

Systematic Evidence-Based Quality Measurement Life-Cycle Approach to Measure Retirement in CHIPRA



Denise Dougherty, PhD; Kamila B. Mistry, PhD, MPH; Olivia Lindly, MA; Maushami Desoto, PhD, MHSA; Karen Llanos, MBA; Francis Chesley, MD

From the Agency for Healthcare Research and Quality, Rockville, Md (Dr Dougherty); Office of Extramural Research, Education, and Priority Populations, Agency for Healthcare Research and Quality, Rockville, Md (Dr Mistry, Dr Desoto, and Dr Chesley); Agency for Healthcare Research and Quality, Rockville, Md (Ms Lindly); and Center for Medicaid and CHIP Services, Centers for Medicare and Medicaid Services, Baltimore, Md (Ms Llanos)

The authors declare that they have no conflict of interest.

Publication of this article was supported by the US Department of Health and Human Services and the Agency for Healthcare Research and Quality.

The views expressed in this article are those of the authors and do not necessarily represent the views of the US Department of Health and Human Services, the Agency for Healthcare Research and Quality, or the Centers for Medicare and Medicaid.

Address correspondence to Denise Dougherty, PhD, Child Health and Quality Improvement, Agency for Healthcare Research and Quality, 540 Gaither Rd, Rockville, MD 20850 (e-mail: Denise.dougherty@ahrq.hhs.gov).

Received for publication December 20, 2013; accepted June 25, 2014.

ABSTRACT

OBJECTIVE: In 2009, Centers for Medicare and Medicaid Services (CMS) publicly released an initial child core set (CCS) of health care quality measures for voluntary reporting by state Medicaid and Children's Health Insurance Program (CHIP) programs. CMS is responsible for implementing the reporting program and for updating the CCS annually. We assessed selected CCS measures for potential retirement.

METHODS: We identified a 23-member external advisory group to provide relevant expertise. We worked with the group to identify 4 major criteria with multiple subcomponents for assessing the measures. We provided information corresponding to each criterion and subcriterion, using a variety of sources such as the 2009 Medicaid Analytic eXtract (MAX), state-level Medicaid and CHIP data submitted to the CMS, and summaries of published literature on clinical and quality improvement effectiveness related to the CCS topics. Using this information, the group: 1) used a modified Delphi process to score the measures in 2 anonymous scoring rounds (on a scale of 1 to 9 in each round); 2) voted on whether each measure should be retired; and 3) provided narrative explanations of their choices (which formed the basis of our qualitative findings). Recommendations were reviewed by CMS before promulgation to state programs.

RESULTS: The Subcommittee of the National Advisory Council on Healthcare Research and Quality (SNAC) recommended that the 4 major criteria be importance, scientific acceptability, feasibility, and usability. The SNAC recommended 3 measures for retirement: access to primary care; testing for strep before recommending antibiotics for pharyngitis; and annual HbA1c testing of children with diabetes. Explanations for suggesting retirement of the measures included: views that the well-visit measures were a better measure of access than the primary care measure; a likely ceiling effect (pharyngitis); and the paucity of clinical evidence and low prevalence (both for HbA1c). CMS recommended that state Medicaid and CHIP programs retire 2 of the recommended measures from the CCS, but retained the access to primary care measure.

CONCLUSIONS: Periodic reassessment of the value of health care quality measures can reduce reporting burden and allow measure users to focus on measures with higher likelihood of leading to improvements in quality of care and child health outcomes.

KEYWORDS: AHRQ; CHIPRA; CMS; quality measures

ACADEMIC PEDIATRICS 2014;14:S97–S103

IN AN ENVIRONMENT with increasing quality measurement demands,¹ attention to the life cycle of quality measures is essential.² A life-cycle approach to quality measurement may involve the periodic examination of measures to determine whether or not they continue to be important, valid, and feasible for use.³ Unthinking continued use of quality measures—the alternative to a life-cycle approach—creates unwarranted burdens on reporting entities and can limit opportunities for more valuable measures to enter the marketplace and drive improvements in quality, equity, and child health.⁴ Here we report on an effort to reexamine selected health care quality measures from a child core set (CCS) voluntarily reported on by a number

of state Medicaid and Children's Health Insurance Program (CHIP) programs over the 3 federal fiscal years (FFY) from 2010 through 2012.

