MEDICARE
HEALTH OUTCOMES SURVEY

HOS/CAHPS®
SURVEY INTEGRATION
FORMATIVE STUDY DESIGN
Final Report

ORIGINALLY ISSUED: DECEMBER 2004
PUBLIC DISTRIBUTION: MARCH 2005

HSAG
HEALTH SERVICES
ADVISORY GROUP

CMS
CENTERS for MEDICARE & MEDICAID SERVICES

TABLE OF CONTENTS

HOS/CAHPS® SURVEY INTEGRATION FORMATIVE STUDY DESIGN

FINAL REPORT

MARCH 2005

EXECUTIVE SUMMARY

Background

The Centers for Medicare & Medicaid Services (CMS) continuously reviews the beneficiary health surveys conducted under its auspices in terms of the value provided for the resources invested. The Medicare Health Outcomes Survey (HOS) and the Medicare + Choice (M+C) CAHPS®[1,2] surveys are both targeted at beneficiaries enrolled in M+C plans. It is therefore reasonable for CMS to consider whether or not these two survey instruments should be integrated into a single questionnaire.

If feasible, the integration of these two surveys will be likely to reduce survey costs as well as reduce the burden on the beneficiaries. An integrated survey also would allow the survey's users to perform additional analyses of the survey data, such as:

- Determining the relationship between self-reported health outcomes and reports of health care experiences at the beneficiary level
- Incorporating health status measures into the case mix adjustment of health plan comparisons
- Identifying plans that are high performers in providing quality health care experiences to members of specific beneficiary subgroups, such as beneficiaries with diabetes, or beneficiaries with low self-reported health status

CMS requested that Health Services Advisory Group (HSAG) conduct a preliminary assessment to determine the feasibility of integrating the HOS and the M+C CAHPS surveys into one survey instrument. In response to this directive, HSAG gathered and reviewed information from a variety of sources, and this report summarizes our findings.

To evaluate the feasibility of integrating the HOS and M+C CAHPS surveys, we examined seven key aspects of these two surveys:

[1] CAHPS® is a registered trademark of the Agency for Healthcare Research and Quality.
[2] Formerly referred to as the Medicare Managed Care (MMC) CAHPS Survey.

- Questionnaire Content
- Survey Administrative Protocols
- Analytic Strategies
- Sampling Methods
- Cost and Burden
- Dissemination of the Results
- Uses of the Results

Sources of Information

Information was gathered from 1) key stakeholders, including both researchers and end users; 2) published and unpublished literature; and 3) analyses of data from the current HOS, M+C CAHPS, and Medicare Fee-For-Service (MFFS) CAHPS surveys.

Key Findings

Questionnaire Content

- An acceptable length for an integrated questionnaire is approximately 100 items. Since the current HOS and M+C CAHPS questionnaires each contain approximately 100 items, a substantial number of items would need to be deleted in order to achieve a practical survey length.

- The current HOS and M+C CAHPS surveys have distinctly different formats. If these distinct formats are maintained in an integrated questionnaire, the questionnaire may be confusing to the respondents. On the other hand, attempting to harmonize the formats may affect the comparability of the integrated survey items to their counterparts in the existing separate survey instruments.

- One of the versions of the Short-Form 12 Health Survey (SF-12) can replace the Medical Outcomes Trust (MOT) Version of the Short-Form 36 Health Survey (SF-36®[3]) in an integrated survey, with little loss of information or reliability. This by itself would reduce the integrated survey's length by 24 questions.

- Items currently used for accreditation purposes should have a high priority for retention in an integrated survey questionnaire.

- The formats of the MFFS and M+C CAHPS surveys have been kept highly similar to facilitate the comparison of CAHPS results between the MFFS and M+C beneficiary populations. If the M+C CAHPS survey is modified and integrated with the HOS, these same changes should be seriously considered for the MFFS CAHPS survey as well.

[3] SF-36® is a registered trademark of the Medical Outcomes Trust.

Survey Administrative Protocols

- The current HOS and M+C CAHPS surveys utilize a similar sequence of questionnaires and cover letters, reminder post cards, and follow up telephone calls. This will facilitate integration of the administrative protocols for the two surveys.

