



Enrollee Satisfaction Survey System: Options and Considerations

April 26, 2013

Under the Affordable Care Act (ACA), Exchanges must administer and report on an enrollee satisfaction survey¹. The U.S. Department of Health and Human Services intends to issue, but has not yet publicly released, proposed rules on this reporting requirement. MNsure will implement an enrollee satisfaction system in the future and the earliest point at which this can practically occur is in early 2015.

Currently, there are few health plan enrollee satisfaction surveys in the public domain. There are proprietary instruments, such as J.D. Power and Associates Member Health Plan Study, as well as surveys used internally by health plans for quality improvement efforts; however, the only NQF-endorsed patient-reported survey on health plan quality is the Consumer Assessment of Healthcare Providers and Systems (CAHPS®) Health Plan Survey. This instrument is the most likely choice for an enrollee satisfaction survey for a number of reasons. Not only is it one of few health plan experience surveys in the public domain, but it is also required for health plan accreditation, both by the National Committee for Quality Assurance (NCQA) and URAC. As QHPs are required under ACA to be accredited by either NCQA or URAC², the CAHPS survey will have to be conducted already for carriers offering qualified health plans on MNsure. Parsimonious with this requirement, 15 of the proposed measures for MNsure's Quality Rating System (QRS) are drawn from the CAHPS Health Plan Survey.

The CAHPS Health Plan Survey is maintained by the Agency for Healthcare Research and Quality (AHRQ) under the CAHPS program, an initiative "to support and promote the assessment of consumers' experiences with health care."³ The survey was first published in 1998 through a collaborative effort between AHRQ, Westat, Harvard Medical School, the Research Triangle Institute, and RAND. Since that time, the survey has been tested with consumers, reviewed by stakeholders, and accordingly revised from the original instrument. The survey has now expanded from focusing solely on the commercial

¹ <http://www.healthcare.gov/law/resources/regulations/guidance-to-states-on-exchanges.html>

² MNsure will require issuers to obtain the appropriate level of accreditation in the third year after the issuer offers a QHP on MNsure. Minnesota Health Insurance Exchange Plan Certification Guidance, October 9, 2012.

³ <http://cahps.ahrq.gov/about.htm>

health plan adult population to include Medicaid and pediatric versions of the survey as well. The CAHPS instrument is NQF-endorsed and considered a national standard for measuring health plan performance, most notably used by NCQA in health plan accreditation and public reporting of health plan performance in NCQA's annual plan rankings and plan report cards. The Office of Personnel Management (OPM) annually reports CAHPS results for all Federal Employee Health Benefits (FEHB) plans with 500 or more members so that federal employees can use this information when choosing a health plan⁴. The Centers for Medicare & Medicaid Services (CMS) also conducts a modified version of the CAHPS Health Plan Survey for measurement and public reporting of Medicare Advantage and prescription drug plans.

When the CAHPS Health Plan Survey is used by health plans for Healthcare Effectiveness Data and Information Set (HEDIS) reporting or NCQA accreditation, it is referred to as the 5.0H ("H" indicating "HEDIS") version of the survey instrument, as opposed to the 5.0 version of the instrument. The surveys are largely the same, but there are stricter protocol requirements to conduct the HEDIS/CAHPS survey, as well as some additional required questions. This memo will focus on the HEDIS/CAHPS version of the survey because a) this is the version of the survey required for QHP accreditation by NCQA (which is currently the market standard for accreditation among Minnesota carriers)); and b) there are proposed measures in the QRS straw model that are required in the HEDIS version of the survey, but not in the AHRQ version of the survey.

Conducting the HEDIS/CAHPS Survey

NCQA's final recommendations around the CAHPS survey for Exchanges are currently unknown. Current documentation seems to indicate that they will recommend the survey be conducted at the Exchange-product level. NCQA has strict rules that must be followed in order for CAHPS data to be considered HEDIS. These rules have been largely unchanged since they were first introduced in 1999 and it is likely that NCQA will continue with the current CAHPS rules for survey results to be considered HEDIS for carriers participating in Exchanges.