BACKGROUND

As described fully by Mangione-Smith et al,⁵ under the auspices of the 2009 Children's Health Insurance Program Reauthorization Act of 2009 (CHIPRA), an expert panel (the 2009 Subcommittee of the National Advisory Council on Healthcare Research and Quality [2009 SNAC]) established and used 3 principal criteria (importance, validity, and feasibility) and a series of Delphi processes and voting to recommend an initial CCS relevant to all child age

groups and all settings in which children received health care services. The measures were for potential future voluntary use by Medicaid and CHIP programs. By the FFY 2012 reporting period, at least 1 of the 24 measures was being reported to the Centers for Medicare and Medicaid Services (CMS) for all states and the District of Columbia; in turn, as required by CHIPRA, the secretary of the US Department of Health and Human Services released this information in a September 2013 report.⁶ In early 2013, as part of the continuing partnership between the Agency for Healthcare Research and Quality (AHRQ) and CMS, CMS asked for AHRQ's assistance in exploring whether some of the CCS measures should be retired from the CCS set so that CMS could reflect the changes in its annual update due January 2014. By early 2013, there were several reasons to consider retiring 1 or more quality measures from the CCS. These included: lessons to be learned from 3 years of state and CMS experience with the CCS⁶; the possibility of updated science to inform consideration of the measures' validity and reliability; and an emerging imperative for alignment with other public and private policies on the updating of quality measure sets.^{3,7,8}

Retirement of measures has become a theme nationally as the number of quality indicators, many of them overlapping, has grown. The National Committee for Quality Assurance (NCQA) and the National Quality Forum (NQF) both reevaluate measures every 3 years, with measure retirement or removal as a possible outcome of the reviews.^{8,9} Review of measures for potential retirement is a component of the CMS blueprint.⁷ Beyond key national entities, some health care professional societies recognize the importance of considering retirement or removal of measures as part the quality-measure life cycle.¹⁰ However, in comparison with the level of detailed criteria related to measure endorsement, guidance for the retirement or removal of measures from measure sets is still somewhat broad. For example, NCQA notes that it may retire measures when it is "clear... that a measure no longer adds value commensurate with the cost of data collection and reporting."³ NQF recommendations are based on a measure approaching being "topped out" or current development of "superior" measures.⁷

METHODS

OVERVIEW

The authors and a group of experts (called the 2013 SNAC for Subcommittee on Children's Healthcare Quality Measures of the AHRQ National Advisory Council on Healthcare Research and Quality¹¹) undertook the reexamination of 20 CMS-selected CCS measures. As shown in the Figure, the steps for considering CCS measures for potential retirement included: 1) selection by CMS of a subset of the initial CCS measures for consideration for potential retirement; 2) appointment of the 2013 SNAC; 3) a collaborative effort across AHRQ, CMS, and the 2013 SNAC to identify relevant criteria; 4) identification, consideration, and use of relevant information sources by which to assess adherence to the agreed-upon criteria; 5)

the 2013 SNAC's application of agreed-upon criteria and criteria-relevant information to the selected measures in 2 rounds of a Modified Delphi approach⁵ scoring and voting on each measure; 6) transmittal of the SNAC guidance to CMS; and 7) CMS consideration and transmittal of its recommendations in its January 2014 update. In addition, members of the SNAC provided comments explaining their scoring and voting decisions. These comments formed the basis of our qualitative findings.

MEASURE SELECTION

Of the 26 CCS measures as of January 2013, CMS selected 20 for consideration for possible retirement. As shown in Table 1, the measures selected included 4 related to perinatal care, 9 related to clinical preventive services for children and adolescents, 3 related to management of acute conditions, and 4 related to management of chronic conditions. Of the CCS measures not considered, 2 dental measures were excluded because data came from states' Early and Periodic Screening Diagnosis and Treatment reports¹²; the Child Medicaid CAHPS was excluded because it can be used to fulfil a CHIPRA requirement for state CHIP programs¹³; a measure related to antibiotic overuse in otitis media with effusion was removed from the CCS in January 2013 because data proved too challenging to collect¹⁴; and 3 measures had just been added to the CCS in January 2013.¹⁵

SELECTION OF THE 2013 SNAC

The 2013 SNAC members were selected by AHRQ and CMS from 2 overlapping pools of experts and stakeholders: subject matter experts and Medicaid/CHIP experts (primarily officials working for Medicaid or CHIP). All 2013 SNAC members¹⁶ signed a form certifying they had no conflict of interest that would affect their assessments of the 20 CCS measures, and agreed to participate actively in the entire process.