- However, the *research designs* for the two surveys are quite different: the HOS is a longitudinal survey consisting of a baseline survey and a follow up survey of the same respondents two years later; whereas, the M+C CAHPS is a cross-sectional survey that assesses the respondents at only one point in time. Integrating the two surveys will require choosing one of several options:

  1. Administering the HOS items on a longitudinal basis, and the M+C CAHPS items on a cross-sectional basis;

  2. Administering both surveys on a cross-sectional basis; or

  3. Administering both surveys on a longitudinal basis.

- A longitudinal survey research design requires a more complex administrative procedure than a does a cross-sectional design.

Analytic Strategies

- Integrating the two surveys means that data on both experiences of care and health status will be available at the beneficiary level. This permits a more granular analysis of the relationship between experiences of care and health status that would otherwise only be available at the plan level. An integrated survey would also make it possible to develop plan performance norms for beneficiaries with specific medical conditions, as well as norms for beneficiaries with specific levels of physical and mental health functioning.

- Analysis of HOS results at the plan level shows that some plans perform distinctly better or worse than others in maintaining the health of their beneficiaries over time. This indicates that the longitudinal change scores are measuring more than simple aging effects.

- A longitudinal design offers significant analytic advantages, including the ability to analyze determinants of change in health status. This design increases the precision of comparisons over time, compared to time trends calculated on different samples of beneficiaries. This increase in precision greatly enhances the ability to monitor the success of quality improvement (QI) initiatives. Another analytic advantage is the ability to determine the impact of experiences of care on subsequent health status, as well as the impact of health status on subsequent experiences of care.

  - However, focusing solely on change scores provides an inaccurate picture of plan performance, due to the attrition of sicker and less satisfied beneficiaries. Combining results from longitudinal, cross-sectional, and disenrollee surveys will provide the most accurate picture of plan performance.

- If a cross-sectional design is employed, demographic matching of separate samples of respondents at different time periods could be used to improve the accuracy of time trend comparisons.

  - The relative merits of cross-sectional, matched cross-sectional, and longitudinal designs can be tested with currently available data from the HOS.

- An integrated survey provides more variables for testing in case mix adjustment models. However, only those variables that make significant contributions to the adjustment of plan means should be incorporated. Also, model assumptions should be tested.

- Integrating the two surveys will require a review of the methods each uses for imputing missing values. These methods currently differ. Furthermore, the current HOS results are not weighted for non-response bias, whereas the M+C CAHPS results are. Integrating the two surveys does not require that a single method for missing value imputation and for non-response bias adjustment be adopted, but not doing so will complicate the analysis of the survey results.

Sampling Methods

- Once a survey research design has been chosen, the choice of sample size is straightforward. If a cross-sectional design is used, the current M+C CAHPS sample size of approximately 500 returned surveys will be adequate. If a longitudinal survey design is used, the current HOS sample sizes of approximately 750 scorable returned surveys at baseline and approximately 400 scorable returned surveys at follow up will provide meaningful results.

- The current M+C CAHPS is distributed to disenrollees as well as enrollees, whereas the HOS is distributed to enrollees only. If disenrollees were to be included in the survey sample for an integrated survey, this would further increase the sample sizes needed.

- Eligibility requirements differ slightly between the two surveys, and would need to be reconciled for an integrated survey.

Cost and Burden

- If the integrated questionnaire is kept at approximately the same length as each of the current questionnaires, then merging the two questionnaires will produce substantial cost savings for CMS, as well as a substantial reduction in respondent burden. Cost savings to CMS for a one-year period are estimated to be approximately $5,000,000. The annual reduction in respondent burden is estimated to be approximately 48,000 hours.

    - The costs of starting up the new survey would reduce the above cost savings to some extent.

- If disenrollees were to also be included in the integrated survey sample, then there would no longer be a need for the Medicare CAHPS Disenrollee Assessment Survey. This would result in additional cost savings.

- Since the plans already pay a portion of the HOS survey cost, survey integration should be cost neutral from the plans' point of view, while lowering the cost to CMS.

- Use of a longitudinal survey research design would be more costly than a cross-sectional design, but also offers analytic advantages (see above). Costs can be further reduced by using public domain instruments and by consolidating the Technical Expert Panels (TEPs) and educational conferences for each survey.

Dissemination of the Results

- Distribution of the data and reports via the electronic media of the Health Plan Management System (HPMS) and the QualityNet Exchange (QNE) will facilitate the dissemination of results from an integrated survey. Both distribution channels are already in use for the current version of the HOS, and the HPMS is already used to distribute data from the M+C CAHPS.

