SUPPORTING STATEMENT

Part A

Care Coordination Quality Measure for Patients in the Primary Care Setting

April 2014

Agency for Healthcare Research and Quality (AHRQ)

Table of Contents

A. JUSTIFICATION

1. Circumstances that Make the Collection of Information Necessary

The mission of the Agency for Healthcare Research and Quality (AHRQ) set out in its authorizing legislation, The Healthcare Research and Quality Act of 1999 (see http://www.ahrq.gov/hrqa99.pdf), is to enhance the quality, appropriateness, and effectiveness of health services, and access to such services, through the establishment of a broad base of scientific research and through the promotion of improvements in clinical and health systems practices, including the prevention of diseases and other health conditions. AHRQ shall promote health care quality improvement by conducting and supporting:

- research that develops and presents scientific evidence regarding all aspects of health care;
- the synthesis and dissemination of available scientific evidence for use by patients, consumers, practitioners, providers, purchasers, policy makers, and educators; and
- initiatives to advance private and public efforts to improve health care quality.

Especially pertinent to the proposed information collection, AHRQ's purview in conducting these activities includes:

- the development and assessment of methods for enhancing patient participation in their own care and for facilitating shared patient-physician decision-making;
- the outcomes, effectiveness, and cost-effectiveness of health care practices, including preventive measures and long-term care;
- the ways in which health care services are organized, delivered, and financed and the interaction and impact of these factors on the quality of patient care; and
- methods for measuring quality and strategies for improving quality.

Compared to other developed nations, the United States ranks highest in health care spending and lowest in many indicators of health care quality, prompting demands for health care reform[1]. At the heart of most proposals to enhance the efficiency of the health care system are plans to improve the coordination of patient care. Improvements to care coordination should result in cost savings (e.g., reducing redundancy in tests and procedures) and improve health outcomes (e.g., reducing medication errors caused by miscommunication across providers). These hypotheses are supported by prior studies demonstrating an association between care coordination and rates of preventive screening, emergency department (ED) utilization, and hospitalizations.[2]

Coordination of care is one of four key attributes of primary care as well as one of seven joint principles of the Patient-Centered Medical Home (PCMH)[3], both of which figure prominently in current policy to reform health care. AHRQ defines the medical home as encompassing five functions and attributes: comprehensive care, patient-centered care, coordinated care, accessible services, and an ongoing commitment to quality and safety. In addition, AHRQ recognizes the central role of health IT/electronic health record systems in successfully operationalizing and implementing the key features of the medical home, and that full medical home implementation will require significant workforce development and fundamental payment reform. The Patient

Protection and Affordable Care Act (PPACA) addresses the use of Accountable Care Organizations (ACOs) as agents of responsibility for health care quality and cost, supports PCMHs to coordinate community care, and includes health care financing reform that moves reimbursement from a volume-based system to a quality-based system.[4] Despite the importance of coordination of care, the field continues to struggle with how to define and apply metrics to care coordination. National organizations including AHRQ, the Institute of Medicine (IOM), the National Quality Forum (NQF), and the American College of Physicians (ACP), among others, have sponsored research on care coordination definitions, practices, interventions, and measurement. Evidence is building about the mechanisms by which care coordination contributes to patient-centered high value, high quality care, yet the health care community is currently struggling to determine how to measure the extent to which this vital activity is or is not occurring in primary care settings. Measuring care coordination in primary care settings for both research and accountability purposes is becoming increasingly important as new models of care such as the PCMH and ACOs, are implemented and tested, and as payment structures are increasingly supportive of integrated care delivery.

AHRQ recognized the need to support rigorous methodological work to address this gap through its Primary Care Transformation strategy within the Prevention and Care Management portfolio. AHRQ's care coordination work under this portfolio is now in its third phase. In collaboration with AHRQ, Stanford's Center for Health Policy/Primary Care and Outcomes Research (CHP/PCOR) achieved, under phases I and II:

- a conceptual framework for the measurement of care coordination[5] (see http://www.ahrq.gov/qual/careatlas/careatlas3.htm#elements),
- a searchable database of information on care coordination measures (the *Care Coordination Measures Atlas*),[6]
- technical reports of best available care coordination measures for assessing care coordination in primary care practices, and
- assessment of the potential use of emerging data sources in the construction of care coordination measures[7,8].

The framework incorporates the key care coordination *activities* that have been hypothesized or demonstrated to facilitate care coordination and broad *approaches* that have been used to improve the delivery of health care, including improved care coordination.

In phase III, supported by the American Institutes for Research (AIR) and its subcontractors, Stanford CHP/PCOR, and the National Committee for Quality Assurance (NCQA)—i.e., the "AIR team"-- AHRQ is developing and testing a new care coordination measure that builds off of the previous phases. , The survey, the Care Coordination Quality Measure for Primary Care (CCQM-PC), measures care coordination in the primary care setting from the perspective of patients and families, filling a gap identified in prior work.

