcommittee on Children's Healthcare Quality Measures for Medicaid and CHIP Programs (SNAC). The AHRQ NAC had agreed to provide advice to AHRQ and CMS to facilitate their work to recommend an initial core set of measures of children's health care quality for Medicaid and CHIP programs. To provide the requisite expertise and input from the range of stakeholders identified in the CHIPRA legislation, the NAC established the Subcommittee on Children's Healthcare Quality Measures for Medicaid and CHIP Programs (SNAC).

The SNAC was charged with: a) providing guidance on criteria for identifying an initial core measurement set; b) providing guidance on a strategy for gathering additional measures and measure information from State programs and others; and c) reviewing and applying criteria to a compilation of measures currently in use by Medicaid and CHIP programs to begin to select the initial core measurement set. SNAC recommendations were to be provided to the NAC, which in turn advises the Director of AHRQ.

Nominations for SNAC members to represent the range of stakeholders were sought from CMS and the CHIPRA Federal Quality Workgroup. An emphasis was placed on identifying Medicaid and CHIP officials because of their unique role as potential implementers of the initial core set. Although more were invited, 4 State Medicaid program officials (from Alabama, Minnesota, Missouri, District of Columbia), and 1 State CHIP official were able to participate as SNAC members. Others represented Medicaid, CHIP, and other State programs more generally (i.e., representatives of the National Academy on State Health Policy, National Association of State Medicaid Directors, and the Association of Maternal and Child Health Programs).

Representatives of health care provider groups came from the American Academy of Family Physicians, American Academy of Pediatrics, American Board of Pediatrics, the National Association of Children's Hospitals and Related Institutions, the National Association of Pediatric Nurse Practitioners, and a Medicaid health plan representative. The interests of families and children were represented by the March of Dimes. Individual SNAC members provided expertise in children's health care quality measurement, children's health care disparities, tribal health care, dental care, substance abuse and mental health care, adolescent health, and children's health care delivery systems in general. Two members of the NAC also participated in the SNAC. SNAC members are listed in the Appendix.

The SNAC Co-chairs Rita Mangione-Smith, MD, MPH and Jeffrey Schiff, MD, MBA were selected because of their expertise in children's health care quality measurement and leadership roles in the Medicaid Medical Directors Learning Network, respectively. The SNAC charter expires December 31, 2009. [3, 4]

The SNAC held two public meetings (July 22-23 and September 17-18, 2009) and accomplished a substantial amount of work outside the meetings in order to help the NAC, AHRQ, CMS, and the Secretary meet the CHIPRA legislative deadline of January 1, 2010.

Public Input

Multiple ongoing opportunities for public input were provided as part of this process. In June 2009, AHRQ established a website to provide information on its role in CHIPRA implementation, in close collaboration with CMS, and an email address through which the public could comment on the process. In addition, both SNAC meetings were open to the public and provided opportunities on each day for anyone to make formal public comments. Additional opportunity for public comment came during the July 24, 2009 NAC meeting at which the SNAC Co-chairs presented on the process used and results of the July 22-23, 2009, SNAC meeting.[5] In addition, the SNAC co-chair, Dr. Schiff, arranged for a conference call for members of the Medicaid Medical Directors Learning Network (MMDLN) to seek input on the measure identification and nomination process. Several members of the MMDLN responded by nominating children's health care quality measures in use by their States for consideration for the initial core measure set. Finally, on September 30, 2009, CMS led a listening session for Medicaid and CHIP officials to provide an opportunity for comment on the initial, recommended core measure set.

Those making public comments through these mechanisms included individual health care practitioners, additional Medicaid and CHIP programs, representatives of industry groups, child and family advocates, and members of the CHIPRA Federal Quality Workgroup.

First SNAC Meeting July 22-23, 2009

The first SNAC meeting was held July 22-23, 2009, in Washington, DC. The meeting was open to the public. This section describes preparation for the first SNAC meeting, the focus of SNAC discussions, presentations to the SNAC, refinements to methodology made during the meeting, and the identification of a preliminary group of measures to further consider for inclusion in the final core set, as well as needs for additional information and work.

Preparation

AHRQ and CMS staff and the subcommittee Co-chairs began conferring prior to the first scheduled SNAC meeting. Seventy-seven measures in use by Medicaid and CHIP programs were identified by AHRQ staff with the assistance of CMS and a process to initially evaluate those measures was agreed upon by AHRQ and CMS.

Prior to the July meeting, SNAC Co-chairs, working through AHRQ, provided subcommittee members with standard definitions and criteria recommended for use in evaluating the validity and feasibility of quality measures. SNAC members were asked to apply these evaluation criteria to the 77 measures using the RAND Corporation's modified Delphi process.[1] Previous work has shown this method of evaluating quality measures to be reliable and to have content, construct, and predictive validity in other applications.[2-4]

The modified Delphi process involved individual SNAC members scoring the initial identified set of Medicaid and CHIP quality measures for validity and feasibility on a 1 to 9-point scale (with 1 denoting the measure was not valid or feasible and 9 indicating it was definitely valid and feasible). Objective information (e.g., on underlying scientific soundness of

the measures) related to both measure validity and feasibility was provided to the extent it was available. However some measures were scored in this round without adequate identification of numerators, denominators, or measure specifications. Measure specifications are essential for evaluating feasibility. Instructions to the SNAC for Delphi I noted that scores for validity could be guided by professional consensus when published evidence to support the measure's validity was insufficient.

The RAND modified Delphi method outlines cut-points for passing scores on validity and feasibility. For validity, the median passing score used is more stringent, i.e. 7-9 on the 9-point scale, than the median passing score for feasibility which requires a median score of 4-9 to pass. The rationale for this difference is that for validity, either the evidence exists to support the measure or it does not which results in relatively objective information being available to make this assessment. Feasibility is a more subjective assessment than validity. Some Medicaid or CHIP programs may find a measure quite feasible to implement (due to their infrastructure, amount of available funding, etc) while others will not.

Median scores and a display of the distribution of scores across voting members were calculated and prepared for SNAC review by AHRQ staff prior to the July meeting. The median scores summarized the individual scores of SNAC members on these two domains (i.e., validity and feasibility). The median scores and the display of distribution across voting SNAC members were presented at the July SNAC meeting and used to determine whether candidate measures would be discussed further. For the purposes of the July meeting, measures with a median validity score of 6 or 7 and a median feasibility score of ≥ 4 were discussed by the SNAC. Measures with a validity score of 6 or 7 were selected for discussion as these measures were deemed controversial and in need of further consideration by the group.