- Since the M+C CAHPS results are currently being publicly reported, an integrated survey that contains both health status and CAHPS questions will renew interest in public reporting of the health status results as well.

Uses of the Results

- Unfortunately, efficient and rapid dissemination of the results via electronic means does not guarantee that the results will be acted upon. A combined educational conference can help to ensure that techniques for using the data to drive improvements are shared with the Quality Improvement Organization (QIO) and M+CO communities. Webcasting is another emerging technology that can facilitate use of the data for quality improvement (QI) activities, while minimizing travel costs. A guide to quality improvement strategies, similar in scope and

content to the *CAHPS® Improvement Guide*, would also help to propagate the use of QI techniques.

- Incorporating the use of the integrated survey results into formal CMS requirements for M+COs and QIOs via the Performance Assessment Monitoring Tool will provide further motivation for acting upon the survey results.

Summary of Preliminary Recommendations

Based on a review of the issues described above, the following preliminary recommendations are presented for review and further discussion by key stakeholders:

1. Convene a task force of stakeholders to review the preliminary recommendations outlined in this document, decide whether to proceed towards integration, and build a consensus regarding next steps.

   - The task force will need to carefully consider the potential impact of other survey initiatives, such as Version 2 of the HOS, the Ambulatory CAHPS (A-CAHPS) survey, and the prescription drug plan CAHPS survey.

   - Since it may be some time before the impacts of these other survey initiatives are fully known, it is suggested that the task force consider adding the SF-12 and the Activities of Daily Living (ADL) questions to the current M+C CAHPS questionnaire for the next possible survey cycle.

2. For the HOS:

   - Substitute the SF-12 for the SF-36®. This eliminates 24 items while still retaining the ability to measure the principal health status measures. The VA version of the SF-12 may be less costly to use than the MOT version, due to the fact that this questionnaire and its associated scoring methods are in the public domain.

   - Retain selected chronic medical conditions items. These provide M+CO staff with a quick snapshot of their beneficiary populations, and also permit the identification of high performing plans for specific subgroups of beneficiaries.

   - Retain the ADL questions to provide data for frailty-based payment systems.

3. For the M+C CAHPS:

   - Retain all items needed to create the CAHPS global rating and composite scores.

   - Consider elimination of the gate questions. There is evidence that these items improve data quality, but their retention will add substantially to survey length.

- Eliminate most of the appeals and grievances questions. Only a small number of these questions are needed to document attempts to address the mandates of the Grijalva decision.

4. Retain the demographic items that are needed for case mix adjustment for both the health status and experience of care measures. When selecting adjuster variables, use only those variables that produce statistically important contributions to the overall adjustment model.

5. Use multiple vendors to manage the data collection process. This helps to lower vendor fees by fostering competition, while ensuring that CMS benefits from an ability to compare vendor performance.

6. Continue to collect data by both mail and telephone, to ensure representation from those subgroups of beneficiaries that are more likely to respond to a telephone survey.

7. Administer the integrated survey on a longitudinal basis, and include disenrollees in the survey sample. This will enable researchers and decision makers to understand the impact of beneficiary attrition on the results. It will also give a more accurate picture of the performance of individual M+COs.

8. Each year, develop a formal analytic plan for the integrated survey. An analytic plan will help focus analytic resources on topics of the greatest value to plans, QIOs, and CMS.

9. Merge the TEPs and the educational conferences for the HOS and M+C CAHPS surveys. This will save money and foster the use of both experiences of care and health outcomes information by the same community of users.

10. Share the cost of the integrated survey with the M+COs. Since the plans already pay for a portion of the HOS, their total survey costs will remain roughly the same.

11. Establish an interactive page on the HPMS Web site for dissemination of data and results to plans and QIOs. Plans and QIOs could then use Online Analytical Programming (OLAP) tools to query and explore their results.

12. Publicly report the results for both the health outcomes and experiences of care questions. This will give the M+COs additional motivation to review and respond to the results.

13. Develop a User's Guide for the integrated survey that provides concrete methods for using the data to drive decisions. Use Webcasting to cost-effectively train a large number of users in these methods.

14. Tie use of the integrated survey results to CMS requirements. This will help to maintain QIO and plan focus on the survey.

Conclusion

Integration of the HOS and M+C CAHPS surveys appears to be quite feasible. However, a successful integration will require the key stakeholders to make some difficult choices. These choices include: which questions to retain and which questions to drop; what formatting changes to make; which survey research design to use; whether results should be publicly reported and how best to report them; and whether to incorporate use of the survey results into CMS contract requirements.