Thus, the overall goal of the project is to develop the CCQM-PC. To achieve this goal, we will:

1) Draft the survey with consumer and stakeholder input
2) Conduct a pilot administration of the survey to establish psychometric properties of the CCQM-PC and understand how its scores relate to an extant practice-level measure of processes of care that support a medical home model

3) Revise the survey as necessary based on the pilot survey findings

On behalf of AHRQ, the CCQM-PC measure has been drafted by the AIR team and is provided in Appendix A. To achieve the second and third goals, AHRQ will pilot the survey with 4,500 adult patients in 30 primary care practices (PCPs) across the United States. Participating practices will also be asked complete the Medical Home Index-Long Version (MHI-LV)[9], a validated self-assessment and classification tool based on indicators of processes of care associated with the medical home model. The MHI-LV is provided in Appendix B.

There are four explicit objectives for our analysis of the pilot-test data:

- Evaluate the quality of the responses to the CCQM-PC survey (through item functioning analysis).
- Determine how the items that ask for reports of patient experiences could be summarized into a smaller set of composite measures (through factor analysis).
- Evaluate the measurement properties of the composite scales (assessment of reliability, validity, and variability of the measure).
- Identify information (i.e., case mix adjusters) that should be used to adjust scores to ensure valid comparisons among primary care practices (PCPs).
- Determine how CCQM-PC scores vary among practices that self-report processes of care that are more or less aligned with a medical home model.

The survey cover letter and reminder letter for nonresponders are included in Appendix C, along with recruitment letters for use with obtaining practices and communicating with professional organizations and networks (e.g., the American Academy of Family Physicians [AAFP], the ACP, the Society of General Internal Medicine [SGIM], and AHRQ's Practice-based Research Networks [PBRNs]) to solicit support for the survey's pilot data collection. We also provide in Appendix C information that will be reviewed by survey participants as part of consent procedures.

This study is being conducted by AHRQ through its contractor, AIR, pursuant to AHRQ's statutory authority to conduct and support research on healthcare and on systems for the delivery of such care, including activities with respect to the quality, effectiveness, efficiency, appropriateness and value of healthcare services and with respect to quality measurement and improvement. 42 U.S.C. 299a(a)(1) and (2).

2. Purpose and Use of Information

The information collected in the pilot survey will be used to test and improve the draft survey. The pilot design will support the standard suite of psychometric analyses conducted to identify and develop composite scoring algorithms as well as to provide evidence of the reliability and construct validity of the composite scores and any scores based on individual items.

Additionally, the variations in composite scores and total CCQM-PC scores will be examined for any differences that may be correlated with variations in the practice's self-assessment of its

engagement in processes of care that are consistent with the medical home model. The analyses will include the following components:

- Item functioning analysis
- Confirmatory Factor Analysis
- Exploratory Factor Analysis
- Evaluation of the reliability, validity, and variability of composite and single-item Scores
- Case mix adjustment (if the data indicate this is needed).

Because the survey items are being developed to measure specific aspects of care coordination in accordance with the domain framework developed through previous phases of AHRQ's Care Coordination Measure Development portfolio, the factor structure of the survey items will be evaluated through multilevel confirmatory factor analysis. On the basis of the data analyses, items or factors may be dropped. Exploratory factor analysis is also planned.

Data from the pilot survey will be used to make final adjustments to the CCQM-PC. The final survey instrument will be made publicly available, at no charge to prospective users, for use in research projects that aim to assess care coordination as it relates to quality care and healthcare outcomes, thereby helping to expand the evidence base for the care coordination construct and its associated processes. There is value, given where the field is now, in developing a survey of reasonable length that can be used for research purposes, but also can serve as the "parent" survey from which a smaller subset of items appropriate for quality improvement could be drawn.

A well-developed, psychometrically-sound, practical survey of adult patients' experiences of care coordination in primary care settings, that covers key conceptual domains articulated through AHRQ's past work in this area, will help generate evidence that is needed to understand the relationship between care coordination processes and health outcomes, in addition to offering a way to explore other critical questions regarding care coordination.

The development of this research-focused survey is a critical step in moving toward the future development by the field of measures of care coordination in primary care settings that can be used for accountability purposes, including those submitted for consideration of endorsement by the NQF. This will ensure that the measure is useful from a public reporting perspective to a variety of potential stakeholders, including patients seeking providers that engage in care coordination practices supported by the evidence base. The key target audiences for the use of the survey are researchers and, ultimately, payers (including health insurance plans, employers, and entities such as the Centers for Medicare and Medicaid Services), although use by health systems and individual primary care practices is also envisioned.

3. Use of Improved Information Technology

This proposed collection of information for the CCQM-PC pilot survey will be used to revise the survey and examine its psychometric properties. A single survey vendor will administer the survey using two modes of administration—mail and telephone interview. The primary administration mode will be a paper- and-pencil survey sent through the U.S. Mail and will involve no information technology except for the scanning of surveys for creating the electronic data file. Telephone follow up for nonresponders to the mail survey will be conducted using computer-assisted telephone interviewing (CATI).

It is anticipated that participating PCPs will have electronic health record (EHR) systems in place that facilitate the drawing of an appropriate sampling frame at low burden to the participating practices.