SNAC Meeting July 22nd-23rd, 2009

The SNAC spent most of the first day reviewing the criteria for validity and feasibility; identifying criteria for importance; and discussing the measures that were deemed "controversial" after Delphi Round 1, i.e., measures with a median validity score of 6 or 7, median feasibility of ≥ 4, and a relatively wide distribution of scores across members, suggesting little consensus among the group. Forty-five of 77 measures met these criteria. On the second day, the SNAC heard presentations by experts commissioned by AHRQ and CMS to provide further input into the overall process.

Additional input and discussion: Presentations to SNAC and the participating public

At the July 22-23, 2009, SNAC meeting, members and the public present at the meeting heard several presentations and engaged in discussions with presenters. Presentations by the AHRQ Director, Carolyn Clancy, CMS's Director of the Center for Medicaid and State Operations (CMSO), Cindy Mann, and Director of the Division of Evaluation, Quality and Health Outcomes in CMSO, Barbara Dailey, set the stage for the meeting. The AHRQ Director provided the charge to the SNAC and the CMSO Director expressed a strong desire for the SNAC to recommend a grounded and parsimonious core set which could be implemented voluntarily by State programs, health plans, or provider groups.[6, 7] Representatives of the National Quality Forum, the National Committee on Quality Assurance, and the Center for

Health Care Strategies spoke on the challenges of implementing health care quality measures for children.

In addition, several experts who had been asked to write federally-supported white papers on specific aspects of measurement in the legislation presented their early thoughts about their work. These experts addressed the charges to them of conceptualizing and assessing the validity, feasibility, and importance of measures of mental and behavioral health care, family experiences of care, duration of enrollment and coverage, availability of services, and the "most integrated health care setting." AHRQ and CMS also asked that papers be prepared analyzing data sets of the National Academy for State Health Policy, Health Management Associates, and the Child and Adolescent Health Measurement Initiative (CAHMI) database from the 2007 National Survey on Children's Health. An additional environmental scan of Medicaid and CHIP websites to identify additional children's health care quality measures that may have been missed in the first effort by AHRQ staff and CMS had also been commissioned. Not all authors could participate in the July SNAC meeting. All presentations are included in the transcript of the July meeting posted at http://www.ahrq.gov/chip/chipraact.htm. (to come).

Refinements to methodology

During the July meeting the SNAC agreed upon refinements to the methodology to be used for future rounds of the modified Delphi process. Importance was added as a third domain to consider when evaluating potential measures in addition to validity and feasibility. The SNAC worked to establish consensus on the criteria to use to rank the importance of measures under consideration. To be considered important at least some of the following criteria had to be met by the measure. The criteria are listed in order of decreasing weight as determined through a voting process by SNAC members on July 23, 2009:

1. The measure should be **actionable**. State Medicaid and CHIP programs, managed care plans, and relevant health care organizations should have the ability to improve their performance on the measure with implementation of quality improvement efforts.
2. The **cost** to the nation for the area of care addressed by the measure should be substantial.
3. Health care systems should clearly be **accountable** for the quality problem assessed by the measure.
4. The **extent** of the quality problem addressed by the measure should be substantial.
5. There should be documented **variation** in performance on the measure.
6. The measure should be **representative** of a class of quality problems, i.e., it should be a "sentinel measure" of Quality of Care (QOC) provided for preventive care, mental health care, or dental care, etc.
7. The measure should assess an aspect of health care where there are known **disparities**.
8. The measure should contribute to a final core set that represents a **balanced portfolio** of measures and is consistent with the intent of the legislation.

9. Improving performance on measures included in the core set should have the potential to **transform care** for our nation's children.

Similar to feasibility, the threshold for a passing score on importance was also set at ≥4 on the 9-point scale as this was felt to be the most subjective of the three evaluation domains.

The SNAC members were asked to score each of the measures that had passed the first round of Delphi scoring for validity and feasibility on the new criterion of importance. AHRQ staff then summarized these scores using the median value. Measures were considered to pass the importance criterion if the median score was ≥ 4.

The refinement process further involved reviewing, discussing and reaching consensus on criteria the SNAC would use to evaluate the validity and feasibility (including reliability) of candidate measures that would be considered for potential inclusion in the recommended core set.

Other steps and decisions

The SNAC's discussion of controversial measures resulted in the recommendation that further information related to measure validity, feasibility and importance would be needed prior to further consideration of these controversial measures. The SNAC asked AHRQ staff to obtain that information.

During their July deliberations, the SNAC also determined that a call for nominations of additional pediatric quality measures in use (either within or outside of the Medicaid and CHIP programs) should be used to identify a larger set of measures to consider for the final core set.

SNAC members expressed a strong desire to recommend a grounded and parsimonious core set of measures that could be implemented voluntarily by State programs, health plans, and provider groups, and agreed on a target number of no more than 25 measures. The SNAC acknowledged that such a core set would be incomplete, but efforts would be made to balance the set to accomplish the legislative goals and the goals articulated in the SNAC discussion of measure importance. The SNAC agreed to bring forth to the NAC's attention measures not accepted into the core set and aspects of child health for which current measures do not exist.

Conclusions

By the end of the July SNAC meeting, SNAC members had identified a preliminary list of 24 measures that had clearly passed criteria for validity and feasibility in the first round of Delphi scoring and also passed scoring for importance using the criteria agreed to by the SNAC at the July meeting. This preliminary list of measures is available at the AHRQ CHIPRA website as part of the SNAC Co-chairs presentation to the NAC on July 24 (see below).[5] The Co-chairs made clear that this preliminary group of measures would be subject to further research by the AHRQ staff as needed and included in the second round of Delphi scoring prior to the September SNAC meeting. In addition, SNAC members were invited to nominate additional measures for consideration.

First SNAC Report to the NAC

The SNAC Co-chairs reported to the NAC immediately after the July meeting (on July 24, 2009).[5] This presentation included a review of the SNAC-refined criteria for the measure evaluation domains (validity, feasibility, and importance) as well as the preliminary list of 24 measures passing all three domains after the initial round of Delphi scoring. The SNAC report is available at http://www.ahrq.gov/chip/chipraact.htm. [6]

Second SNAC meeting September 17-18, 2009

The SNAC held its second meeting on September 17-18, 2009 in Washington, DC. In addition to being open to public participation on site, the meeting was Webcast. The technology allowed for greater participation and public comment. A link to the Webcast is available at http://www.connectlive.com/events/ahrq2009/.