SELECTION OF CRITERIA

We based our initial suggestions to the 2013 SNAC for measure retirement criteria primarily on the 3 criteria used in 2009⁵ and the desirable measure attributes codified in the CHIPRA Pediatric Quality Measures Program Candidate Measure Submission Form.^{17,18} The desirable measure attributes were informed by criteria used by NQF but were modified in order to emphasize the CHIPRA foci on children, public insurance, and the existence of racial and ethnic and other socioeconomic and special needs disparities. The retirement criteria we proposed to the 2013 SNAC also took into account emerging guidance on recommending measures for retirement.^{3,7,8}

AHRQ initially proposed 5 criteria to the SNAC initially: importance, measure reliability and validity, measurement performance, whether a better measure was available, and feasibility. In response, the 2013 SNAC expressed a strong preference for the following: 1) reducing the number of major criteria; 2) relabeling the measure reliability and validity criterion as scientific

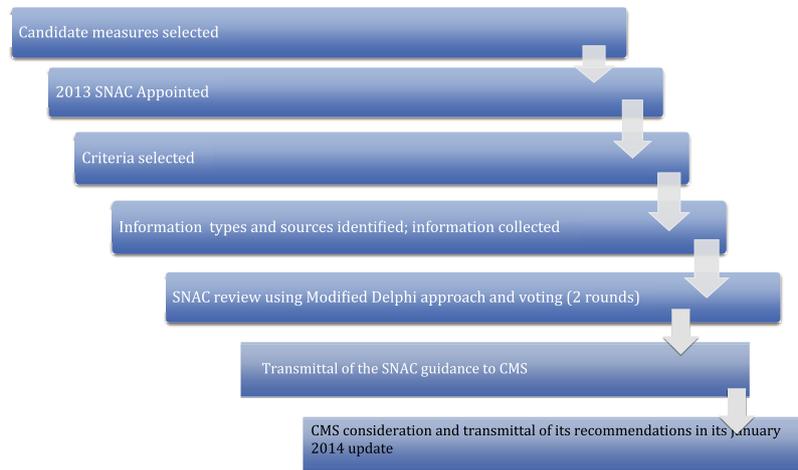


Figure. 2013 CCS measure retirement process. CMS, Centers for Medicare & Medicaid Services; SNAC Subcommittee on children’s health-care quality measures of the AHRQ National Advisory Council on Healthcare Research and Quality.

acceptability to be consistent with NQF usage; 3) incorporating state performance into the importance criterion; 4) adding the criterion of usability; and 5) removing attempts at a systematic consideration of whether better measures were available. To afford some guidance on potential alternative measures, we gave the 2013 SNAC a list of the measures previously submitted and in development under the CHIPRA Pediatric Quality Measures Program,¹⁸ but we relied primarily on SNAC knowledge and opinion to determine if a better measure was available for the CCS topics. In summary, AHRQ and the 2013 SNAC arrived at the following criteria by consensus: importance, scientific acceptability, feasibility, and usability (Box).

IDENTIFICATION OF INFORMATION TYPES AND SOURCES

Once criteria were agreed to, we identified relevant available data sources and provided the 2013 SNAC members with a draft template showing types and sources of information that we proposed to use to report on each criterion and subcriterion across the measures (as data were available). Many of the SNAC’s ideas for additional information were unable to be accommodated with the resources available. Supplemental document 1 at the AHRQ CHIPRA Web site¹⁹ provides methodological detail on the types and sources of data we made available for each of the major criteria and their subcriteria as well as the template we used to provide the information.

Table 1. Child Core Set Measures Under Consideration for Retirement

Measure	Measure Steward per CMS	Brief Measure Label
Perinatal measures	NCQA	Timeliness of prenatal care
	NCQA	Frequency of ongoing prenatal care
	CDC	Live births weighing less than 2,500 grams
		Cesarean rate for nulliparous singleton vertex
Child and adolescent clinical preventive services	NCQA	Childhood immunization status
	NCQA	Adolescent immunization status
	NCQA	Weight assessment and counseling for nutrition and physical activity for children/adolescents: body mass index assessment for children/adolescents
	OHSU	Developmental screening in the first 3 y of life
	NCQA	Chlamydia screening
	NCQA	Well-child visits in the first 15 mo of life
	NCQA	Well-child visits in years 3, 4, 5, and 6 of life
	NCQA	Adolescent well-care visit
Management of acute conditions	NCQA	Child and adolescent access to primary care practitioners
	NCQA	Appropriate testing for children with pharyngitis
	CDC	Ambulatory care–ED visits Pediatric central line–associated bloodstream infections—neonatal intensive care unit and pediatric intensive care unit
Management of chronic conditions	Alabama Medicaid	Annual percentage of asthma patients 2 to 20 y old with ≥1 asthma-related ED visits
	NCQA	Follow-up care for children prescribed ADHD medication
	NCQA	Annual pediatric hemoglobin (HbA1c) testing
	NCQA	Follow-up after hospitalization for mental illness

CMS indicates Centers for Medicare and Medicaid Services; NCQA, National Committee for Quality Assurance; ED, emergency department; CDC, US Centers for Disease Control and Prevention; and ADHD, attention-deficit/hyperactivity disorder.