4. Efforts to Identify Duplication

The survey development process for the CCQM-PC closely follows principles of survey design established by the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Consortium. CAHPS® principles help ensure that the data are psychometrically sound (reliable and valid, address the key domains while minimizing respondent burden), credible (feasible), and useful (for research, quality reporting, and accountability). Thus, extant CAHPS® surveys were examined for potential inclusion or modification of items well-suited to the conceptual framework. The most relevant of these had been mapped to the project's care coordination conceptual framework through the catalogue of care coordination measures included in the *Care Coordination Measures Atlas* (discussed above) and update to the Atlas performed under the current phase of the AHRQ portfolio. To begin measure development, AHRQ's contractor, AIR, created a database (i.e., an "item library") of all items included in extant measures of care coordination that were both identified by the Atlas and assessed care coordination from the patient/family perspective. Included items also were specific to those measures that were identified in the Atlas and other work in the care coordination portfolio as being applicable to primary care. To create the item library, measure profiles created for the Atlas were used to identify items that had been "mapped" in the past to one or more of the care coordination activity domains in the Atlas conceptual framework. AHRQ contends that no single measure found to date in the course of developing the Atlas provides the comprehensive view of care coordination that is necessitated by the extant conceptual framework.

AHRQ's contractors for all phases of this portfolio have recruited and consulted with stakeholder panels to identify any additional measures that may have been missed by the environmental scan and literature reviews conducted in conjunction with the Atlas work. No measures were identified that weren't included in the Atlas. Thus, although there are many measures of care coordination, none has the breadth to assess the quality of care coordination that patients experience in primary care, when care coordination is holistically approached as a construct encompassing the domains defined in the Atlas conceptual framework.

5. Involvement of Small Entities

Some of the PCPs participating in this pilot test will be small practices, possibly with a solo provider and limited office staff capacity. The sampling and data collection procedures are designed to minimize burden on individual PCPs (e.g., through the use of EHR-based data to identify the sampling frame, through the provision of technical assistance and support in the identification and production of the sampling frame [this will be provided by the AIR contract], and through the use of a survey vendor to administer the survey vs. involving the PCPs in data collection) and should not have a significant impact on them. Our recruitment plan targets a range of practice sizes and configurations. Small practices will likely only be found among one segment of recruited practices: physician-owned, single site PCPs. There will only be nine PCPs recruited of this practice type, only some of which will be very small (i.e., two or fewer providers). Thus, fewer than nine practices are expected to be small entities.

6. Consequences if Information Collected Less Frequently

This effort is a one-time pilot test.

7. Special Circumstances

This request is consistent with the general information collection guidelines of 5 CFR 1320.5(d)(2). No special circumstances apply.

8. Federal Register Notice and Outside Consultations

8.a. Federal Register Notice

As required by 5 CFR 1320.8(d), a notice was published in the Federal Register on July 30th, 2014, for 60 days, and again on October 16th, 2014 for 30 days (see Attachment D. AHRQ received and responded to two comments from the comments (see Attachment F).

8.b. Outside Consultations

AHRQ is working with AIR and its subcontractors, Stanford's CHP/PCOR, and NCQA. In addition, a panel of stakeholders composed of consumer advocates, primary care providers, researchers/methodologists, survey experts, community-based healthcare professionals, and health care organization leaders is involved with this phase of the work. The panel provided input to our formative research activities, advised on the domain framework, provided feedback on the draft survey, and will help envision successful uptake of the CCQM-PC by users. See Appendix E for a list of those consulted both within and outside the Agency thus far. There are no unresolved issues.

9. Payments/Gifts to Respondents

Pilot Survey Respondents. We are offering no incentive to participants taking the pilot survey. Our recruitment and consent materials frame the perspective that taking the survey contributes to the knowledge about how providers can better serve their patients. We will provide PCPs with a report on their own performance on the CCQM-PC including benchmarks based on other participating practices. The report will provide a tool by which participating practices can understand where they excel and where they could improve with respect to the various composite scores. AHRQ's contractor, AIR, has had considerable success using such a report as a non-financial incentive for participating entities in CAHPS and CAHPS-like surveys.

10. Assurance of Confidentiality

Individuals and organizations will be assured of the confidentiality of their replies under Section 934(c) of the Public Health Service Act, 42 USC 299c-3(c). They will be told the purposes for which the information is collected and that, in accordance with this statute, any identifiable information about them will not be used or disclosed for any other purpose.

AIR will apply for a Health Insurance Portability and Accountability Act (HIPAA) waiver of authorization from the AIR Institutional Review Board (IRB) to enable the participating practices to share contact information as part of the recruitment and survey operations processes without obtaining prior permission from patients. Practices will be concerned about sharing the identities of their patients without informed consent and HIPAA authorization. In all of AIR's many CAHPS survey development projects, AIR's IRB has granted a waiver of documentation of consent and a partial waiver of HIPAA authorization, which permit practices to release patient contact information for the purpose of making initial contact. The respondents are then asked to provide consent in the cover letter that accompanies the survey in self-administered modes or in the initial telephone introduction. We will include a copy of AIR's IRB approval and HIPAA waiver forms with the invitation letter to practices, along with an explanation of the HIPAA regulation for research uses of protected health information to demonstrate that releasing this information for this project is ethical and legal.

If a waiver of documentation of consent is granted, AIR's survey vendor will be the only party with access to identifiable information (patient name and contact information) provided by practices for the purposes of creating the sampling frame. The analytic team at AIR will only receive a de-identified data set of survey responses.

11. Questions of a Sensitive Nature

We do not believe there are questions of a sensitive nature included in the survey. If sensitivities are discovered in cognitive testing, they will be modified so they are not sensitive nature.