Preparation for the Meeting

Additional Measure Nominations

Shortly after the July meeting, the AHRQ staff in collaboration with the SNAC Co-chairs developed a measure nomination template. This template was created in order to collect a standardized set of information on all measures nominated for potential inclusion in the core set (see Appendix – ***Nomination Template***). The nomination template was made available in early August 2009. Nominations were accepted until August 24, 2009. In addition to measure nominations by SNAC members, public nominators included members of the Medicaid Medical Directors Learning Network, the American Medical Association Physician Consortium for Performance Improvement, the National Partnership for Women and Families, and the Child and Adolescent Measurement Initiative on behalf of The Commonwealth Fund. Additional nominations were obtained through e-mail to the AHRQ public comment e-mail address. CHIPRA Federal Quality Workgroup nominations also came from CMS and HRSA.

In addition to all newly nominated measures, each measure that either 1) passed Delphi round one or 2) was considered controversial by the SNAC during their first meeting in July was entered into the measure template, with required information, by AHRQ staff. Authors of the CHIPRA-commissioned papers also recommended measures for consideration and additional sources of data for quality measurement based on their works in progress. Measures recommended by the contractors included a measure of medical home (for "most integrated health care setting") using items from the Healthcare Effectiveness Data and Information Set (HEDIS) Consumer Assessment of Healthcare Providers and Systems (CAHPS®) surveys; a preliminary measure of availability also using items from the HEDIS CAHPS®; and measures of duration of enrollment based on work done by researchers primarily using Medicaid and CHIP enrollment data. In addition, one of the works in progress focused on the type of data (e.g., race/ethnicity) and measures that could be obtained from the Medicaid Statistical Information Statistics (MSIS).

At a minimum, nominators were asked to identify the measure numerator and denominator; measure specifications; and current use of the measure. Substantial effort was put into obtaining all of the information requested in the template for every measure under

consideration. The nominators entered information into the nomination template. Each template was then supplemented with additional information where necessary by AHRQ staff and the SNAC Co-Chairs. Through this work, a standardized set of information was made available for almost all measures for consideration by the SNAC members during their second round of Delphi scoring. One-page summary sheets that abstracted information from the measure nomination templates were provided for each measure under consideration (see **Appendix –One Page Summary Template**).

By mid-September 2009, the SNAC had 121 measures to consider during a second modified Delphi process.

Delphi II scoring by the SNAC

Using a second modified Delphi scoring process prior to the September meeting but including the SNAC-identified criteria for importance (see **Appendix – Instructions for Delphi Round 2 for AHRQ SNAC Members**), SNAC members selected 65 of the 121 measures as meeting criteria for validity, feasibility, and importance. As in Delphi I, SNAC members were instructed to use professional consensus on the underlying scientific soundness of the measures in cases of insufficient published evidence.

SNAC September meeting deliberations

As at the second SNAC meeting, members first heard opening remarks from the Directors of AHRQ and CMSO, and an overview of the meeting agenda and process.[8] Unlike the first meeting, there were no invited presentations (other than during public comment periods on Days 1 and 2). Due to time constraints and the need to identify for NAC consideration a reasonable core set of measures near the SNAC's target number of 25, the initial plan was to only discuss and consider the 65 measures that passed the second modified Delphi scoring process as candidates for the core set. However, initial discussions at the September 17-18, 2009, SNAC meeting resulted in adding back 5 measures that did not strictly pass the second Delphi round (i.e., those with high median feasibility and importance scores [$\geq$ 7] and median validity scores of 6 or 6.5 rather than the cutoff of 7) to the list of measures to be discussed and voted on during the meeting. Thus, 70 of the 121 measures scored in Delphi round two were discussed and considered for the core set. **Table B** of the **Appendix** provides a list of nominated measures that did not meet the criteria threshold for validity during the Delphi II scoring process and were not discussed at the September meeting.

Electronic voting process

Throughout the one and a half-day meeting in September, a method of electronic confidential voting was used extensively by SNAC members. This method was chosen because in small groups some members may dominate a discussion, leading to group decisions that do not reflect the true sense of the group membership.[5] Through private electronic voting, the SNAC process was most likely to obtain the candid individual preferences of members, accumulating to a consensus of the SNAC.

Balancing measures across multiple domains

The SNAC reviewed and prioritized measures based on several characteristics pertaining to legislative and feasibility criteria, including: data source (administrative, medical record, HIT,

survey); site of care (primary care, specialty care, inpatient, emergency, mental health, substance abuse, dental); measure type (outcome, process, structural); care continuum (screening, prevention, diagnosis, treatment, care coordination); accountable entity (state program, health plan, provider); child ages to which the measure applied; and availability of data to report disparities.

Elimination of multiple overlapping measures, merging of some measures within specific categories, and voting

On day one of the meeting, SNAC members engaged in detailed discussions of measures felt to have substantial overlap. For example, multiple measures pertaining to premature birth passed the criteria for validity, feasibility and importance, as did multiple dental measures. After discussions were completed, a series of votes was conducted which resulted in elimination of multiple measures and merging of some measures within a given category. For example, three separate well-child care visit (WCV) measures that apply to different age groups were combined into one measure for voting purposes. Similarly, multiple measures of premature birth were eliminated, narrowing measures in this area to one measure of low birth weight. Measures in each category (e.g. prevention/health promotion, care of children with chronic disease) were rank ordered within the category. Lowest scoring measures were eliminated from further consideration. This process resulted in 31 measures for final consideration on the second day of the meeting.

Getting to 25 measures to recommend to NAC

On day two of the meeting, three rounds of voting were conducted in succession. SNAC members could vote for their top 20 measures out of the 31 that remained. In round one, SNAC members individually voted for their top 10 measures; in round two their next 5 measures; and in round three their final 5 measure choices. Measures voted for in the 1st round received 3 points per vote, measures voted for in the second round received 2 points per vote, and measures voted for in the 3rd round received 1 point per vote. A priority score was then calculated for each measure which represented the total points assigned to that measure by SNAC members after the three rounds of voting. The final rank order of the measures based on priority scores was examined by the SNAC to assess how the acceptance of various cut-points (i.e. 10, 15, 20, 25 total measures) would fulfill the goal of arriving at a grounded, parsimonious, balanced core set of measures. The SNAC voted to recommend the top 25 measures on the list (see **Table 1**). **Table A** in the **Appendix** provides the list of measures that met criteria for validity, feasibility and importance during the Delphi II scoring process but were not ultimately recommended for inclusion in the core set.

