

IMPLEMENTATION OF MEDICARE CAHPS FEE FOR SERVICE SURVEY

FINAL REPORT FOR YEAR 1 Executive Summary

Contract No. 500-95-0061/T.O.#7
RTI Project No. 7903

Submitted to:

Edward S. Sekscenski
HCFA/CBS/BPPAG
7500 Security Boulevard
Mail Stop S1-15-03
Baltimore, MD 21244

Prepared by:

RTI
and
RAND

Under Subcontract to:

Center for Health Systems Research and Analysis
University of Wisconsin-Madison

October 29, 2001

Executive Summary

This report summarizes the methodology and findings of the 2000 Medicare Consumer Assessment of Health Plan Surveys (CAHPS) Fee-for-Service (MFSS) Survey conducted for the Centers for Medicare & Medicaid Services (CMS) (formerly Health Care Finance Administration (HCFA)) by RTI with the assistance of RAND, NCS Pearson, and Discovery Research Group (DRG). The work was performed under subcontract to the Center for Health Systems Research and Analysis at the University of Wisconsin-Madison under UW's contract with CMS/HCFA. More detailed information on many of the topics presented in this report is available in the individual project reports referenced in **Chapter 9** of this report.

Questionnaire Development

The 2000 MFSS Survey used questions based on the CAHPS questionnaire for adult, privately-insured populations developed as part of the CAHPS development project sponsored by the Agency for Healthcare Research and Quality (AHRQ). A field test was conducted prior to the survey to test field procedures and to evaluate the psychometric performance of standard CAHPS questions. Data collection for the field test was conducted between August and December of 1998 on a sample of Medicare fee-for-service (FFS) beneficiaries.

Field test participants received one of two randomly administered versions of the survey instrument: The majority (n=1971) were asked to report on experiences during the past 6 months, while the rest (n=381) were asked to report on experiences during the past 12 months or without a specified recall period. In addition, some experimental phrasings of the CAHPS core items and new items were tested. Although similar response patterns were obtained for the CAHPS report and rating items for both versions of the instrument, cognitive testing prior to the field test revealed that a 12-month recall period could impose a greater cognitive and response burden for Medicare beneficiaries. These findings, in combination with a lower response rate among persons randomized to the 12-month or unspecified recall period, led to the recommendation of a 6-month recall period for health care experiences for the 2000 MFSS Survey.

Based on the field test, the wording of some CAHPS questions was slightly revised to make them more applicable to the Medicare population. In addition, some questions in the Medicare CAHPS Managed Care (MMC) Survey questionnaire were excluded from the MFSS

Survey questionnaire, and vice versa, depending on the analysis needs of the project prior to the national implementation of these surveys. Subsequent changes made to the CAHPS questionnaire for the 2001 MFFS Survey are presented in **Chapter 2** of this report.

Sample Selection and Weighting

The 2000 MFFS selected a sample of 167,993 from a sampling frame constructed from the August, 2000 version of CMS/HCFA's Enrollment Database (EDB). The frame comprised 30.1 million persons enrolled in Medicare fee-for-service for at least the prior 6 months and who resided in the U.S. or Puerto Rico. A total of 280 geographical units (geounits) were constructed (275 in the U.S. and 5 in Puerto Rico) to allow CAHPS outcomes to be compared both within the MFFS subpopulation and between the MFFS and MMC subpopulations for small, meaningful areas.

The goal was to obtain a minimum of 300 responses in each sampling unit so that ratings and composites can be calculated. This resulted in the need for beneficiaries in rural counties and less populous states to be sampled at higher rates than beneficiaries in urban counties and populous states. Therefore, an initial sampling weight was assigned to each selected beneficiary as the inverse of the selection probability, reflecting the differential selection rates that were used to select beneficiaries from each geographic area.

The response rates for the MFFS Survey varied considerably with respect to urbanicity (rural counties higher than urban), race (Whites higher than other races), age (younger beneficiaries higher than older), dual Medicare/Medicaid eligibility, and region of the country (Midwest higher than others). As a result, the respondent distribution is composed of too many Whites, too few dual eligibles, and too many beneficiaries from the Midwest when compared to the original sample distribution. To reduce the potential biasing effects of differential nonresponse, the initial sampling weights of respondents were post-stratified to 338 separate counts of the total number of Medicare FFS beneficiaries obtained from the October 22, 2000 version of the EDB. The counts included totals for each of the geographic areas in the U.S. and Puerto Rico as well as totals formed by the intersection of the age, gender, race, and dual Medicare/Medicaid eligibility factors. Details of the sample selection and weighting activities are presented in **Chapter 3**.