Box. NQF Definitions of Major Criteria Used to Consider Retirement of CCS Measures

Criterion	Description
Importance	Extent to which the specific measure focus is evidence based, important to making significant gains in health care quality, and improving health outcomes for a specific high-priority (high-impact) aspect of health care where there is variation in or overall less than optimal performance.
Scientific acceptability	Extent to which the measure, as specified, produces consistent (reliable) and credible (valid) results about the quality of care when implemented.
Feasibility	Extent to which the specifications, including measure logic, required data that are readily available or could be captured without undue burden and can be implemented for performance measurement.
Usability	Extent to which potential audiences (eg, consumers, purchasers, providers, policy makers) are using or could use performance results for both accountability and performance improvement to achieve the goal of high-quality, efficient health care for individuals or populations.

NQF indicates National Quality Forum; CCS, Child Core Set.

Source: National Quality Forum. Measure evaluation criteria. Available at: http://www.qualityforum.org/docs/measure_evaluation_criteria.aspx. Accessed June 2, 2014.

SUMMARY CHART OF INFORMATION

In addition to providing a detailed measure report for each measure, we also prepared a summary table in response to the SNAC's request for a format that would facilitate easier scanning of key data points across the 20 measures. As shown in the summary table at the AHRQ CHIPRA Web site,²⁰ the key data points related to the importance criterion were: numbers of Medicaid/CHIP enrollees potentially affected by the measure; variation in the number of enrollees potentially affected by race/ethnicity; costs to Medicaid/CHIP of the service that was related to the measure; the evidence grade for the clinical focus of the measure; up to 5 components of performance on the measure (median performance in the latest year; 25th and 75th percentile performance in the latest year; performance trends for FFY 2010–2012; and, if available, the means and 10th and 90th percentiles in Medicaid and Commercial Healthcare Effectiveness Data and Information Set (HEDIS) performance for calendar year 2012). For the scientific acceptability criterion, key data points reported were measure reliability and measure validity. For the feasibility criterion, key data points were 3 indicators of the level of state reporting on the measures (number of states reporting in FFY 2012, state reporting trend FFY 2010–2012, and number of states reporting that data were not available to them in FFY 2012) and the data source or sources for the measure. For the usability criterion, the 1 key data point was a summary of the evidence for improvability, based on the AHRQ analysis of published literature.

As detailed in the summary table,²⁰ most of the importance information related to numbers of Medicaid/CHIP enrollees affected and costs to Medicaid/CHIP came from analyses of the 2009 Medicaid Analytic eXtract (MAX) data, sometimes supplemented by data from the AHRQ Healthcare Cost and Utilization Project (HCUP). Among the measures for which costs to Medicaid/CHIP were available, the highest costs were for discharges of all live-born infants (\$5.9 billion when fee-for-service and managed care are combined as in HCUP; relevant to the 2 prenatal care measures), followed by the costs for hospitalizations for C-sections (\$3 billion), and fee-for-service payments only for inpatient and residential treatment admissions for child and adolescent mental health care (\$2.6 billion). The greatest numbers of Medicaid/CHIP enrollees potentially

affected by a measure were affected by the access to primary care measure, which pertains to all 1- through 19-year-olds. Although white beneficiaries accounted for a plurality (about 35%) of enrollees across measures for which race and ethnicity were available, between 20 and 28% of enrollees were black and 21% to 30% of enrollees were Hispanic, depending on the measure.

Continuing with the summary table,²⁰ new evidence grades, a key subcriterion of importance, were approximated for 5 measures for which the grades were lower than a B in 2009^{5,20} (AHRQ CHIPRA's supplemental document 1¹⁹). Among the updated grades, the 2 prenatal care measures were elevated to a B– from the B and D grades in 2009^{5,20} and the attention-deficit/hyperactivity disorder (ADHD) follow-up measure was upgraded to a C from a D. However, the evidence grades for asthma emergency department (ED) visits and for HbA1c testing remained what they were in 2009 (grades of C and D, respectively).