Important Considerations

The SNAC did not recommend that the measures in Table 1 be implemented "as is." Rather, the group emphasized that the measure denominators should be re-specified, insofar as needed, so that the measures can be made feasible for use across all Medicaid and CHIP programs, providers, consumers, and intermediaries (e.g., health plans contracting with State Medicaid programs). For example, HEDIS CAHPS (FEC 1 and 5) as currently specified is used primarily by Medicaid Managed Care health plans that report to National Committee on Quality Assurance (NCQA). The SNAC recommended that in the future the CAHPS measures should be used by

all Medicaid and CHIP programs so that family experiences of care across a broader spectrum of covered children can be understood, compared, and, when needed, acted upon.

Additionally, the SNAC agreed that identifying the entities accountable for multi-level layers of service delivery (e.g., providing the service, facilitating the service) and for quality measure data reporting is critical. For all measures, a common duration of enrollment calculation is essential to make valid and reliable comparisons across institutions, programs, and states. In implementing these measures, however, evaluators might explore the use of standard "person-enrollment-month" methods for making these calculations, rather than simply drop enrollees who do not remain enrolled continuously for an entire assessment period.

SNAC members also emphasized that further attention to improving the capacity of measures and datasets to assess disparities is needed. Few of the proposed measures are used, at least at present, to report data that distinguish care quality by race, ethnicity, tribe, socioeconomic status, or special health care need status among children.

Finally, the SNAC recognized the critical importance of several topics that are essentially missing in the recommended set of quality assessment measures for Medicaid and CHIP purposes; they stressed the need for developing valid and feasible health care quality measures to fill these gaps. These included measures of specialty care, inpatient care, substance abuse care, mental health treatment, measures that link mainstream clinical care with other services that children receive (i.e., coordination of care), health outcome measures, and measures of the medical home.

Next Steps

A final recommended set of quality measures for children enrolled in Medicaid and CHIP programs will be posted for general comment by January 1, 2010. To that end, the CHIPRA Federal Quality Workgroup discussed the SNAC-recommended core set of measures on September 23, 2009. In addition, comments on them were solicited at a CMS listening session for Medicaid and CHIP officials on September 29, 2009.

CMS-AHRQ-recommended initial core set to the Secretary, HHS, for her consideration, and then to the White House Office of Management and Budget for its review for general comment, by January 1, 2010.

Public comments on the process of identifying the initial core measure set for voluntary use by Medicaid and CHIP programs are continuously invited via the email address at http://www.ahrq.gov/chip/chipraact.htm.

Legend:
ABCD – Assuring Better Child Health and Development
AC – Acute care
AL – Alabama
AMA – American Medical Association
AS – Access
CAHPS® - Consumer Assessment of Healthcare Providers and Systems
CC - Chronic condition
EPSDT – Early Periodic Screening, Diagnosis and Treatment
FEC – Family Experiences of Care
NCQA – National Committee for Quality Assurance
PCPI – The Physician Consortium for Quality Improvement®
PHP – Prevention and Health Promotion
S/ – State of
SNAC – Subcommittee on Children's Healthcare Quality Measures for Medicaid and CHIP Programs

APPENDIX

APPENDIX
Members of the AHRQ NAC Subcommittee on Children's Healthcare Quality Measures for Medicaid and CHIP Programs

Individual Members

Rita Mangione-Smith, M.D. M.P.H. (Co-Chair)
Associate Professor of Pediatrics
University of Washington
Seattle, WA

Anne Beal, M.D. M.P.H.
President
Aetna Foundation
Hartford, CT

James Crall, Sc.D.
Professor and Chairman
Section of Pediatric Dentistry
Director
National Oral Health Policy Center
University of California, Los Angeles (UCLA) School of Dentistry
Los Angeles, CA

Doreen Cavanaugh, Ph.D.
Research Associate Professor
Georgetown University Health Policy Institute
Washington, DC

Glenn Flores, M.D., F.A.A.P.
Professor of Pediatrics and Public Health
Director, General Pediatrics Division
Judith and Charles Ginsburg Chair in Pediatrics
University of Texas Southwestern Medical Center and Children's Medical Center, Dallas
Dallas, TX

Paul Melinkovich, M.D., F.A.A.P.
Executive Director
Denver Community Health Services
Denver Health
Denver, CO

Robert St. Peter, M.D.
President and CEO
Kansas Health Institute
Topeka, KS

Lisa Simpson, M.B., B.Ch., M.P.H., F.A.A.P.
Director, Child Policy Research Center
Cincinnati Children's Hospital Medical Center
Professor, Department of Pediatrics
University of Cincinnati
Cincinnati, OH

Phyllis Sloyer, R.N., Ph.D., P.A.H.M., F.A.A.P.
Division Director
Children's Medical Services
Florida Department of Health
Tallahassee, FL

Organizational Representatives

State Medicaid Programs

Mary McIntyre, M.D. M.P.H.
Medical Director
Office of Clinical Standards and Quality
Montgomery, AL

George Oestrich, PharmD., M.P.A.
Deputy Division Director
Missouri HealthNet Division
Jefferson City, MO

Ann Page, R.N., M.P.H.
Director
Health Care Accountability Administration
Washington, DC

Jeffrey Schiff, M.D., M.B.A. (Co-Chair)
Medical Director
Health Care Services and Medical Management Division
Minnesota Department of Human Services
St. Paul, MN

State CHIP

Cathy Caldwell, M.P.H.
Director
Alabama Children's Health Insurance Program
Montgomery, AL

AHRQ National Advisory Council on Healthcare Research and Quality (NAC)

Timothy Brei, M.D., F.A.A.P.
Riley Hospital for Children
Indianapolis, IN

Kathleen Lohr, Ph.D.
Distinguished Fellow (Health Services Research)
RTI International
Research Triangle Park, NC

American Academy of Family Physicians (AAFP)

Bruce Bagley, M.D.
Medical Director for Quality Improvement
American Academy of Family Physicians
Leawood, KS

American Academy of Pediatrics (AAP)

Jonathan Klein, M.D., M.P.H.
Associate Executive Director and Director
Julius B. Richmond Center
American Academy of Pediatrics
Elk Grove Village, IL

Xavier Sevilla, M.D., F.A.A.P.
Chair, Steering Committee on Quality Improvement and Management
American Academy of Pediatrics
Whole Child Pediatrics
Lakewood Ranch, FL