I. INTRODUCTION

CMS continuously reviews each of the beneficiary health surveys conducted under its auspices to identify opportunities for increasing the value received, as well as opportunities to reduce the associated costs and burdens. Since the HOS and the M+C CAHPS surveys are both targeted at beneficiaries of M+C plans, CMS asked HSAG to conduct a preliminary assessment of the feasibility of integrating these two survey instruments into a single questionnaire.

If feasible, the integration of these two surveys would be likely to reduce costs to the plans and at the same time reduce the burden on the beneficiaries being asked to respond. An integrated survey also would permit a variety of new uses of the data, such as:

- Determining the relationship between self-reported health outcomes and reports of health care experiences at the beneficiary level.

- Incorporating health status measures into the case mix adjustment of health plan comparisons.

- Identifying high performing plans in terms of experiences of care for specific beneficiary subgroups, such as beneficiaries with diabetes or beneficiaries with low self-reported health status.

To begin an examination of this question, CMS structured the current Statement of Work for the HOS to include a preliminary assessment of the feasibility of integrating the HOS and the M+C CAHPS surveys into one survey instrument. In response to this directive, HSAG gathered and evaluated information from a variety of sources. This report summarizes our findings.

Successful integration of the current HOS and M+C CAHPS surveys requires a careful review of the following aspects of the two surveys:

- Questionnaire Content
- Survey Administrative Protocols
- Analytic Strategies
- Sampling Methods
- Cost and Burden
- Dissemination of the Results
- Use of the Results

Purpose of This Report

The purpose of this report is to review and assess the potential impact of survey integration on each of the seven areas listed above. For each area, this report will 1) review the current status; 2) identify and discuss the key issues that must be addressed to successfully integrate the two surveys; and 3) make preliminary recommendations for resolving these issues in order to achieve integration of the HOS and M+C CAHPS questionnaires into a single instrument.

Sources of Information for This Report

The information presented in this report comes from three primary sources:

Interviews with Key Stakeholders: HSAG staff interviewed a total of 27 stakeholders who are users of and/or experts in the HOS and/or CAHPS surveys. The interviews were conducted by telephone using a standardized set of questions. The individuals that we interviewed represent 19 different organizations, including CMS, survey research organizations, universities, the Veterans Administration (VA), health plans, and QIOs. Attachment B contains a listing of the individuals interviewed and the organizations they represent.

Literature Review: A comprehensive review of the literature was also conducted. Documents reviewed included published papers, unpublished manuscripts, conference presentations, internal reports, survey questionnaires, survey administrative protocols, and Web sites.

In addition to the HOS and M+C CAHPS surveys, there are several other major Medicare survey initiatives that may be instructive regarding integration issues. Information regarding the following surveys was also gathered from stakeholders and relevant literature:

- The Medicare Fee-for-Service (MFFS) CAHPS Survey
- The Medicaid CAHPS Surveys
- The Medicare CAHPS Disenrollment Assessment Survey
- The M+C CAHPS Disenrollment Reasons Survey
- The Medicare Health Survey for PACE[4] and Evercare (MHSPE[5])
- The Survey of Health Experiences of Veterans (SHEP)

Statistical Analyses of Current Surveys: Currently available data from the MFFS CAHPS survey, the M+C CAHPS survey, and the HOS can provide some guidance as to the feasibility and utility of integrating the latter two surveys. The MFFS CAHPS survey is particularly useful for this purpose, as it is the only one of the three surveys that currently contains items that assess both experiences of care and health status (the latter in the form

[4] Program of All-Inclusive Care for the Elderly.
[5] This survey will be integrated into the HOS in 2005.

of the SF-12®[6]). To help shed light on some of the issues posed by attempting to integrate the HOS and M+C CAHPS surveys, data from the HOS for the years 1998 through 2002, and data from the M+C and MFFS CAHPS surveys for the years 2000, 2001, and 2002 were analyzed. The results of these preliminary analyses are presented and discussed in this report.

[6] SF-12® is a registered trademark of the Medical Outcomes Trust.