Survey Population

The NCQA HEDIS/CAHPS system includes six surveys:

- Adult commercial

⁴ <http://www.opm.gov/healthcare-insurance/healthcare/quality-healthcare/fehb-member-survey-results/>

- Child commercial
- Child commercial with chronic conditions
- Adult Medicaid
- Child Medicaid
- Child Medicaid with chronic conditions

The commercial and Medicaid survey instruments differ slightly in content. The commercial survey asks respondents to consider the last 12 months while the Medicaid survey asks respondents to consider the last six months. This difference is reflected in the eligibility criteria for the survey which specify that commercial members must be enrolled for the previous 12 months while Medicaid members need only be enrolled for the previous 6 months.

Children are defined as plan members 17 years old or younger at of the end of the measurement year. The child survey is rarely fielded for commercial plans. The child survey is not part of NCQA accreditation for commercial plans. Several of the measures included in the proposed QRS like the flu shot and smoking cessation measures are not asked on the child survey. The child survey takes the same concepts addressed on the adult survey and adapts the question to refer to parents' experiences with their child's health plan. Child commercial survey results are highly correlated with adult commercial survey results.

The survey for children with chronic conditions also is not commonly conducted for the commercial population. This survey returns two sets of results. The first set of results is for the general population and these results are directly comparable to the child commercial survey.

The second set of results is for children with chronic conditions and addresses the patient's experience in getting the care they need in managing these conditions. These measures are not included in the adult commercial survey and are not proposed in the QRS straw model. The chronic condition survey is much longer than the standard child survey and requires a much larger sample size, thus making the child commercial with chronic conditions survey considerably more expensive to conduct than the child commercial survey.

This discussion will focus on the NCQA requirements for the HEDIS/CAHPS adult commercial survey, as it is the most relevant survey for the MNSure population.

Sampling

In order to be eligible to be included in the survey, a plan member must be 18 years of age or older as of the end of the measurement year and have been continuously enrolled in the plan for the measurement year (one gap of up to 45 days is allowed). The standard sample size for the adult commercial survey is 1,100. Carriers are allowed to oversample in increments of five percent. Plans choose to oversample for two reasons. First, the carrier may simply want to increase the number of respondents (especially when the carrier has had a historically low response rate on surveys). Another common reason is to offset the estimated number of members that will be identified as disenrolled (and subsequently removed from the sample); this option is utilized by carriers that are unable to identify disenrollees in a timely manner before the sample is first drawn.

Random samples are drawn from the eligible sample frame. Only one plan member per household can be included in the sample.

Survey Fielding

Plans have the option of selecting one of two basic protocols – mail-only and mixed methodology. The mail-only protocol consists of a survey mailing, a reminder postcard, a second survey mailing, a second reminder postcard, and a third survey mailing. Surveys fielded following a mail-only protocol must be in the field for at least 81 days. The mixed methodology protocol consists of a survey mailing, a reminder postcard, a second survey mailing, and a second reminder postcard followed by three attempts to contact the sample member by phone. Surveys fielded following the mixed methodology protocol must be in the field for at least 71 days.

NCQA allows plans to add NCQA-approved enhancements to help increase response rates. Steps cannot, however, be dropped from the standard protocol. Commonly employed enhancements include extra mailings or phone attempts, use of the Internet for data collection, or use of Spanish language surveys. Spanish is the only non-English language that the survey can be conducted in and NCQA-provided translations must be used.

NCQA also allows carriers to add up to 20 NCQA-approved supplemental questions to the survey. NCQA reviews the questions to try to ensure that they will have minimal impact on response rate and how the core questions will be answered.

NCQA Rules

NCQA requires that the survey be conducted by an NCQA HEDIS/CAHPS certified survey vendor. There are 16 certified vendors eligible to conduct the survey in 2013. Certified vendors are required to undergo NCQA training, pay an annual certification fee, and pay a submission fee for each survey sample submitted to NCQA.

For CAHPS survey results to be considered HEDIS, they must be submitted by the NCQA-certified vendor to NCQA in the NCQA-specified format. Survey results can be considered HEDIS but not used for NCQA accreditation or inclusion in NCQA's Quality Compass unless the carrier-provided sample frame is approved by an NCQA-approved auditor prior to sampling.

The survey package includes a cover letter, survey instrument, and postage-paid reply envelope. The cover letter is printed with the carrier logo and carries the signature of a high-level carrier executive.