American Board of Pediatrics

Paul Miles, M.D., F.A.A.P.
Senior Vice President for Quality and Maintenance of Certification
The American Board of Pediatrics
Chapel Hill, NC

March of Dimes

Marina Weiss, Ph.D.
March of Dimes
Washington, DC

National Academy on State Health Policy (NASHP)

Alan Weil, J.D., M.P.P.
Executive Director
National Academy for State Health Policy
Washington, DC

National Association of Children's Hospitals and Related Institutions

Marlene Miller, M.D., M.Sc.
Vice Chair
Quality and Safety Initiatives
Johns Hopkins Children's Center
Alexandria, VA

National Association of Pediatric Nurse Practitioners (NAPNAP)

Linda Lindeke, Ph.D., R.N., C.N.P.
National Association of Pediatric Nurse Practitioners
Minneapolis, MN

National Association of State Medicaid Directors (NASMD)

Ann Clemency Kohler
Director of Health Services
American Public Human Services Association
National Association of State Medicaid Directors
Washington, DC

Medicaid Health Plan

Carroll A. Carlson, R.N.
Director of Government Programs
Group Health Cooperative, Eau Claire
Altoona, WI

Tribal Representative

Victoria Warren-Mears, Ph.D., R.D.
Director, NW Tribal Epidemiology Center
Northwest Portland Area Indian Health Board
Portland, OR

	APPENDIX - TABLE A
Internal SNAC/AHRQ Control #	**Measures that *Met Threshold for Validity, Feasibility, and Importance Criteria during* DELPHI II Scoring Process Ending 9/13/2009 OR that were Added Back on 9/17/2009 but were not Recommended by the SNAC on 9/18/2009 for Core Set Inclusion**
	Measure Label
	PREVENTION AND HEALTH PROMOTION ***Prenatal/Perinatal***
PHP-3	Smoking Cessation and Prevention: Pregnant women
PHP-26B	HRSA MCH Health Status Indicator #01B - % of live singleton births weighing less than 2,500 gms
PHP-26C	HRSA MCH Health Status #02A - % of live births weighing lt 1500 gms
PHP-26D	HRSA MCH Health status #02B - % of live singleton births weighing lt 1500 gms
PHP-31	MCHB National performance measure #8 - the rate of birth (per 1,000) for teenagers aged 15-17 years
PHP-32	Proportion of infants 22-29 weeks gestation treated with surfactant who are treated within 2 hours of birth
PHP-34	Health systems capacity indicator #04-% of women 15-44 with a live birth during the year whose observed to expected prenatal visits are greater than or equal to 80% on the Kotelchuck Index
PHP-36A	Rate of elective delivery prior to 39 completed weeks gestation
	Immunizations
PHP-5B	Two year old Immunization Measure - Assessing immunizations by timeliness and the ACIP/AAP/CDC schedule versus HEDIS dose counting (OR)
	General Screening
PHP-17	Newborn hearing screening
PHP-18	Vision screening - use MEPS description and performance data
	Social/Behavioral Health Screening
PHP-33A	Administration of SDBS (Standardized screening tools social and emotional (CMWF/CAHMI))
	Dental/Oral Health
PHP-42	Total eligibles receiving any dental services (EPSDT measure Line 12A)
PHP-43A	HRSA Oral Health Measures WG Measure - % of children age 12-72 mos with 1 or more fluoride varnish apps documented
	MANAGEMENT OF ACUTE CONDITIONS ***Acute Upper Respiratory Tract Illness***
AC-1	Upper respiratory infection - appropriate treatment
	Acute Otitis Externa
AC-11	Acute Otitis Externa - Topical therapy
AC-13	Acute Otitis Externa - Systemic antimicrobial therapy-avoidance of inappropriate use
	Otitis Media with Effusion
AC-14	OME Diagnostic evaluation - assessment of tympanic membrane mobility

	APPENDIX - TABLE A
Internal SNAC/AHRQ Control #	**Measures that *Met Threshold for Validity, Feasibility, and Importance Criteria during* DELPHI II Scoring Process Ending 9/13/2009 OR that were Added Back on 9/17/2009 but were not Recommended by the SNAC on 9/18/2009 for Core Set Inclusion**
	Measure Label
AC-18	OME - systemic corticosteroids - avoidance of inappropriate use
	Inpatient Care
AC-5	Foreign body left after procedure (PDI 3)
AC-6	Iatrogenic pneumothorax in non-neonates (PDI)
AC-20	Care transitions - transition record with specified elements received by discharged patients - Inpatient
AC-23	Central line associated bloodstream infection (PDI 12)
AC-24	Accidental puncture and laceration (PDI)
AC-25	Decubitus ulcer (PDI)
AC-26	AHRQ Pediatric Quality Indicator Composite Measure (Patient Safety Composite)
	ED Care
AC-21	Care transition - transition record with specified elements received by discharged patients-ED
	MANAGEMENT OF CHRONIC CONDITIONS **Mental and Behavioral Health/Substance Use** ***ADHD Care***
CC-1	Follow-up care for children prescribed attention-deficit/hyperactivity disorder - (ADHD) medication (Initiation Phase (NCQA measure)
	HIV
CC-23	HIV AIDS Bureau Measure - % of clients with HIV infection who had 2 or more CD4 T-cell counts performed in the measurement year
CC-24	HRSA HIV/AIDS QPR % of clients with HIV who had 2 or more medical visits in an HIV setting in the measurement year
CC-8A	HRSA HIV/AIDS Bureau quality performance measure - % AIDS who are prescribed HAART
	Asthma
CC-10	Asthma -- appropriate medications
CC-10B	Use of Appropriate Medications for People 5-20 years of age with Asthma – Average number of member controller months
CC-18	Annual influenza vaccination (all children and adolescents diagnosed with asthma)
CC-20	Annual number of asthma patients (> 1 year-old) with > 1 asthma related hospitalization

	APPENDIX - TABLE A
Internal SNAC/AHRQ Control #	**Measures that *Met Thresholds for Validity, Feasibility, and Importance Criteria during* DELPHI II Scoring Process Ending 9/13/2009 OR that were Added Back on 9/17/2009 but were not Recommended by the SNAC on 9/18/2009 for Core Set Inclusion**
	Measure Label
	Diabetes
CC-14	Annual lipid profile (adolescents with diabetes ≥ 16 years-old)
CC-16	Annual eye examination (adolescents with diabetes > 16 years-old)
CC-17	Annual influenza vaccination (all children and adolescents diagnosed with diabetes)
	End Stage Renal Disease
CC-33	Pediatric End stage renal disease - Plan of care for inadequate hemodialysis
CC-34	Pediatric End stage renal disease - influenza immunization
	DURATION OF ENROLLMENT/COVERAGE
D-2	Retrospective duration measure

Source: AHRQ, based on SNAC Delphi II scoring and September 17-18, 2009, SNAC meeting.