12. Estimates of Annualized Burden Hours and Costs

Exhibit 1 shows the total estimated annualized burden hours for the CCQM-PC pilot survey (2,022 hours), including burden for survey respondents (1,890 hours) and practice staff (132 hours). With respect to the burden on CCQM-PC survey respondents, thirty practices will be sampled, with the survey sent to 375 prospective respondents per sample. A 40% response rate (in keeping with response rates on other CAHPS® and CAHPS®-like surveys of similar length and mode) will yield 150 respondents per practice. Total respondents were calculated by multiplying the number of practices by the respondents per practice, for a total of 4,500 (i.e., 150x30=4,500). The survey has 102 items (79 assessment items, 4 items about healthcare services sought in the past 12 months, and 19 items that assess participant characteristics such as demographics), with an estimated completion time of 25 minutes (.42 hours) per survey response. This estimate is based on the length of previous CAHPS® surveys of comparable length that have been administered to similar populations.

Burden hours for participating practices are calculated based on the total burden to one physician/ administrator and one other clinician to complete the MHI-LV. The measure author recommends that both physician and non-physician viewpoints are considered in the PCP's response, thus the estimate is based on an assumption that two clinicians per practice will complete the MHI-LV process of care items together, with only one of the clinicians (i.e., the physician/administrator) completing the items on practice characteristics. Contract staff from AIR will ensure that practices realize there is no burden to them on the MHI-LV other than the time required to fill out the MHI-LV tool (i.e., they can ignore the measure author's reference in the instructions to a companion patient tool associated with the MHI-LV).

Exhibit 1. Estimated Annualized Burden Hours for CCQM-PC Survey Pilot Test by Entity

Data collection activity	Number of Respondents	Number of responses per respondent	Hours per response	Total Burden hours
CCQM-PC survey	**4,500**	**1**	**0.42**	**1,890**
MHI-LV[1]: Physician/administrator	**30**	**1**	**2.33**	**70**
MHI-LV: Non-physician clinician	**30**	**1**	**2.08**	**62**
Total				**2022**

[1]The instructions for completing the MHI-LV recommend that a physician/administrator and a non-physician clinician each fill out the index separately. So, even though it is one form as reproduced in Appendix B, we have two rows in the table to describe the burden of the two individuals. There are a series of questions on the first two pages of the index which simply require administrative information and would only need to be completed once. We assume that the administrator would complete these and so the time required for the administrator to complete the MHI-LV is longer than that required for the clinician.

Exhibit 2 shows the estimated annualized cost burden associated with the pilot survey administration. The total cost burden is estimated to be $51,228 for the one-time survey pilot.

Exhibit 2. Estimated Annualized Cost Burden for CCQM-PC Survey Pilot Test by Entity

CCQM-PC Survey	Total Burden Hours	Average Hourly Wage	Total Cost Burden
Survey Respondents	**1,890**	**$22.33**[1]	**$42,204**
Physician/Administrator	**70**	**$88.43**[2]	**$6,190**
Non-physician Clinician	**62**	**$45.71**[3]	**$2,834**
TOTAL OVERALL	**2022**	**n/a**	**$51,228**

[1] Average wage for civilian workers, http://www.bls.gov/news.release/ocwage.htm.
[2] Average wage for family and general practitioners, http://www.bls.gov/news.release/ocwage.htm
[3] Average wage for nurse practitioners, http://www.bls.gov/news.release/ocwage.htm

13. Estimates of Annualized Respondent Capital and Maintenance Costs

There are no direct costs to respondents other than their time to participate in the pilot study.

14. Estimates of Annualized Cost to the Government

The cost to the Government of this information collection that would not otherwise have been incurred is the cost of the AIR contract ($1,689,948) as well as the labor of Federal employee(s) who provide oversight to that contract ($16,824). These costs comprise the annual (and total) cost to the Federal government of $1,706,772. Exhibit 3 shows the estimated total and annualized cost for this project. Although data collection will last for less than one year, the entire project is a three-year effort.

Costs, including pre-OMB approval costs, have been calculated to manage and support tasks for the survey, from survey development and administration through the development of reports and presentation of findings to AHRQ staff.

Federal Government staff time is required to conduct research and development, PRA/OMB Clearance development, and to lead and support this study. The Health Scientist Administrator (Task Order Officer) is responsible for project management, planning, and oversight for the study.

Exhibit 3. Estimated Total and Annualized Cost

Cost Component	Pre-OMB	Post OMB	Total Cost	Annualized Cost[2]
Contractor Costs				
Survey Development	$585,373	$0	$585,373	$195,124.33
Data Collection Activities (Pilot)		$261,083	$261,083	$87,027.67
Data Processing and Analysis		$118,933	$118,933	$39,644.33
Reports and Publication of Results		$336,710	$336,710	$112,236.67
Project Management	$110,550	$56,950	$167,500	$55,833.33
Overhead	$104,345	$116,003.57	$220,349	$73,449.67
SUBTOTAL CONTRACTOR COSTS	*$800,268*	*$889,679.57*	*$1,689,948*	*$563,316.00*
Government Personnel and Non-Personnel Costs				
Health Scientist Administrator—GS 14, Step 9[1]	$6,730	$10,095	$16,824	$5,608
Non-personnel Costs	$0	$0	$0	$0.00
SUBTOTAL GOVERNMENT PERSONNEL AND NON-PERSONNEL COSTS	*$6,730*	*$10,095*	*$16,824*	*$5,608*
Total Costs	$806,998	$899,774	$1,706,772	*$568,924*

[1] Based on 2014 OPM Pay Schedule for the Washington-Baltimore-Northern VA locality pay area: http://www.opm.gov/policy-data-oversight/pay-leave/salaries-wages/salary-tables/pdf/2014/DCB.pdf, annual salary of $134,595; 5% time pre-OMB and 7.5% post-OMB

[2] Annualized cost divides each line item by the three (i.e., the three years of the project).