NCQA does not require carriers to achieve a specific response rate or number of completed surveys. Response rates for adult commercial samples averaged 30 percent in 2012 and have been declining in recent years. With a standard sample size of 1,100, this means an average of 330 completed surveys. NCQA does require that each individual measure have a minimum of 100 responses in order to be considered reportable. The survey instructs respondents to skip questions where they did not have a certain experience. Some questions are skipped by as many as 70 percent of respondents.

Survey Cost

The cost of the survey is dependent on the protocol chosen, the enhancements added to the basic protocols, the number of supplemental questions added, and the oversampling rate. The cost to carriers of fielding the basic mail-only survey is approximately \$11,000. The cost of fielding the basic mixed-methodology survey is approximately \$15,000. This does not include the cost to the carrier of having their sample frame audited if they would like their survey data used for NCQA accreditation.

Measures in the CAHPS Health Plan Survey

The primary indicators of enrollee *satisfaction* on the CAHPS survey are referred to as the "Global Ratings" items and ask enrollees to rate their health care on a 0-10 scale. These include: Rating of All Health Care, Rating of Personal Doctor, Rating of Specialist, and Rating of Health Plan. Satisfaction ratings often reflect enrollees' expectations and emotions regarding their health care, and are difficult to interpret in regards to how health plan performance could be improved. For this reason, the CAHPS survey goes beyond satisfaction to measure enrollee *experience*, which focuses on how often certain

events occurred during an enrollee's interaction with the health care system. Most questions ask about the enrollee's experience in the last 12 months so that responses are time-bound. For example, "In the last 12 months, how often did you get an appointment to see a specialist as soon as you needed?" Asking specific, time-bound questions about an enrollee's experience makes it easier for carriers to identify actionable areas for improvement and allows plan scores to reflect the quality improvement initiatives introduced by the carrier more quickly.

Many of these questions are rolled up into topic areas that summarize the carrier's performance in those areas. These summary measures are referred to as "composites" and reflect performance in the areas of: Claims Processing, Customer Service, Getting Care Quickly, Getting Needed Care, How Well Doctors Communicate, Shared Decision Making, and Plan Information on Costs. Each composite is the average of two or more "question summary rates" that indicate the proportion of enrollees that selected a favorable response choice (generally "Always" or "Always" + "Usually"). Most individual question summary rates are included in a composite, but there are two exceptions of individual measures standing alone in their respective topic areas on the survey: Health Promotion and Education & Coordination of Care.

In addition to experience of care measures, the survey also includes "effectiveness of care" measures that assess whether or not enrollees are receiving the appropriate treatment to prevent certain diseases. The Aspirin Use and Discussion measure assesses whether aspirin use is being managed effectively to prevent cardiovascular disease. The Medical Assistance with Smoking and Tobacco Use Cessation measures assess whether health care providers are encouraging patients to quit smoking. The third effectiveness of care measure, Flu Shots for Adults Ages 50-64, represents the number of enrollees in that age bracket that have received an influenza vaccination. The effectiveness of care measures use a rolling average methodology. This method allows carriers to collect data on the measure for two years before calculating results, in order to obtain sufficient sample size (a denominator of 100 is required to calculate and report on a HEDIS/CAHPS measure). The rolling average methodology is also used by NCQA for the Plan Information on Costs composite, as this composite only includes enrollees that have looked for information about how much they would have to pay for a health care service, equipment, or a prescription medicine.

The effectiveness of care measures have been added by NCQA to the HEDIS/CAHPS survey; they are not included in the core AHRQ CAHPS survey. Similarly, the questions that make up the Shared Decision Making, Claims Processing, and Plan Information on Costs composites are also present on the

HEDIS/CAHPS instrument but not the AHRQ instrument. All of the composites and effectiveness of care measures from the HEDIS/CAHPS survey are proposed in the QRS, with the exception of the Aspirin Use and Discussion measure.

Scoring CAHPS Health Plan Survey Measures

There are two basic methods for calculating scores for questions on the CAHPS Health Plan Survey: proportional scoring and mean scoring.