	APPENDIX - TABLE B **Measures that *Did Not Meet Criteria Thresholds during* DELPHI II Scoring on Validity, Feasibility, and/or Importance**
Internal SNAC/AHRQ Control #	**Measure Label**
	PREVENTION AND HEALTH PROMOTION
	Prenatal/Perinatal
PHP-1A	HRSA MCH MPR #18 - % of infants born to pregnant women receiving prenatal care beginning in the first trimester
PHP-26	Percentage of low birth weight (PDI --hospital discharge data)
PHP-27	Postpartum care visit NCQA measure
PHP-30	MCHB National performance measure #17 - percent of VLBW infants delivered at facilities for high-risk deliveries and neonates
PHP-30A	Under 1500g infant Not Delivered at Appropriate Level of Care
PHP-36	Elective delivery prior to 39 weeks gestation
PHP-40	HRSA MCH NPR MEASURE #11 - % of mothers who breastfeed their infants at 6 mos of age
PHP-41	HRSA MCH NPR #15 - % of women who smoke in the last 3 mos of pregnancy
	Immunizations
PHP-8	Immunization reporting (provider registry)
	Adolescent Preventive Services
PHP-15	Adolescent receipt of the following seven components of care during the measurement year: BMI percentile, assessment/counseling/education on nutrition, physical activity, risk behaviors associated with sexual health/activity/preventive actions, depression screening
PHP-16	Smoking Cessation and Prevention: adolescent tobacco users
	General Screening
PHP-19	Weight Assessment and Counseling for Nutrition and Physical Activity for children and adolescents
PHP-19B	Nutrition counseling - NCQA (fix label)
PHP 19C	Physical Activity Counseling (NQCA)
PHP-29	Lead screening rate (NCQA - 1st year measure -- see confidential data); note USPSTF rec
PHP-29A	Lead screening 2d year - % of members who turned two during the reporting period and received a lead test (data from MaineCare claims and the Maine CDC)
PHP-29B	Lead screening 1st year - 5 of members who turned one during the reporting period and received a lead test (data from MaineCare Claims and Maine CDC)
PHP-29C	HRSA Lead screening in children by 2 years of age (data source differs from NCQA)

	APPENDIX - TABLE B **Measures that *Did Not Meet Criteria Thresholds during* DELPHI II Scoring on Validity, Feasibility, and/or Importance**
Internal SNAC/AHRQ Control #	**Measure Label**
PHP-39	EPSDT - Percentage of members 0-20 years old who had one or more EPSDT procedure(s) during the reporting period
	Social/Behavioral Health Screening
PHP-37	Percent of members under age 21 with a WCC visit by any provider during the measurement period who had a BH screen.
	Dental/Oral Health
PHP-22A	HRSA oral health measure - the percentage of patients who had at least one dental visit during the measurement year (differs from NCQA - data source is CHCs; broader age range)
PHP-35	HRSA oral health measure - percentage of all dental patients with a comprehensive or periodic recall oral exam within a 12 month period
	MANAGEMENT OF ACUTE CONDITIONS
	Acute Otitis Externa
AC-12	Acute Otitis Externa - Pain Assessment
	Otitis Media with Effusion
AC-15	OME Hearing testing
	Dental
AC-8	HRSA oral health measure - percentage of all dental patients for whom the Phase I treatment plan is completed within a 12 month period
AC-9	HRSA oral health performance measure - percentage of all dental patients with a comprehensive or periodic recall oral exam, for whom the Phase I treatment plan is documented
	Inpatient Care
AC-7	PICU pain assessment on admission
AC-19	Care transitions - reconciled medication list received by discharged patients
AC-22	Care transitions - timely transmission of transition record (from inpt)

	APPENDIX - TABLE B **Measures that *Did Not Meet Criteria Thresholds during* DELPHI II Scoring on Validity, Feasibility, and/or Importance**
Internal SNAC/AHRQ Control #	**Measure Label**
	ED Care
AC-10	ER Utilization - Average number of emergency room visits per member per reporting period
AC-21	Care transition - transition record with specified elements received by discharged patients-ED
	MANAGEMENT OF CHRONIC CONDITIONS
	Mental and Behavioral Health/Substance Use
	ADHD Care
CC-6	Diagnosis of ADHD in primary care for school age children and adolescents (using DSM)
	Depression Care
CC-27	Child and adolescent Major depressive disorder (MDD)- interview of adolescent or child
CC-28	Child and adolescent Major depressive disorder (MDD)- diagnostic evaluation
CC-30	Child and adolescent Major Depressive Disorder (MDD) - psychotherapy
CC-32	Child and adolescent Major Depressive Disorder (MDD) - follow-up care
	Other Mental Health/Behavioral Care
CC-4	Initiation and Engagement of Alcohol and Other Drug Dependence Treatment (ages 3-17 and 18+)
CC-35	Medication adherence as measured by maximal gaps in days during a calendar year for ADHD and anti-psychotics medication in children
	Children with Special Health Care Needs
CC-37	CSHCN number of missed school days due to illness or injury
	HIV
CC-8	Highly active anti-retroviral treatment
	Asthma
CC-25	Percentage of patients for whom there is documentation that a written asthma management plan was provided to the patient or the patient's caregiver OR at minimum instructions on under what conditions the patient's doctor should be contacted or the patio (t
	Chronic Disease Identified through Newborn Screening
CC-12	Timely follow-up of positive newborn screens
	Diabetes