15. Changes in Hour Burden

This is a new information collection activity.

16. Time Schedule, Publication and Analysis Plans

Schedule and Publication Plans. AHRQ targets fall to winter 2014 as the timeframe for administration of the pilot survey. Data collection will be open for a three month period, including initial sendout of the survey by mail, reminder cards, follow up by phone for non-responders, and CATI-supported phone interviews where needed to supplement the mail responses. As soon as OMB approval is received, pilot survey activities will begin. Cognitive testing of the survey will already have occurred (and was designed in a way that would not require its clearance through OMB). The estimated time schedule to conduct the data collection activities is shown below:

1. Finalize recruitment (2 months)
2. Pilot study data collection (3 months)
3. Data analysis, development of draft technical report (2 months)
4. Final technical report, plan for item revisions and supplemental analyses (2 months)
5. Supplemental analyses; final survey item set and supporting documentation (3 months)

The final version of the CCQM-PC and accompanying documentation will be made available in the public domain on the AHRQ Web site.

Analysis Plans. The primary purpose of the pilot survey is to test the CCQM-PC, assess any issues with item functioning, and establish its psychometrics. Psychometric analyses of the pilot data will involve an examination of item-level non-response and variability. After the pilot survey has been conducted, we will conduct psychometric analyses based on the item responses to identify the most important and appropriate items for each care coordination domain identified in the conceptual framework and formative research with consumers. The analyses will be conducted using the classical factor analysis approach, which is the CAHPS standard. The analyses will include the following components—

- Item functioning analysis
- Confirmatory Factor Analysis
- Exploratory Factor Analysis
- Test of the Reliability of the Survey
- Scoring of the Composites
- Evaluation of the reliability, validity, and variability of the composite and any single-item Scores
- Case mix adjustment (if the data indicate this is needed).

Each of these components is described in subsequent sections.

A secondary goal of the information collection is to collect information from practices about their philosophies and engagement in processes of care that are more or less aligned with the principles of the medical home model. The planned analysis for examining the MHI-LV in relation to CCQM-PC will involve, at minimum, comparison of patients' CCQM-PC scores for practices that self-assess as low vs. high on MHI-LV overall and/or domain scores. We do not know what the distribution of MHI-LV scores will be among the selected practices, thus we may restrict the analysis to practices in the top or bottom third relative to other included practices, particularly if there appear to be clusters of practices at the top or bottom of the scoring range. Of primary interest in this analysis is whether patient-reported experiences of care coordination vary in ways that would be expected among practices that have more or fewer processes of care in place that are conducive to supporting care coordination and high quality care.

Item Functioning Analysis. Item-functioning analysis involves calculating the rate of occurrence of various survey item dispositions, including assessments of respondents' ability to follow any skip patterns in the survey (i.e., what proportion of respondents skip items they are supposed to skip, versus those who fail to skip items they are supposed to skip), and the rate of item nonresponse (when patients fail to respond to items they are eligible to answer).

The goal of the item functioning analysis is to identify any survey questions that may have been confusing or burdensome to respondents by flagging items with high rates of nonresponse and identifying areas of the survey where skip instructions were not correctly followed. When correlated with participant characteristics or experience, high rates of nonresponse can potentially bias the survey results. Inordinate rates of nonresponse to a question suggest that the

question was poorly understood by respondents (who skipped the item because they were not sure how to reply), that the item did not apply to the respondent, or that the item asked for sensitive information that respondents may have been unwilling to give. Failure to follow skip instructions can suggest that those instructions were unclear or inappropriate (and thus the respondent ignored or chose not to follow the instructions), or that the format of the survey made it difficult for respondents to understand the skip instructions (and thus they were unable to follow the skip instructions).

The item functioning analysis will be conducted at both the item level and the respondent level. The former is used to identify problem items, while the latter is designed to identify respondents who might have had problems completing the survey.

Confirmatory Factor Analysis. A confirmatory factor analysis (CFA) involves explicitly testing the full set of linkages between survey items and the latent constructs to which they are thought to belong. The latent constructs (i.e., domains or factors), the items, the loadings of the items on the factors, and any correlations among items and among factors are collectively referred to as the hypothesized factor structure. The domain-item map will inform this hypothesized factor structure. We will test the fidelity of observed question responses to this conceptual model by conducting CFA using structural equation modeling (SEM)[10, 11].