II. OVERVIEW OF EXISTING SURVEY INSTRUMENTS

The Medicare HOS

The goal of the HOS program is to gather valid and reliable health status data from Medicare managed care beneficiaries for use in QI activities, public reporting, plan accountability, and improvement of health outcomes. The core of this survey is the SF-36®, a standardized measure of self-reported health status with a long history of psychometric testing and development (Ware, Snow, Kosinski, and Gandek, 1993; Ware and Kosinski, 2001). The SF-36® has been used to estimate relative disease burden for over 200 diseases and conditions including arthritis, back pain, depression, diabetes and hypertension (National Committee for Quality Assurance [NCQA], 2003). The SF-36® produces two summary measures, the Physical Component Summary (PCS) score and the Mental Component Summary (MCS) score, and eight scale scores:

- Physical Functioning
- Role-Physical
- Bodily Pain
- General Health
- Vitality
- Social Functioning
- Role-Emotional
- Mental Health

In addition to the SF-36®, the HOS questionnaire includes questions regarding chronic medical conditions, negative symptoms, ADLs, and some demographic items (e.g., age and household income) that are useful for the case mix adjustment of plan-to-plan comparisons. Case mix adjustment models have been created for PCS change scores, MCS change scores, and death.

The HOS questionnaire is administered by mail, followed by telephone interviews of beneficiaries who do not respond to the mail questionnaires. The HOS is a longitudinal survey; after the initial baseline survey, an attempt is made to survey the original respondents to the baseline survey again two years later. The response rate to the baseline survey is approximately 65%, and the response rate to the follow up survey is approximately 75%.

The HOS questionnaire was developed by CMS, Health Assessment Lab (HAL), NCQA, and a consortium of performance measurement experts. The first round of data collection commenced in 1998 and the survey has been conducted annually since then. Each year a baseline cohort is surveyed, and two years later these same respondents are surveyed again (if they are still enrolled in the same health plan as at baseline). Reports of the results for the baseline survey are disseminated to the plans and the QIOs in the spring following the year in which the sampling occurred. Reports of the results for the follow

up survey, which include change scores based on the difference between the baseline and follow up results, are disseminated to the plans and the QIOs in the summer following the year in which the sampling occurred.

A key indicator of the utility of the HOS is its ability to detect differences between plans. An evaluation of MCS plan performance outliers has shown that the amount of change from baseline to follow up varies significantly from plan to plan (HSAG, 2003). This same study provides evidence of the instrument's validity: the MCS score was found to be strongly related to several of the HEDIS measures, as well as to some of the CAHPS composite scores.

The CAHPS Surveys

The purpose of the CAHPS surveys is to provide a standardized system for the measurement and reporting of health plan enrollees' experiences with the care they receive. In 1995, the Agency for Healthcare Research and Quality (AHRQ; formerly AHCPR) funded the development of the original CAHPS survey by a consortium of researchers at Harvard Medical School, the Research Triangle Institute (RTI), and RAND. In 1997, CMS began collecting CAHPS survey data from enrollees of M+COs. In 2000, CMS initiated the MFFS CAHPS survey to collect information on the experiences of enrollees in the original Medicare program. Three *Medicaid* versions of the CAHPS survey, for adults, children, and Children with Chronic Conditions (CCC), are also available. CAHPS scores are included in the Health Plan Employer Data and Information Set (HEDIS®[7]) requirements for NCQA accreditation of managed care plans.

The Medicare and Adult Medicaid CAHPS survey questions produce nine scores. These include four global rating questions (e.g., how respondents rate health care received from their health plan) and five composite measures. The composite measures are sets of questions grouped together to address a single aspect of care (e.g., getting needed care or getting care quickly). Currently, the MFFS CAHPS survey also contains the SF-12® health status measure, a shortened version of the SF-36®.

The CAHPS questionnaires are administered by mail, followed by telephone interviews of beneficiaries who do not respond to the mail questionnaires. The CAHPS surveys are cross-sectional surveys; no attempt is made to resurvey specific beneficiaries at a later point in time. Response rates to the CAHPS questionnaires are approximately 70% for the MFFS survey, 80% for the M+C survey, and 30% to 40% for the Medicaid surveys.

For the M+C CAHPS survey, the plan comparisons are case mix adjusted for general health perception, mental health perception, age, educational level, and type of respondent (self versus proxy). For the MFFS CAHPS, geographic comparisons are case mix adjusted for these same variables. The Medicaid CAHPS surveys are generally case mix adjusted for general health perception, age, and educational level.

[7] HEDIS® is a registered trademark of the National Committee for Quality Assurance.