	APPENDIX - TABLE B **Measures that *Did Not Meet Criteria Thresholds during* DELPHI II Scoring on Validity, Feasibility, and/or Importance**
Internal SNAC/AHRQ Control #	**Measure Label**
CC-15	Annual urine protein screening (adolescents with diabetes > 16 years-old)
CC-21	Hemoglobin A1c test for pediatric patients
	FAMILY EXPERIENCES WITH CARE
	Other Family/Patient Experiences of Care Measures
FEC-2	Helpfulness of counseling (adolescent reported)
FEC-3	Communication and experience of care (adolescent reported)
FEC-4	Dental CAHPS (overall ratings of dentist, dental plan, dental care, office staff)
	MOST INTEGRATED HEALTHCARE SYSTEMS
MIH-2	Medical home measure using subset of HEDIS CAHPS MEDICAID 4.0 survey items
	AVAILABILITY OF SERVICES
AV-1	Unduplicated members served per provider
	Uses of Services
US-1	Utilization of ambulatory services
US-2	Outpatient drug utilization - per member per year average number of prescriptions
US-3	Utilization of inpatient care
	Access to Services
AS-1	Access to primary care practitioners, by age and total
	HEALTH STATUS
HS-2	PROMIS Pediatric item Banks: physical function, emotional distress, social role relationship, fatigue, pain and asthma
	DURATION OF ENROLLMENT
D-1	Prospective duration measure

KEY
V = Validity
F = Feasibility
I = Importance
IQR = Inter-Quartile Range
N = Number of Subcommittee members rating measure

APPENDIX – ONLINE MEASURES NOMINATION TEMPLATE

APPENDIX – INSTRUCTIONS FOR *Delphi Round 2 for AHRQ SNAC Members*

Members of the AHRQ SNAC are being asked to rate the validity, feasibility, and importance for a set of quality measures that include the following:

1) **Measures that had passing scores for validity, feasibility, and importance in Delphi Round 1.** Modified criteria for validity, feasibility, and importance were established at the first SNAC meeting on July 22nd and 23rd. In addition, substantially more information has been collected on several of these measures since the first round of scoring. As such, these measures require re-assessment by the SNAC;

2) **Measures that were judged to be "controversial" during scoring for validity and feasibility in Delphi Round 1.**

3) **Measures identified through environmental scans** but that were not included on the original list of measures scored during Delphi Round 1;

4) **Measures nominated by SNAC members, Federal partner agencies, and the public**, between July 24th and August 24th.

To accomplish this task, AHRQ staff members have provided the following materials as zip files enclosed with this mailing:

1) A Delphi Round 2 scoring sheet;

2) A one-page summary of key information related to the validity, feasibility, and importance for each measure you need to score;

In addition, they have provided access to the AHRQ SNAC Extranet at: https://ahrqsnac.webexone.com/r.asp?a=3&id=13399 where one folder of information exists for each measure and a summary "guide to the webex" has been posted for your convenience. Everyone should have received a username and password to access the extranet. If you haven't please let Denise Dougherty know as soon as possible. The extranet folders provide substantially more detailed information collected on each measure. Some folders will contain more information than others. This is dependent on how much relevant information was available or provided for each measure by those who nominated them. These folders are provided for you to review *if you feel you need more information than what was provided in the one-page summary* before you make your scoring decisions. Reviewing all of this material is *optional not required*.

Please note that each measure has been assigned a **control #,** e.g., PHP-1. This number will appear on the scoring sheet next to the measure name, on the top of the one-page summary for the measure, and on the folder for the measure found on the AHRQ SNAC Extranet.

Please score each measure according to the criteria outlined below for validity, feasibility, and importance that were established at our meeting in July. SNAC members are asked to please return their scoring sheets by **Sunday, September 13th to XXXXXXXX at the Seattle Children's Research Institute via email:** XXXXXXXXX[1]@seattlechildrens.org. If we do not receive your scores by **Midnight (PDT) 9/13/09** your scores will not be factored into the analysis we will use to determine which measures pass Delphi Round 2. We realize this is an extremely tight time frame in which to complete this task, however, this was the only way we could allow adequate time for additional measure nominations after the July meeting.

Description of Assessment Criteria Agreed upon July 22nd-23rd, 2009

Validity

Validity is the degree to which a quality measure is associated with what it purports to measure (e.g., a clinical decision support system is a measure of structure or capacity; prescribing is a measure of a clinical process; asthma exacerbations are a measure of health outcomes).

A quality measure should be considered *valid* if:

- **It meets criteria for scientific soundness:**

    1. There is adequate scientific evidence or, where evidence is insufficient, expert professional consensus to support the stated relationship between:

        - ***structure and process***[2] (e.g., that there is a demonstrated likelihood that a clinical decision support system (a structural or capacity measure) in a hospital or ambulatory office leads to increased rates of appropriate flu vaccination in the hospital or practice),

        - ***structure and outcome*** (.e.g., higher continuity of care in the outpatient setting (influenced by how appointments are organized) is associated with fewer ambulatory care sensitive hospitalizations, (e.g., hospitalizations for dehydration), or

        - ***process and outcome*** (e.g., that there is a demonstrated likelihood that prescribing inhaled corticosteroids (a clinical process) to specified patients with asthma will improve the patients' outcomes) and vice versa (e.g. that if we measure quality as a health outcome measure there is sufficient demonstrated likelihood that the outcome can be attributed to either health care delivery structures or clinical processes of care or an explicit combination of both)

[1] Name deleted to protect privacy.

[2] Structure of care is a feature of a healthcare organization or clinician relevant to its capacity to provide health care. A process of care is a health care service provided to, on behalf of, or by a patient appropriately based on scientific evidence of efficacy or effectiveness. An outcome of care is a health state of a person resulting from health care.

- **The measure itself is valid – that is, it should truly assess what it purports to measure**

Measures should be scored on a 9-point scale:

7-9 → Measure concept is scientifically sound and the measure itself is definitely valid (i.e., sufficient evidence of scientific soundness and measure validity)
4-2 → Measure concept has uncertain scientific soundness (i.e., insufficient evidence) and the measure itself has uncertain validity (may not measure what it purports to measure).
1-3 → Measure concept is not scientifically sound and the measure itself is not valid (sufficient evidence of lack of scientific soundness and invalidity of the measure itself).

Measures with a median validity rating (taking all submitted ratings into account) of 7-9 will pass and be considered in the final round of assessment at the September 17-18 meeting in Washington DC.

Feasibility
A measure will be considered feasible if:

1. The data necessary to score the measure are available to state Medicaid and CHIP programs;
2. Detailed specifications are available for the measure*
3. Estimates of adherence to the measure based on available data sources are likely to be reliable and unbiased. This allows for meaningful comparisons across states, programs, individual providers or institutional providers.
    a. Reliability is the degree to which the measure is free from random error.

Measures should be scored on a 9-point scale:

7-9 → Measure is definitely feasible
4-2 → Measure has uncertain feasibility
1-3 → Measure is not feasible

Measures with a median feasibility rating (taking all submitted ratings into account) of 4-9 will pass and be considered in the final round of assessment at the September 17-18 meeting in Washington DC.