With large samples, even trivial departures from the specified model may be statistically rejected; therefore, it is customary to use practical fit indices to evaluate the hypothesized model. Specifically, we rely on the comparative fit index (CFI) and non-normed fit index (NNFI) along with the standardized root mean square residual (RMSEA) [12,13,14]. The CFI and NNFI compare the fit of the specified model to that of a model which specifies no covariation (the null model). Both indices run from a value of "0" (no relationship between the predicted and observed correlation matrix) to "1.0" (the predicted correlation matrix is identical to the observed). The NNFI, also known as the Taylor-Lewis Index (TLI), includes a greater correction for the number of parameters in the model (analogous to an adjusted R2) than the CFI. The RMSEA is the amount of variance that is not predicted by the model and has associated confidence intervals (which the CFI and NNFI do not). A CFI and NNFI less than 0.90 and an RMSEA greater than 0.10 indicate that the hypothesized model may not be the best description of the data. Excellent fit of the model to the data is considered if the CFI and NNFI are equal to or greater than 0.95 and the RMSEA is equal to or less than 0.06.

CAHPS surveys typically contain at least some structured missing data, which results from patients skipping questions that do not apply to their experiences. Alternatively, they may respond to a tailored "not applicable" option written into specific items. We will conduct the CFA using Mplus, which implements a pair-wise deletion of missing values and uses the appropriate poly- and tetrachoric correlations as input. This eliminates the need for imputation and uses a more appropriate input for the modeling than other programs do.

Exploratory Factor Analysis. It is often the case in CAHPS surveys that the observed data do not fit the hypothesized factor structure. If this is the case with the CCQM-PC, we will conduct an exploratory factor analysis (EFA) to identify the underlying factor structure in the data. The EFA will use the principal factor method with squared multiple correlations as initial-communality estimates and oblique rotation (promax) with Kaiser normalization. The number of

factors will be determined by parallel analysis[15] as well as the interpretability of the rotated-factor-pattern matrix [16]). Items will be assigned to a factor that have standardized-regression coefficients greater than 0.40 following Stevens[17] and no coefficients on secondary factors that differ from the primary factor by less than 0.20.

Alternative factor structures will be hypothesized using information from the EFA, as well as input from the analysis team and AHRQ regarding the face validity of the factors in the alternative structures. We will test all alternative structures using CFA. If multiple structures display a good fit to the data, we will rely on reliability and validity analyses to help determine the final factor structure (see below).

Test of the Reliability of the Survey. One of the primary purposes of a CAHPS survey is to be able to detect variability across reporting units. Even though the CCQM-PC will not pursue the CAHPS® trademark, it is still relevant to understand how care coordination differs among the units—in this case, primary care practices—because a range of practice types and ownership models will be included in the pilot. The practices vary on many characteristics that past research has shown to be associated with capacity for quality improvement and practice redesign (e.g, practice size, ownership status). Thus, one of the major goals for the composites and other survey items from this measure is to be able to discriminate across PCPs. This ability is referred to as unit-level reliability, and there are two statistics used to assess this reliability. One is a measure of inter-unit reliability (IUR) based on the F-statistic from an analysis of variance (ANOVA). The IUR is equal to F-1/F; it can also be calculated as the between-unit variance minus the within-unit variance over the between-unit variance.[18, 19] The other measure is the intra-class correlation (ICC), which is also calculated using statistics produced by an ANOVA. The ICC is the between-unit variance minus the within-unit variance over the total variance adjusted for the average number of respondents per reporting unit.[20]

The IUR provides the reliability based on the sample size associated with the data, while the ICC indicates the reliability of a measure for a single respondent. Scales with reliability coefficients above 0.70 provide adequate precision for use in statistical analysis of group-level comparisons.[21]

In addition, using the ICC and the Spearman-Brown prophecy formula,[22,23] and adjusting for the item-level response rates due to skip instructions, we can estimate the number of complete responses, as well as the total effective sample size, needed for a composite or item to yield an acceptable IUR.

Determination of Composite- and Single-item Scores. Items assigned to factors based on the results of factor analysis will comprise multi-item composite measures. Those items that do not relate to any of the composites may be retained in the survey as single-item measures depending on their importance to AHRQ, its contractors, the stakeholder panel, and prior qualitative analysis and public comments. In calculating composite scores, we will use the CAHPS Analysis Program.[1]

[1] Comprehensive instructions for using the CAHPS analysis program are provided in ***Instructions for Analyzing CAHPS Data*** (Document No. 15), which is included in the ***CAHPS Survey and Reporting Kit.*** CAHPS Reporting Kit materials are available at: https://www.cahps.ahrq.gov/cahpskit/CAHPSKIT_main.asp

The CAHPS Analysis Program provides a standardized way to—

- Construct composite scores out of individual items
- Adjust for case-mix
- Estimate unit scores
- Test the significance of unit rankings
- Apply sampling weights
- Report unit-level quality results—in the case of the CCQM-PC, this would be the quality of care coordination perceived by patients within practices.

The CAHPS Analysis Program generates output that can be used to calculate the performance of each practice and, where appropriate, show how a practice's performance compares to the overall performance of other practices.

Evaluation of Reliability, Variability, and Validity of Composite Scores. Internal-consistency reliability is a traditional method used to evaluate the amount of systematic variance among the items in a composite and is widely used and understood. We will provide estimates of this reliability using Cronbach's Alpha.