Medicaid/CHIP state median performance rates, also a key subcriterion of importance, ranged from 28% for developmental screening in the first year of life to 85% for timely prenatal care across states using the appropriate technical specifications. Relatively low median performance rates were found for body mass index assessment (42%); chlamydia screening (49%); adolescent well visits (46%); and ADHD follow-up (43% in the initiation phase, 52% in the continuation and maintenance phase). Relatively substantial state-level variations in performance were found for adolescent immunizations (58% to 81%). However, performance showed upward trends over time for many measures, including adolescent access to primary care, and follow-up after mental hospitalization.

For the feasibility domain, information was mixed. More states reported data on the measures to the CMS CARTS (50 states and the District of Columbia) than reported data using the appropriate technical specifications (5 to 40, depending on the measure). The trend for any reporting was upward for all measures but a few showed mixed trends (eg, well-child visits in the first 15 months). Finally, evidence for the effectiveness of quality improvement (QI) interventions was most compelling for increasing up-to-date status for childhood and immunizations, and highly promising for improving chlamydia screening rates. We also found some evidence supporting the use of pay for

performance (along with additional QI interventions) for childhood immunizations. Although we found some promising research on QI effectiveness, for the most part we found few studies, methodologically poor studies, or mixed evidence.

SNAC SCORING AND VOTING

In both preliminary and final rounds of scoring, the 2013 SNAC considered the information provided in the measure reports and summary tables and used a modified Delphi process to score measures on the 4 criteria.⁵ In each round of Delphi scoring, each SNAC member was asked to provide a score between 1 and 9 for each measure for each of the major criteria, with the number 1 representing the lowest possible score and 9 representing the highest possible score. In addition to summary scores on each measure criterion, SNAC members cast a yes or no vote on whether or not to retire each measure, based on the judgment of the SNAC member after considering all the information provided on each measure. The voting component was useful during the process of identifying the initial CCS,⁵ and it functions as a way for a group to step back and consider decisions in the broader context.

ANALYSIS

Before round 2, we provided the 2013 SNAC with de-identified scoring/voting summaries from round 1 including: 1) the distribution of scores on each criterion, 2) the distribution of total scores for each measure, 3) the count of individuals scoring using the Modified Delphi approach and the number who voted to recommend retirement or not, and 4) the comments made on each criterion

and on the measure overall. The 2013 SNAC completed their final round of scoring, voting, and commenting by October 28, 2013. RTI analyzed the data and reported to AHRQ and CMS a range of statistics and graphic data for each measure, including box-and-whisker plots and a histogram that illustrates the distribution of total scores across SNAC members; and a summary of SNAC member comments for each criterion.

RESULTS

QUANTITATIVE FINDINGS

As shown in Table 2,²⁰ SNAC scoring and voting resulted in a near-universal recommendation for retirement of 3 measures: children and adolescent access to primary care practitioners; appropriate testing for children with pharyngitis; and annual pediatric hemoglobin (HbA1c) testing. (For substantial additional detail, see supplemental document 3 at the AHRQ CHIPRA Web site.¹⁹) The 3 measures recommended for retirement had the lowest median total scores out of the 20 measures in the final round (5.75, 5.25, and 4.75; Table 2). As also shown in Table 2, although the number of SNAC members who voted changed between the preliminary round of voting and the final round of scoring and voting, the median total score did not change substantially for any of the measures, and was identical for 7 measures in the 2 rounds. The remaining 17 measures not recommended for retirement ranked relatively high (7 to 9) for importance; median scores for scientific acceptability (5 to 8) were generally lower.¹⁹

Returning to the full list of 20 measures considered for retirement,¹⁹ 11 scored below a 7 on feasibility, with

Table 2. Summary of Rounds 1 and 2 SNAC Scoring Results and Retirement Recommendations