Importance

During the SNAC meeting on 7/23, we worked to establish consensus on the criteria we would use to rank the importance of measures under consideration. To be considered important at least

some of the following criteria should be met by the measure. The criteria are listed in order of decreasing weight as determined through a voting process by SNAC members on 7/23:

- The measure should be **actionable**. States, CHIP managed care plans, and relevant healthcare organizations should have the ability to improve their performance on the measure with implementation of quality improvement efforts;

- The **cost** to the nation for the area of care addressed by the measure should be substantial;

- Health care systems should clearly be accountable for the quality problem assessed by the measure;

- The extent of the quality problem addressed by the measure should be substantial;

- There should be documented variation in performance on the measure;

- The measure should be representative of a class of quality problems, i.e., it should be a "sentinel measure" of QOC provided for preventive care, mental health care, or dental care, etc.;

- The measure should assess an aspect of health care where there are known disparities;

- The measure should contribute to a final core set that represents a balanced portfolio of measures and is consistent with the intent of the legislation;

- Improving performance on measures included in the core set should have the potential to transform care for our nation's children.

Measures should be scored on a 9-point scale:

7-9 → Measure is definitely important and meets several of the above criteria.
4-2 → Measure has an uncertain level of importance and meets some of the criteria above but fails to meet some of the criteria given higher weight by the committee (1-4 above).
1-3 → Measure fails to meet most of the criteria for importance outlined above.

Measures with a median importance rating (taking all submitted ratings into account) of 4-9 will pass and be considered in the final round of assessment at the September 17-18 meeting in Washington DC.

The Nine-Point scale

The nine point scale has been used for more than two decades at RAND in developing explicit measures for evaluating appropriateness and quality.[1] Essentially these methods require individuals who rate quality measures to place them into one of three categories (e.g., valid criterion for quality, equivocal criterion for quality, invalid criterion for quality) and each

category can be rated on a three point scale to allow for some variation within category. The scale is ordinal so that a 9 is better than an 8 and so on. Because quantities (e.g., risk-benefit ratios) are not assigned to each number on the scale, the difference between and 8 and a 9 is not necessarily the same as the difference between a 5 and a 6. Explicit ratings are used because in small groups some members tend to dominate the discussion and this can lead to a decision that does not reflect the sense of the group.[5]

For validity ratings, we use a more stringent level for the passing median score, i.e. 7-9, than we do for feasibility or importance ratings which require a median score of 4-9 to pass. The rationale for this difference is that feasibility and importance are more subjective assessments than validity. For validity, either the evidence exists to support the measure or it does not which results in relatively objective information being available to make this assessment. For feasibility, some states or CHIP programs may find a measure quite feasible to implement (due to their infrastructure, amount of available funding, etc) while others will not. Feasibility of measure implementation can also be field tested. If it is determined that a measure is less feasible to implement than initially assumed, the measure could be deleted from the core set. The importance rating is the most subjective of the three criteria and thus again, we choose to set the bar lower for the passing median score.

The Meeting on September 17th and 18th

At the meeting we will only be discussing and considering measures that pass Delphi Round 2. We will work to fill in a balancing grid that will help us to track how well we are doing in terms of selecting a set of measures that is responsive to the intent of the legislation. This will require much discussion and many rounds of voting. The panel co-chairs, Drs. Mangione-Smith and Schiff, will lead this discussion of the measures.
To facilitate the voting process, AHRQ has arranged for electronic voting to be available at the meeting. For those joining on the phone, AHRQ staff will talk with you off speaker phone and allow you to privately register your votes which they will electronically enter for you. Hopefully, at the end of this process we will have a parsimonious, balanced set of 10-25 measures that we can recommend for inclusion in the Core Set.

We want to thank you for your commitment to this important process and for taking the time to lend your expertise.

Appendix – One-page summary template

MEASURE SUMMARY
CHIPRA Core Set Candidate Measures

A. **Control #**:

B. **Measure Name**:

C. **Measure Definition**

a. Numerator:
b. Denominator:

D. **Measure Type**: ____ Process ____ Outcome ____ Structure ____ Efficiency

E. **Measure collected using:** ____ EMR ____ CPOE ____ Other HIT ____ N/A ____ NR

VALIDITY

F. **Evidence of measure validity** ____ Yes ____ No **submitted?**

G. **Level of evidence supporting the measure (**if submitted): (see Oxford University CEBM Levels of Evidence)

_____

H. **USPSTF Grade** if applicable:

_____

FEASIBILITY

I. **Measure Specifications Submitted**? ____ Yes ____ No ___ Yes, but insufficient detail provided

J. **Data Source**:

___ Admin ____ MR ____ Survey ___ Other (specify): ____ NR

K. **Evidence of measure reliability** ____ Yes ____ No **submitted?**

L. **List of <u>entity types</u> currently <u>using measure</u>:**

_____

IMPORTANCE

M. **Addresses area of care mandated in legislation?**

	Yes (specify):		No

N. **Documented variation in performance (*by race/ethnicity, language spoken, insurance type*, etc)?**

_____ Yes _____ No ____ NR

O. **Measure used/data are collected in <u>racial/ethnic populations</u> *other than non-Hispanic white?***

_____ Yes _____ No ____ NR

APPENDIX
ONLINE MEASURE NOMINATION TEMPLATE

Reference List

1. Brook RH. The RAND/UCLA appropriateness method. In: McCormick KA, Moore SR, Siegel RA, editors. Clinical practice guidelines development:methodology perspectives . Rockville, MD: Agency for Health Care Policy and Research; 1994.

2. Hemingway H, Crook A, Feder G et al. Underuse of coronary revascularization procedures in patients considered appropriate candidates for revascularization. N Engl J Med 2001;344(9):645-654.

3. Kravitz RL, Park RE, Kahan JP. Measuring the clinical consistency of panelists' appropriateness ratings: the case of coronary artery bypass surgery. Health Policy 1997;42(2):135-143.

4. Shekelle PG, Chassin MR, Park RE. Assessing the predictive ability of the RAND/UCLA appropriateness method criteria for performing carotid endarterectomy. Int J Technol Assess Health Care 1998;14(4):707-727.

5. McGlynn EA, Kosecoff J, Brook RH. Format and conduct of consensus development conferences: a multi-nation comparison. In: Goodman C, Baratz S, editors. Improving consensus development for health technology assessment. Washington, D.C.: National Academy Press; 1990.