Lack of variability in scores can attenuate validity coefficients and reduce the amount of information provided by the survey, so we evaluate the distributional properties of the scores. For variability, we are most interested in knowing the percentage of respondents with the highest (ceiling effect) and lowest (floor effect) possible scores. Ceiling and floor effects indicate the percentage of people for whom it would be impossible to assess improvement or decrement, respectively, over time. Composites or single items with high ceiling and/or floor effects should be considered for modification or deletion.

Construct validity will be assessed using the multi-trait, multi-method approach. This approach entails comparing the correlations of items with their composite total (correcting for overlap[24]) to the correlations of items with competing composites. The 'scaling success' statistic (an indicator of discriminant validity) is one of a number of pieces of evidence that bears on the construct validity of the proposed composites—scaling success of 100 percent indicates that all items correlate more highly (at least one standard error higher) with their own composites than with competing composites. The logic for this analysis as an evaluation of construct validity follows that laid out by Campbell and Fiske (1959).[25] In addition, to evaluate the relative importance of the items and composites in predicting the overall evaluations of the practices (i.e., criterion validity), we will regress the overall rating of care coordination on all of the composites. The better a composite can predict the overall rating of care coordination, the higher the validity of that composite would be. In addition, we can look at correlations between care coordination composites from the CCQM-PC and the self-assessment by practices on the MHI-LV.

Identification of Important Case-Mix Adjusters. One opportunity offered by developing the CCQM-PC in the CAHPS® tradition is that primary care practices can be compared to a benchmark, typically the mean of all reporting units in a particular universe. In order to make comparisons, it is important to control for the influence of patient characteristics on the outcome variables (composite scores and overall ratings).

Past research has shown that some types of patients, such as older patients or patients in better health, tend to give higher ratings of their hospital care than patients who are younger or in poor health [26]. Conversely, those patients with more education tend to give lower ratings of their health care experiences. These are characteristics of the patients that are related to the CAHPS scores but are not within the control of the provider, nor are they believed to reflect true differences in the quality of care delivered. With respect to the CCQM-PC, when comparing PCPs to each other or to some benchmark, the differences should derive entirely from differences in the quality of care coordination provided. However, if the differences derive in part from differences in their patient populations, it is important to remove (i.e., adjust for) the portion of the scores that come from patient characteristics so PCPs are not held accountable for factors that are beyond their control. Thus, the three goals of case-mix adjustment are to[27]—

1. Help remove the effects of individual patient characteristics that can affect ratings
2. Remove effects that might be considered spurious (i.e., that reflect something other than quality of care)
3. Remove incentives for practices to avoid aspects of care coordination that are more complicated with particular subgroups of patients.

Zaslavsky [21] outlines three conditions to be met in the selection of variables for case-mix adjustment:

1. Within reporting units, the case-mix variables must be related to the outcome measures (ratings). That is, the variables must have sufficient predictive power in relation to the outcomes (e.g., older patients give higher ratings of their care). These variables are referred to as "predictors" of the outcome being examined.
2. There must be variation between reporting units on these predictor variables. That is, the predictors must be unevenly distributed across reporting units (e.g., one treatment center might have a patient population that tends to be much younger than the patient population of a treatment center in another location). This condition is the heterogeneity factor of the predictor.
3. The case-mix variables must be appropriate for adjustment because they are not themselves determined by the provider's actions. That is, they must be characteristics that are brought to the provider by the patient (e.g., age or education), not characteristics that might be consequences of the patient's satisfaction with, or assessment of, the provider. Predictors that are consequences of the patient's satisfaction with the provider are endogenous. For example, patients who are happy with their treatment at a particular provider tend to stay with their provider longer, and thus the length of the relationship with the provider would be endogenous: the quality of care in this situation predicts the length of the relationship, rather than the reverse. This appropriateness criterion cannot be assessed statistically; rather, it must be determined based on researchers' knowledge of a particular care setting and patient population.

The case-mix analysis follows four steps—

- Selection of potential case-mix adjusters
- Estimation of predictive power of the selected adjusters
- Estimation of heterogeneity (the degree to which the adjusters vary across providers)

- Calculation of explanatory power and impact of each adjuster

Heterogeneity—differing distributions of candidate case-mix adjusters among the PCPs being compared—and impact, which is a function of heterogeneity, are necessary conditions for choosing a case-mix adjuster. Those variables with a significant impact will be recommended as case mix adjusters. The project team will also obtain the input of the stakeholder panel in determining candidate case-mix adjusters.

Revise Survey Questionnaire Based on Field Test Findings. Based on the psychometric evaluation of the pilot survey, we will revise the CCQM-PC as necessary. This evaluation typically indicates that we can eliminate questions that do not contribute substantially to a composite measure or serve as an effective case-mix adjuster.

17. Exemption for Display of Expiration Date

No exemption is being requested.

List of Attachments

Appendix A – Care Coordination Quality Measure for Primary Care (CCQM-PC)

Appendix B – The Medical Home Index: Adult, Long Version (MHI-LV)

Appendix C – Recruitment and Consent Materials

Appendix D – Federal Register Notice

Appendix E – List of CCQM-PC Consultants and Stakeholder Panel Members

Appendix F – Public Comments and AHRQ Response

[1] Davis, K, Schoen, C, Stremikis, K. Mirror, Mirror on the Wall: How the Performance of the U.S. Health Care System Compares Internationally 2010 Update, The Commonwealth Fund, June 2010.