Measure	Total Score, Preliminary Round, Median (n)	Total Score, Final Round, Median (n)	No. of Retirement Votes, Final Round
5. Childhood immunization status	8.25 (15)	8.25 (22)	0
3. Live births weighing less than 2500 g	7.25 (15)	7.25 (22)	1
6. Immunization status for adolescents	8.25 (15)	8.13 (22)	1
10. Well-child visits in the first 15 mo of life	7.75 (15)	7.88 (22)	1
11. Well-child visits in years 3, 4, 5, and 6 of life	7.75 (15)	7.88 (22)	1
20. Follow-up after hospitalization for mental illness	6.75 (15)	6.75 (22)	1
1. Timeliness of prenatal care	6.75 (15)	6.88 (22)	2
12. Adolescent well-care visits	7.75 (15)	7.63 (22)	3
7. Weight assessment counseling for nutrition and physical activity for children and adolescents: body mass index assessment for children/adolescents	6.00 (15)	6.50 (22)	4
8. Developmental screening in the first 3 y of life	6.75 (15)	6.75 (22)	4
2. Frequency of ongoing prenatal care	6.25 (15)	6.25 (22)	5
4. Cesarean rate for nulliparous singleton vertex	6.75 (15)	6.50 (22)	5
9. Chlamydia screening in women	6.75 (15)	6.50 (22)	5
18. Follow-up care for children prescribed attention-deficit/hyperactivity disorder medication	6.25 (14)	6.50 (21)	5
17. Annual percentage of asthma patients with ≥ 1 asthma-related emergency room visits	6.75 (15)	6.50 (20)	6
15. Ambulatory care—emergency department visits	6.25 (15)	6.25 (22)	7
16. Pediatric central line-associated bloodstream infections	6.00 (15)	6.38 (21)	8
13. Children and adolescent access to primary care practitioners	5.75 (14)	5.75 (21)	16
19. Annual pediatric hemoglobin (HbA1c) testing	5.50 (15)	5.25 (22)	19
14. Appropriate testing for children with pharyngitis	5.00 (14)	4.75 (21)	20

SNAC indicates Subcommittee of the National Advisory Council on Healthcare Research and Quality.

developmental screening and HbA1c testing scoring the lowest (5.5 each). Child and adolescent immunizations scored the highest on usability (9 and 8.5, respectively) and pharyngitis the lowest (3.5).

QUALITATIVE FINDINGS

As shown in the comments sections for each measure in supplemental document 3 at the AHRQ CHIPRA Web site,¹⁹ explanations for suggesting retirement of the measures varied but key reasons included the following: views that the well-visit measures were a better measure of access than the primary care measure; a likely ceiling effect for the pharyngitis measure; and the paucity of clinical evidence and low prevalence for HbA1c. Across all measures, many SNAC comments pertained to the importance criterion, noting as rationales for continuation in the CCS high prevalence, substantial use of services, or considerable room for improvement as demonstrated by relatively low state-reported performance rates. SNAC members also commented on troubling variations in performance across states (eg, prenatal care, low birth weight, developmental screening, adolescent immunizations, and well-child visits). With respect to scientific acceptability, concerns were expressed about the absence of measure validation studies in general, as well as specific apprehensions such as lack of risk adjustment for the low birth weight, asthma ED, and general ED measures (although these were not recommended for retirement). Other feedback noted the potentially inappropriate level of accountability for 2 measures. Specifically, SNAC members opined that hospitals rather than states should be held accountable for occurrences of neonatal central line-associated bloodstream infections. Conversely, SNAC members noted that health care providers have little control over the timeliness of prenatal care; instead, timeliness is likely controlled by states' ability to get pregnant women enrolled in Medicaid/CHIP.

With respect to usability, the potential likelihood of improvement should states act on information from the QI evidence reviews (eg, for childhood immunizations) appeared to contribute to retention decisions. However, SNAC members expressed uncertainty about improvement potential for topics such as chlamydia screening and C-section rates for nulliparous singleton vertex deliveries. Concerns about feasibility were less frequent (eg, ability to get data on prenatal care measures due to bundled payments and claims submitted after postpartum care; a low percentage of states voluntarily reporting on some measures). Finally, SNAC members also commented on additional themes not specifically attributable to the 4 criteria. For example, a SNAC member noted that the low birth weight measure is one of the few outcome measures in the set, and another SNAC member noted that chlamydia screening was the only gynecology-related measure in the CCS.

CMS DECISIONS

After the final 2013 SNAC scoring and voting, CMS retired 3 measures as part of its updates to the CCS: pharyngitis testing, annual HbA1c testing, and the asthma ED mea-

sure.²¹ CMS recommended retiring the asthma measure as a result of a lack of an available measure steward.

DISCUSSION

We used a systematic multistakeholder-engaged process to consider carefully 20 child health care quality measures for potential retirement from the CCS of measures for voluntary state Medicaid/CHIP use. The close correspondence between levels of Delphi scores and numbers of retirement votes suggests that the process was largely evidence based. CMS accepted 2 out of 3 SNAC recommendations for its January 2014 CCS update, fulfilling a legislative requirement.

Ours appears to be the first published detailed accounting of an evidence- and stakeholder-informed systematic process for considering measure retirement. Although many of the criteria for retirement are identical to those for measure acceptance or endorsement, considerations for retirement require additional information. Examples include data on the measure collection and reporting experiences of Medicaid/CHIP programs (relevant to assessments of feasibility), performance on the measures (relevant to a component of importance), additional tests of measure validity, and use of the data as stimuli for QI interventions.