Data Collection

The primary mode of data collection for the MFFS Survey was a self-administered mail survey. Respondents were given the option to complete the survey by telephone to facilitate inclusion of the most possible sample members—for example, sample members with vision, reading, or other impairments that might otherwise preclude their participation. A Spanish-language version of the questionnaire was also offered. We followed up with nonrespondents to the mail survey for whom we had a telephone number by telephone (in English and Spanish), and with others by overnight mail.

The data collection period for the MFFS Survey began on October 9, 2000, with the mail-out of the prenotification letter and ended on February 1, 2001, with the close of the telephone follow-up. The overall response rate among all eligible beneficiaries was 63.9 percent. The rates varied somewhat among the geographic areas from which randomized subsamples were drawn; however, response in each area was sufficient to provide measures of CAHPS composites and ratings for all geographic areas in the U.S. and Puerto Rico.

The data collection effort achieved a response rate of 61.6 percent among all sample members (including ineligible) and a response rate among eligible sample members of 63.9 percent. The following table provides a summary of the data collection results by wave.

	First Mailing	Second Mailing	Third Mailing	Inbound Phone	Outbound Phone	Total Activity
Initial Cases	167,993	124,503	7,773	2,191	48,471	167,993
Completed Cases	66,137	23,788	1,168	1,232	11,226	103,551
Response Rate	39.4%	14.2%	0.7%	0.7%	6.7%	61.6%

Notable among these results was the lack of telephone numbers in CMS/HCFA administrative files. Despite a concerted tracing effort, telephone matching yielded current numbers for only about half of the nonrespondents to the mail portion of the MFFS Survey. One reason for this may have been that 25 percent of the sample were 80 years old or older and another 14 percent were identified as “dual-eligible.” The elderly and persons with low incomes can be very difficult to find because they often do not have telephones or credit histories to trace. As a result, tracing methods that match names to public records and search credit bureau

information are often not very successful. Details of the data collection activities are presented in **Chapter 4**.

Case Mix Adjustments

The two applications of Case-Mix Adjustment (CMA) to the 2000 MFFS Survey (within-MFFS comparisons and MFFS-versus-MMC comparisons) suggested two distinct, but similar, CMA models. The case-mix adjusters currently employed in MMC CMA (age, education, self-rated health status, and proxy respondent status) constituted an effective case-mix model for both comparison purposes. An indicator of dual eligibility further enriched the within-MFFS model. A self-rated mental health item demonstrated the potential to improve both models in the future.

The 2000 MFFS Study found that the assumptions behind CMA were satisfied or could be accommodated with simple adjustments. Furthermore, the MFFS CMA models appeared to be suitable presenting results in a variety of forms and at a variety of levels of aggregation.

While the direction of CMA coefficients was similar for MFFS and MMC, the magnitudes of the effects sometimes differed. In particular, the well-established tendency of healthier beneficiaries to rate their care more positively or to report better health care experiences was considerably stronger in MMC than in MFFS. In other words, personal satisfaction with health care was much more sensitive to health status in MMC than in MFFS. Because of this difference and the generally poorer health status of MFFS beneficiaries (even excluding the dually eligible), CMA tended to make small adjustments in favor of MFFS relative to MMC.

The existence of strong and different case-mix effects for health status between MMC and MFFS suggests that we should consider stratified reports by beneficiary health status. In fact, the Subgroup Analysis Report demonstrates that a “cross-over” occurs in many instances: Less healthy beneficiaries are more satisfied with MFFS than with MMC, whereas healthier beneficiaries are more satisfied with MMC than MFFS. Details of the MFFS case-mix adjustment activities are presented in **Chapter 5**.

Analysis of Geographic Units

The results of the geounits analyses, which are consistent across the various procedures used, indicate that the vast majority of variability in the CAHPS outcomes is at the individual level. For higher levels of geographical aggregation, geounits tend to look alike within a

particular state with respect to responses on the CAHPS measures. Those geographic differences present are attributable to differences in the composition of beneficiaries.

While the geounits do not contribute in any statistically meaningful way for purposes of analysis, they are essential for the creation of comparisons to MMC. Because the criteria for aggregating counties imposed many constraints, there are few alternatives for the creation of geounits that will allow comparison to MMC. The current geounits perform well in that respect and should be modified only to conform to the changing MMC landscape. Details of the analysis of geographical units are presented in **Chapter 6**.

Subgroup Analyses

The 2000 MFFS Survey data were analyzed to gain an understanding of the differences in health services experience and satisfaction among Medicare beneficiaries, according to geographic levels, socio-demographics, plan options, and health status. The MFFS population is quite heterogeneous in terms of demographic characteristics, region of residence, supplemental insurance (employer-based, private, or Medicaid), and health-related characteristics. These subgroups may have vastly different experiences with and expectations of the health care system and, thus, may perceive the quality of and access to services differently.