[2] Ferrante JM, Balasubramanian BA, Hudson SV, Crabtree BF. Principles of the Patient-Centered Medical Home and Preventive Services Delivery. Annals of Family Medicine, 8(2):108-116, 2010.

[3] Stange KC, Nutting PA, Miller WL, Jaén CR, Crabtree BF, Flocke SA, et al. Defining and measuring the patient-centered medical home. Journal of General Internal Medicine, 25(6):601-612, 2010.

[4] U.S. PUBLIC LAW 111 - 148 - Patient Protection And Affordable Care Act - H.R. 3590

[5] McDonald KM, Sundaram V, Bravata DM, Lewis R, Lin N, Kraft S et al. Care coordination. Vol 7 of: Shojania KG, McDonald KM, Wachter RM, Owens DK, editors. Closing The Quality Gap: A Critical Analysis Of Quality Improvement Strategies. Technical Review 9 (Prepared by Stanford-UCSF Evidence-Based Practice Center under contract No. 290-02-0017). AHRQ Publication No. 04(07)-0051-7. Rockville, MD: Agency for Healthcare Research and Quality. June 2007.

[6] McDonald KM, Schultz E, Albin L, Pineda N, Lonhart J, Sundaram V et al. *Care Coordination Measures Atlas.* AHRQ Publication No. 11-0023-EF, January 2011. Agency for Healthcare Research and Quality, Rockville, MD. http://www.ahrq.gov/qual/careatlas/.

[7] McDonald KM, Schultz E, Pineda N, Lonhart J, Chapman, T. and Davies, S. Care Coordination Accountability Measures for Primary Care Practice (Prepared by Stanford University under subcontract to Battelle on Contract No. 290-04-0020). AHRQ Publication No.12-0019-EF. Rockville, MD: Agency for Healthcare Research and Quality. January 2012.

[8] McDonald KM, Schultz E, Chapman T, Davies S, Pineda N, Lonhart J et al. Prospects for Care Coordination Measurement Using Electronic Data Sources. (Prepared under contract number 290-04-0020 AHRQ SQI-II.) AHRQ Publication No. 12-0014-EF. Agency for Healthcare Research and Quality, Rockville, MD. March 2012 https://www.pcpcc.net/content/joint-principles-patient-centered-medical-home.

[9] Center for Medical Home Improvement (CMHI) Web site. Available at: http://www.medicalhomeimprovement.org/knowledge/practices.html#measurement Accessed: April 14, 2014.

[10] Hoyle RH (ed.). Structural equation modeling: concepts, issues, and applications. Thousand Oaks, CA: SAGE Publications, Inc., 1995.

[11] Keller SD, O'Malley AJ, Hays RD, et al. Methods used to streamline the CAHPS® Hospital Survey. Health Services Research, 40(6):2057-7, 2005.

[12] Hu L, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. Structural Equations Modeling, 6(1):1-55, 1999.

[13] Kenny, DA. Measuring Model Fit. http://davidakenny.net/cm/fit.htm. Revised July 5, 2012. Accessed September 11, 2013.

[14] Suhr, DD (2006). Exploratory or confirmatory factor analysis? Paper 200-31. In: SAS Institute Inc. Proceedings of the thirty-first annual SAS® Users Group International Conference. 2006 March 26–29; San Francisco, CA. Cary (NC): SAS Institute Inc.; 2006:1-17.

[15] Horn JL. A rationale and a test for the number of factors in factor analysis. Psychometrika, 30: 179-85, 1965.

[16] Thurstone LL. *Multiple-factor analysis*. Chicago: University of Chicago Press, 1947.

[17] Stevens J. Applied multivariate statistics for the social sciences. 2nd ed. Hillsdale, NJ: Erlbaum, 1992.

[18] Winer BJ. Statistical principles in experimental design. New York: McGraw-Hill, 1970.

[19] Zaslavsky AM, Buntin MJB. Using survey measures to assess risk selection among Medicare Managed care plans. Inquiry, 39(2):138-51, 1965.

[20] Hays RD, Revicki D. Reliability and validity (including responsiveness). In: Fayers P, Hays R, eds. Assessing quality of life in clinical trials: Methods and practices. 2nd ed. Oxford: Oxford University Press; 2005:41-53.

[21] Nunnally JC. Psychometric theory. 2nd edition. New York: McGraw-Hill Book Company; 1978.

[22] Spearman CC. Correlation calculated from faulty data. Brit J Psychol 1910;3:271-95.

[23] Brown W. Some experimental results in the correlation of mental abilities. British Journal of Psychology, 3:296-322, 1910.

[24] Howard KI, Forehand GG. A method for correcting item-total correlations for the effect of relevant item inclusion, Educational and Psychological Measurement, 22(4):731-5, 1962.

[25] Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psycholical Bulletin, 56:81-105, 1962.

[26] O'Malley AJ, Zaslavsky AM, Elliott MN, Zaborski L, Cleary PD. Case-mix adjustment of the CAHPS Hospital Survey. Health Services Research, 40(6, pt 2):2162–81, 2005.

[27] Zaslavsky AM. Proceedings of the Joint Statistical Meetings. (Section on Government Statistics and Social Statistics), American Statistical Association; 1998. Issues in Case-Mix Adjustment of Measures of the Quality of Health Plans; pp. 56–65.