The M+C CAHPS survey results are publicly reported, and the QIOs, CMS Central Office, and CMS Regional Offices all receive reports of the results. Since 2000, results from the MFFS CAHPS survey have been posted on the Medicare Web site for consumers to view. Data from both surveys have proven useful for comparing the results across M+C and FFS beneficiaries.

Since the fall of 2000, CMS has implemented two separate surveys of beneficiaries who voluntarily left their M+COs. These surveys are designed to gather information about these disenrollees' experiences with their plan and their reasons for leaving. The Medicare CAHPS Disenrollment Assessment Survey is conducted annually, and the results are combined with those from the M+C CAHPS survey for reporting to the public and the plans. Reporting the information in this way is thought to provide a more accurate description of all Medicare beneficiaries' experiences with their M+COs. The Medicare CAHPS Disenrollment Reasons Survey is designed to collect data about the reasons why Medicare beneficiaries leave their M+COs. Sampling and data collection occur quarterly, with analysis and reporting to plans and beneficiaries on an annual basis. The rate of response to each of these surveys is approximately 60%.

The ability of the M+C CAHPS instrument to detect differences between plans has been well documented (Zaslavsky, Zaborski, and Cleary, 2003). Evidence of the instrument's validity comes from a study by Schneider et al. (2001) that found strong relationships between HEDIS measures and CAHPS composite scores.

The Medicare Health Survey for PACE and Evercare (MHSPE)

PACE health plans are organizations that provide enhanced services to individuals who are 55 or older, certified by their state as eligible for nursing home care, able to live safely in the community, and live in a PACE geographical catchment area. The PACE program has the overall goal of keeping beneficiaries in their homes and communities. The Evercare program offers intensive primary care services to long-stay nursing home patients, with the goal of reducing hospitalizations (Kane, Keckhafer, and Robst, 2002).

The Medicare Health Survey for PACE and Evercare (MHSPE) is a brief survey instrument comprised of the SF-12®, a series of questions regarding ADLs, and questions about the use of a proxy respondent, and the types of help provided by the proxy if one was used. The PACE population is older and has significantly worse physical and mental health than the population enrolled in regular M+C plans (Grace, Fowler, Li, and Shannon, 2003). The purpose of the MHSPE survey is to provide a measure of beneficiary frailty that can be used to adjust payment rates for PACE and Evercare plans (Walsh, Nason, Moore, and Khatutsky, 2003).

The MHSPE is administered by mail, followed by telephone interviews of beneficiaries who do not respond to the mail questionnaires. The response rate for the first year of administration (2003) was close to 80%. In 2005, the MHSPE survey will be integrated into the HOS.

The Survey of Health Experiences of Veterans (SHEP)

Each month, the VA surveys approximately 12,000 enrollees who have used the VA's inpatient services, and also surveys approximately 36,000 enrollees who have used the VA's outpatient services. The instrument used is the SHEP. This instrument contains the VA version of the SF-12 (Veterans SF-12), and questions regarding patient satisfaction, health care utilization, health behaviors such as smoking and alcohol consumption, insurance coverage, usual sources of care, and complaint resolution. The inpatient version of the questionnaire contains 99 questions, and the outpatient version of the questionnaire contains 102 questions.

The SHEP is fielded by mail only, and produces a 60% to 70% response rate. There is no longitudinal follow up of enrollees at a later point in time. The survey is fielded in English and Spanish. The results are case mix adjusted for self-reported health status and age, and the results are also weighted to adjust for non-response biases. Reports of the results are posted on the VA's Web site. The results are also provided to Congress. A small portion of the facility and network administrators' salaries is based on the results for their respective areas of responsibility.

III. QUESTIONNAIRE CONTENT

Overview of Current HOS and CAHPS Survey Content

The current HOS instrument contains 99 items, which fall into the following main categories: the SF-36® health survey, questions about chronic medical conditions, and questions about the presence and severity of various symptoms.

The current M+C CAHPS survey contains 95 items, and the current MFFS CAHPS survey contains 92 items. Four global rating questions capture the respondents' overall satisfaction with their care:

- Rating of all health care
- Rating of health plan
- Rating of personal doctor
- Rating of specialist seen most often

Specific questions about key areas of the health care experience are summarized into six composite measures:

- Claims processing (MFFS CAHPS survey only)
- Courteous and helpful office staff
- Customer service
- Getting care quickly
- Getting needed care
- How well doctors communicate

In addition, the M+C CAHPS survey contains questions on appeals and grievances, as well as several indicators useful for the calculation of HEDIS indicators (advising smokers to quit, flu immunization, and pneumonia immunization).