1. Proportional Scoring

- a. Proportional scores reflect the number of respondents that gave a favorable response out of the total number of respondents that answered the question. For example, a proportional score would reflect the number of respondents who said “Always” when given the options of “Never,” “Sometimes,” “Usually,” or “Always” to the question, “In the last 12 months, how often did you get an appointment to see a specialist as soon as you needed?” This type of score is called a “question summary rate” in the HEDIS specifications, but is often referred to as the “top box score.”
- b. Proportional scoring is the recommended method described in the AHRQ CAHPS Kit. It is also how scores are displayed in NCQA’s Quality Compass, a tool to analyze and compare health plan performance on HEDIS data.
- c. The proportional scoring method is consistent with how clinical HEDIS measures are calculated and displayed for public reporting by community health collaboratives such as MN Community Measurement. Top box scores are often used for public reporting because the results are intuitive to consumers (“How often did other patients have a favorable experience?”) and easier to understand than mean scoring.
- d. The top box score reflects how often the plan performs well, as opposed to a reflection of performance on-average (as in mean scoring). Proportional scoring is flexible in that there can be a looser or stricter definition of “good” performance. For example, on a 0-10 rating question, the top box score could be: i) Responses of 8, 9, or 10; ii) responses of 9 or 10; or iii) responses of 10 (a very strict standard that is not commonly used).
 - i. The ability to fine-tune the definition of “good” performance is ideal when all carriers perform well on a measure, as is the case for many measures in Minnesota. For example, if most carriers score high when the proportional score

is defined as responses between 8 and 10, then variation between plans can be created by narrowing the definition to responses between 9 and 10.

2. Mean Scoring

- a. Mean scores reflect the average response among all responses to a question. Each response choice is assigned a numeric value (e.g. “Never” = 1; “Sometimes” = 2; “Usually” = 3; and “Always” = 4) and the mean of these values is calculated. For example, if four health plan members answered a “Never/Sometimes/Usually/Always” question, each with a different response, then the mean score would be $(1+2+3+4)/4 = 2.5$.
- b. CMS employs mean scoring to calculate CAHPS scores that are the basis for star ratings of Medicare contract performance, which in turn are publicly reported and used in pay for performance initiatives.
- c. Mean scoring is also the basis for NCQA Accreditation scoring. This mean scoring method is different from the CMS mean scoring method in that all responses are re-coded on a 1 to 3 scale (1= worst; 3 = best) before the responses are averaged together. For example, the response choices are assigned numeric values between 1 and 3 (e.g. “Never” = 1; “Sometimes” = 1; “Usually” = 2; and “Always” = 3). If four health plan members answered a “Never/Sometimes/Usually/Always” question, each with a different response, then the mean score would be $(1+1+2+3)/4 = 1.75$.

Risk Adjustment and CAHPS

Data on a limited set of demographic characteristics are collected as part of the CAHPS survey. The relationship between enrollee demographic characteristics and responses to the CAHPS survey has been well documented. Individuals in better health and older individuals tend to give higher ratings, whereas individuals with higher education and women tend to give lower ratings. Sometimes, there may be substantial differences in these enrollee characteristics across plans. Case-mix adjustment is a way to control for differences in enrollee characteristics when comparing plans so that differences in scores reflect quality of care differences rather than enrollee characteristic differences. In the CAHPS consortium documentation, it is recommended that results be case-mix adjusted. Examples of variables that could be adjusted for include age, education, gender, and health status. Although there is also a well documented relationship between race and enrollee responses (e.g. Asian respondents tend to give

lower ratings), the CAHPS team does not recommend case-mix adjusting for race and, according to the CAHPS team, race historically has not been used as an adjustor in public reporting.

NCQA does not adjust CAHPS scores. Historically there would not have been a large impact on scores for most commercial plans as a result of case-mix adjustment due to the homogeneity of enrollee characteristics across plans.

MNsure should consider the impact of case-mix adjustment in Minnesota and determine whether the value added from case-mix adjustment offsets potential confusion created if carriers have multiple sets of scores (NCQA-unadjusted scores and MN-adjusted scores).

Discussion Questions

- Should the protocol for MNsure's enrollee satisfaction survey be more standardized? For example, should all carriers be required to use the mixed-methodology for fielding the survey (as opposed to plans having the option to do mail-only or mixed)?
- Should MNsure adopt proportional or mean scoring?
- Should results of the survey be case-mix adjusted?