The data generated by the 2000 MFFS Survey should be usable for quality improvement, accountability, and beneficiary information. These goals required that data be reported on a number of levels of aggregation, including geographic sampling units, state, region, and nation. In markets where there was sufficient MMC penetration to offer choices to beneficiaries, the aggregation enabled MFFS and MMC comparisons.

Ratings and composites were constructed using the CAHPS 3.0 macros, case-mix adjusted and weighted, for the following measures:

- Rate Personal Doctor (Q 7)
- Rate Specialist (Q 11)
- Rate Health Care* (Q 30)
- Rate Medicare* (Q 46)
- Needed Care composite* (Q 21, 22, 4, 9)
- Good Communication composite* (Q 26, 27, 28, 29)

* Indicates composites or ratings featured on the Medicare Compare website (<http://www.medicare.gov/mphCompare/home.asp>).

- Care Quickly composite* (Q 14, 16, 18, 23)
- Respectful Treatment composite (Q 24, 25)
- Medicare Customer Service composite (Q 41, 43, 45).

Notable findings from these analyses include the following:

- Across geounits, states, and CMS/HCFA regions, a consistent pattern emerged among MFFS beneficiaries with the Needed Care composite having the highest percentage of most positive responses and Rate Medicare having the lowest percentage of most positive responses.
- Ratings and composites vary by subgroups of MFFS beneficiaries; differences in ratings and composites were found by insurance status (dually eligible, with versus without insurance in addition to Medicare), self-reported health status, race, and age. However, these differences were not always consistent.
- With the exception of Medicare Customer Service, no more than 20 percent of MFFS beneficiaries responded negatively to all CAHPS performance indicators and ratings.
- MFFS beneficiaries who are younger, more educated, in poorer health, and who do not have a personal doctor are generally less satisfied with MFFS than their counterparts.
- On a national level, neither MFFS nor MMC beneficiaries consistently provided more positive responses across all indicators.
- Beneficiaries in excellent/very good health perceive their plans and the care they receive differently than those in fair/poor health. Generally, a larger proportion of beneficiaries in fair/poor health give MFFS higher ratings while a larger proportion of those who rate their health as excellent/very good give MMC higher ratings.

Conclusions of the Subgroup Analysis

Findings from the geographic-level analyses indicate that notable differences exist among MFFS beneficiaries across all geographic aggregation options for Rate Medicare, Rate Personal Doctor, Rate Specialist, and the Medicare Customer Service composite. Also, when beneficiaries' responses to ratings and composites are aggregated to state, region, and national levels, the differences across the geographic levels are still present but mitigated.

Findings from the individual-level analyses suggest that satisfaction and experience with MFFS are affected by socio-demographic characteristics, health status, and insurance type. Younger beneficiaries are less satisfied than older beneficiaries; beneficiaries with lower levels

of education rate Medicare higher than the more highly educated; and males rate Medicare lower than females. Healthier beneficiaries were more satisfied and rated Medicare higher than less healthy beneficiaries. Beneficiaries with a personal doctor were more satisfied than those without a personal doctor. Beneficiaries living in metropolitan areas were less satisfied and rated Medicare lower than those living in rural areas. MMC penetration rates were also associated with satisfaction and ratings, with those living in areas with higher MMC penetration reporting higher levels of satisfaction and rating Medicare higher than those living in areas with less than 25 percent managed care penetration.

The comparative analysis of the five composites and ratings, along with the Flu Shot indicator, illustrate statistically significant differences ($p < 0.05$) in the satisfaction and experience reported by beneficiaries in MFFS and MMC. In general, a higher percentage of beneficiaries in MFFS (compared to MMC enrollees) reported “Not a Problem” for the Needed Care composite and reported “10” for Rate Medicare. On the other hand, a higher percentage of beneficiaries in MMC (compared to those in MFFS) reported “Always” for the Good Communication composite, reported “Yes” for the Flu Shot indicator, and assigned a “10” for Rate Health Care. For one composite, Care Quickly, neither MFFS nor MMC was clearly better or worse. These findings suggest that MFFS beneficiaries are more satisfied with Medicare and health services access in general, while MMC beneficiaries are more satisfied with their health provider interaction and may receive more preventive measures.

Findings from the analysis comparing MFFS with MMC by health status suggest that beneficiaries who are in fair/poor health and those in excellent/very good health perceive their plans differently. In general, beneficiaries in fair/poor health reported better experiences and higher levels of satisfaction with MFFS than with MMC. On the other hand, beneficiaries in excellent/very good health rated MMC higher than MFFS most of the time. These data can be used to examine state-level trends and initiatives that can influence beneficiaries’ experience with and perceptions of their choice of health plan. Details of the MFFS subgroup analysis are presented in **Chapter 7**.