Deliberations of the 2013 SNAC were somewhat limited by the paucity or inadequacies of key data sources. For example, the MAX data date from 2009 and come primarily from fee-for-service claims, even though 64% of children enrolled in Medicaid get care through managed care plans.⁶ Therefore, prevalence, utilization, and cost data available to the SNAC were not current. CMS is working to improve its ability to more quickly collect, clean, and report Medicaid data to interested communities.²² The main advantage of MAX is its detail on Medicaid/CHIP. In addition, because MAX includes enrollment data, we could provide some information on variations in the numbers of children potentially affected by a measure by race and Hispanic ethnicity. A limitation of our study and of the field of quality measurement is the absence of measure validation studies. NCQA generously provided new information about measure reliability (for confidential 2013 SNAC purposes only) but told us that resources are generally not available for validation studies (Sarah Scholle, personal communication). Searches to inform potential changes in evidence grades were also limited; there is a paucity of rigorous studies of the effectiveness of much of routine health care, including health care for children.^{23,24} The pediatric QI literature is limited,²⁵ and we found very few studies of payment reform and health information technology in children's health care despite considerable excitement about the potential for these strategies to improve quality. Given how quickly we were required to do literature searches for and reviews of the effectiveness of clinical and QI interventions, we had to rely more heavily on expert opinion than may be ideal.²⁶ Finally, the information we were able to gather on state activities to improve performance on measure topics was deemed too general to be useful as an indicator of measure usability.

With quality measurement here to stay, it is essential to find ways to thoughtfully and comprehensively consider the breadth, depth, and evidence base for quality measures across quality domains, age groups, settings, and other foci.^{27,28} Periodic consideration of measures for potential retirement is one approach, but it may not always be able to include reflection on the full scope of potentially better measures or take the most systematic approach to gathering and reviewing evidence.

CONCLUSIONS

As part of a life-cycle approach to quality measurement, careful and systematic periodic reexamination of measures for the purpose of retirement is highly recommended for measures that are in widespread use. Revisiting measures for potential retirement is an essential component of generating and maintaining interest in children's health care quality measurement and leading toward true improvement. The expected remaining duration of a child's life compared to any adult's means that high quality care can provide substantial returns on investment.²⁹ When resources for the measurement and improvement of children's health care are available, they should be used wisely. In settings where resources are scarce and the populations vulnerable, it is critical that we measure what matters.

ACKNOWLEDGMENTS

We especially acknowledge the contributions of the 2013 SNAC and Research Triangle International. The views expressed in this article represent those of the authors only, and do not necessarily represent those of the US Department of Health and Human Services or its components, the AHRQ or the CMS.