Encouraging PROs to Use CAHPS Data for Quality Improvement

The original goal of this task was to gain a better understanding of how CAHPS was viewed and understood by the Peer Review Organizations (PROs), and then to develop a model

for enabling them to use CAHPS data in their Quality Improvement (QI) projects. By the time of the November 30, 2000 Technical Expert Panel (TEP) meeting held in Baltimore, TEP members had been briefed at the previous year's American Health Quality Association annual meeting on the results of focus groups on a very similar topic conducted by the Picker Institute with PRO staff. This focus group report suggested that PRO staff did not have much knowledge of or experience with CAHPS data, and that there was no great interest in becoming more familiar with the data in the context of QI. In the discussion surrounding our task plan presentation at the TEP meeting, two new, more promising possible directions emerged. One was to see how private health plans are using CAHPS data for QI purposes and to assess parallels for PROs. The second was to make an effort to tie CAHPS data in some way to clinical or preventive care.

To identify private health plans that were using CAHPS data for QI purposes, we reviewed the 1999 project summaries of the 39 organizations with a summary posted on the CAHPS Users Group website (<http://www.cahps-sun.org/>). Private health plans generally indicated that CAHPS was performed in order to obtain accreditation and for promotional reasons. However, three health plans did mention using CAHPS for QI purposes. We contacted and interviewed representatives of those three health plans, and we found that CAHPS data were being used by health plans along with other information to identify areas needing improvement *within* the health plans. Most typically, health plan attention was directed at improving their overall health plan ratings in order to raise their accreditation score. Also, we found that other surveys were often conducted with "CAHPS-like" items in order to get closer to identifying operational problems. It was felt that the CAHPS survey identified problem areas but was "too high level" to actually identify the roots of the problems.

We also examined whether variations in two of the CAHPS service quality measures in the 2000 Medicare Fee-for-Service CAHPS Survey were related to variations in health behavior. In particular, we explored the extent to which the CAHPS measures of the communication skills of primary care physicians and the helpfulness and respectfulness of their office staff are associated with better compliance in the use of screening mammography. This analysis was done at the level of the geographic areas (the 275 counties and county aggregates) used to sample for the Medicare CAHPS. The rates of mammography use for the same 275 geounits were obtained from a report prepared by Health Economics Research (HER) entitled *Performance Measurement in Medicare Fee-for-Service: Biennial Mammography Screening Rates for 1998-*

1999. In addition to these variables, a number of others could confound the analysis of the relationship between service quality and mammography use using data obtained from the CAHPS and other sources.

We analyzed the relationships between service quality and mammography use employing the SAS multiple linear regression procedure with a dichotomous value for each level of categorical predictor variables. We tested two models using slightly differently calculated CAHPS quality measures of physician communication and staff helpfulness. In the first model (to predict the rate of screening mammography use among 52- to 69-year-old Medicare beneficiaries in the 275 geographic areas), the overall model analysis of variance was highly significant (F-Value = 8.40, DF 17/256, $p < 0.0001$), with an adjusted $R^2 = 0.31$. This model accounted for just less than one-third of the variance in mammography use rates. The second model overall analysis of variance was highly significant as well (F-Value = 9.23, DF 17/257, $p < 0.0001$), with an adjusted $R^2 = 0.34$. The second model accounted for just over one-third of the variance in the geographic area rates of mammography use.

We also repeated the estimation of both models on two population subgroups for which we had mammography use rates (all white women, and all women aged 65 to 69). The subgroup models were also all significant and the R^2 s ranged from 0.24 to 0.38. Results were largely consistent with the two models for the overall group of women aged 52 to 69.

The models we have estimated successfully explain considerable variance in the rates of mammography use at the county or county group level at which the Medicare Fee-for-Service CAHPS data were collected. However, census division was the most consistently significant variable and likely accounts for most of the differences in rates that the model explains. Only one of the two CAHPS service quality measures rates—physician communication—was associated with the mammography rate, but it was only significant in half of the models.

Our analyses conducted thus far are not conclusive with respect to the association of CAHPS service quality measures. In the future, we propose to obtain individual-level preventive service use outcomes to analyze with individual-level CAHPS scores rather than the geographic area or ecological measures we examined in this analysis.

Among the preventive health behaviors we will focus on next are some included in the CAHPS survey (receipt of a flu shot, pneumonia immunization, and smoking cessation

counseling for smokers), and others extracted from Medicare claims data (mammography screening, treatment of depression, and diabetes care). Because person-level analysis may capture relationships between individual beneficiaries and their providers, the objective of this analysis will be to determine whether selected dimensions of service quality collected in the CAHPS survey are associated with the use and receipt of primary and secondary preventive services at the individual (person) level. To the extent that the selected service quality dimensions are associated with the use of preventive services at the person level, we will have established an empirical basis for recommending to policymakers the more widespread use of CAHPS for health care quality improvement purposes. More details of the MFFS quality improvement activities are presented in **Chapter 8**.