REFERENCES

- Panzer RJ, Gotomer RS, et al. Increasing demands for quality measurement. *JAMA*. 2013;310:1971–1980.
- Lee T. Eulogy for a quality measure. *N Engl J Med*. 2007;357:1175–1177.
- National Committee for Quality Assurance. HEDIS life cycle. <https://www.ncqa.org/tabid/425/Default.aspx>. Accessed December 14, 2013.
- Institute of Medicine. Future directions for the national healthcare quality and disparities reports. April 2010. Available at: <http://www.iom.edu/Reports/2010/Future-Directions-for-the-National-Healthcare-Quality-and-Disparities-Reports.aspx>. Accessed May 26, 2014.
- Mangione-Smith R, Schiff J, Dougherty D. Identifying children's health care quality measures for Medicaid and CHIP: an evidence-informed, publicly transparent expert process. *Acad Pediatr*. 2011;11:S11–S21.
- Sebelius K. The Department of Health and Human Services 2013 annual report on the quality of care for children in Medicaid and CHIP. Available at: <http://medicaid.gov/Medicaid-CHIP-Program-Information/By-Topics/Quality-of-Care/Downloads/2013-Ann-Sec-Report.pdf>. Accessed December 14, 2013.
- Centers for Medicare and Medicaid Services. *A Blueprint for the CMS Measures Management System Health Services Advisory Group*. Baltimore, Md: Centers for Medicare and Medicaid Services; 2013.
- National Quality Forum. Maintenance of NQF-endorsed performance measures: measuring performance. 2014. Available at: <http://www.qualityforum.org>. Accessed May 24, 2014.
- National Quality Forum Measure Applications (MAP) Partnership. *MAP 2014 Recommendations on Measures for More Than 20 Federal Programs—Final Report*. Washington, DC: National Quality Forum; 2014.
- Shyu J, YL, Burleson J, et al. Performance measures in radiology. *J Am Coll Radiol*. 2014;11:456–463.
- Agency for Healthcare Research and Quality. National Advisory Council on Healthcare Research and Quality. Available at: <http://www.ahrq.gov>. Accessed December 14, 2013.
- Centers for Medicare and Medicaid Services. Early and periodic screening, diagnostic, and treatment. Available at: <http://www.medicaid.gov/>. Accessed December 14, 2013.
- US Congress. Public Law 111-3, The Children's Health Insurance Program Reauthorization Act. February 4, 2009. Available at: http://frwebgate.access.gpo.gov/cgi-bin/getdoc.cgi?dbname=111_cong_public_laws&docid=f:publ003.111. Accessed January 9, 2010.
- Agency for Healthcare Research and Quality. Improving the initial core set: CHIPRA Pediatric Quality Measures Program. Children's Health Insurance Program Reauthorization Act. Available at: <http://www.ahrq.gov/policymakers/chipra/index.html>. Accessed December 14, 2013.
- Centers for Medicare and Medicaid Services. SHO 13–002—Re: 2013 Children's core set of health care quality measures, State Health Official (SHO) letters. January 24, 2013. Available at: <http://www.medicaid.gov/Federal-Policy-Guidance/downloads/SHO-13-002.pdf>. Accessed May 24, 2014.
- Agency for Healthcare Research and Quality. Appendix A. 2013 Subcommittee of the National Advisory Council on Healthcare Research and Quality (SNAC) members. Background paper. February 2014. Available at: http://www.ahrq.gov/policymakers/chipra/measure_retirement/measure_retirementapa.html. Accessed June 23, 2014.
- Agency for Healthcare Research and Quality. Candidate measure submission form (CPCF). CHIPRA. February 2014. Accessed May 24, 2014.
- Mistry KB, Chesley F, Llanos K, Dougherty D. Advancing children's healthcare and outcomes through the Pediatric Quality Measures Program. *Acad Pediatr*. 2014;14(5 suppl):S19–S26.
- Available at: www.ahrq.gov/PQMP-supplemental-materials. Accessed August 14, 2014.
- Agency for Healthcare Research and Quality. Table 1. Initial, recommended set of children's health care quality measures. Background report on request for public comment on initial, recommended core set of Children's Healthcare Quality Measures for voluntary use by Medicaid and CHIP programs. December 2009. Available at: <http://www.ahrq.gov/policymakers/chipra/overview/background/tables.html>. Accessed June 13, 2014.
- Centers for Medicare and Medicaid Services. 2014 updates to the child and adult core health care quality measurement sets. CMCS informational bulletin. December 19, 2013. Available at: <http://medicaid.gov/Federal-Policy-Guidance/Downloads/CIB-12-19-13.pdf>. Accessed June 13, 2014.
- Centers for Medicare and Medicaid Services. SMD 13–004—Re: Transformed Medicaid Statistical Information System (T-MSIS) data. State Medicaid Director letters. August 23, 2013. Available at: <http://www.medicaid.gov/Federal-Policy-Guidance/Downloads/SMD-13-004.pdf>. Accessed June 4, 2014.
- Melnik B, Grossman D, et al. USPSTF perspective on evidence-based preventive recommendations for children. *Pediatrics*. 2012;130:e399–e407.
- Springer S, Silverstein J, Copeland K, et al. Management of type 2 diabetes mellitus in children and adolescents. *Pediatrics*. 2013;131:e648–e664.
- McCormick M, Co J, Dougherty D. Quality improvement in pediatric health care: introduction to the supplement. *Acad Pediatr*. 2013;13(6 suppl):S1–S4.
- Institute of Medicine. Finding what works in health care: standards for systematic reviews. March 2011. Available at: <http://www.iom.edu/reports/2011/finding-what-works-in-health-care-standards-for-systematic-reviews.aspx>. Accessed June 13, 2014.
- Institute of Medicine. Child and adolescent health and health care quality: measuring what matters. April 25, 2011. Available at: <http://iom.edu/Reports/2011/Child-and-Adolescent-Health-and-Health-Care-Quality.aspx>. Accessed June 13, 2014.
- National Quality Forum. Measure applications partnership. Available at: <http://www.qualityforum.org/map/>. Accessed June 13, 2014.
- Dougherty D, Chen X, Gray DT. Children and adolescent health care quality and disparities: are we making progress? *Acad Pediatr*. 2014;14:137–148.