All three surveys contain questions regarding demographic characteristics such as age, gender, racial background, and marital status. For each of the surveys, these items form part of the case mix adjustment process.

Comparison of Current HOS and CAHPS Survey Content

Despite the differences in the goals and historical development of these three surveys, some of their content is similar. Attachment C identifies the similarities among the current HOS, M+C CAHPS, and MFFS CAHPS surveys. The HOS and M+C CAHPS surveys share a total of four items with identical wording and format, and five items with similar wording. The HOS and MFFS CAHPS instruments share a total of ten items with identical wording, and seven items with similar wording.

Review of Key Issues

Survey Length

In the opinion of many of the stakeholders interviewed, the existing HOS and CAHPS surveys are quite lengthy. However, a study of Medicaid beneficiaries revealed that response rates were relatively insensitive to dramatic increases in CAHPS survey length (Gallagher and Fowler, 1998). A majority of stakeholders agree that approximately 90 items is the upper limit for a health care survey, and that adding further items will compromise response rates and the validity of the responses. There is some evidence that the reliability of the later items in a long survey is lower than that of the earlier items (Kazis, 2003). Therefore, to achieve an integrated survey of an appropriate length, a substantial number of items will need to be removed from the existing surveys. However, some items are currently shared between the two surveys, and this somewhat reduces the need to delete items.

Criteria for Determining Which Items to Retain

In deciding which items to include on an integrated survey, a number of criteria need to be considered.

Comparability Issues

For any item not retained in a new integrated survey, the ability to compare results from this item to results from its counterpart in the existing versions of the HOS or CAHPS surveys will be lost. Even for items that are retained, differences in format or order may compromise the ability to compare results from the current survey to results from an integrated survey.

One way to ensure valid comparisons between an item on one of the current questionnaires and its counterpart on an integrated questionnaire is to compare the two items in a randomized experiment, where one group of beneficiaries is randomly selected to receive the current version and another group is randomly selected to receive the integrated version. If differences are revealed by this experiment, then these differences can be used to adjust the new item scores for comparison to the scores for the current items.

A related issue involves comparability of the integrated survey to items from other surveys. For example, the "Healthy Days" questions that were added to the HOS in 2003 are also found in the Behavioral Risk Factor Surveillance System (BRFSS) questionnaire (Centers for Disease Control and Prevention, 2003). This makes possible direct comparisons between the Medicare managed care population and the population at large. In deciding which items to retain, consistency with other surveys also needs to be considered.

Certification Requirements

The HOS is currently not part of the NCQA accreditation process. In 1998, NCQA incorporated CAHPS 2.0H into the HEDIS reporting requirements. Currently, CAHPS performance accounts for 12.5% of a health plan's accreditation score. For example, the pneumonia and influenza immunization questions are required for HEDIS reporting. Deleting any items that are required for current HEDIS reporting would require a modification of current NCQA HEDIS accreditation requirements.

Prevalence in the Survey Sample

Items that involve only a small subset of the respondents, such as the item concerning the loss of the ability to talk on the CAHPS surveys, often do not yield enough observations per plan to be useful for plan-level analyses. Ware, Gandek, and Kosinski (2002) found that a number of the items on the HOS also suffered from a lack of observations at the plan level, including:

- Paralysis/weakness (1 item)
- Loss of ability to talk (1 item)
- Smoking questions (4 items)
- Provision of medical services in a retirement community (2 items)
- Chest pain (2 items)
- Shortness of breath (4 items)
- Problems with feet (4 items)
- Arthritis pain (1 item)

They recommended that these items be considered for elimination from the HOS. Since then, three of the smoking items and the two items concerning medical services in a retirement community have been dropped.

Current Performance of Survey Items

Items that have not performed well on the existing HOS and CAHPS surveys are candidates for omission from an integrated survey. For example, on the MFFS CAHPS survey, the overall rating of the Medicare program is difficult to interpret, because many of the respondents do not think of Medicare as a "health plan" that is comparable to a private sector health plan. Also, some respondents who are also enrolled in Medicaid or a Medigap plan find it difficult to differentiate between the Medicare FFS program and these other programs.

The CAHPS questionnaires contain a number of "gate" questions (questions which, when answered in a certain way, direct the respondent to skip a question or group of questions). For example, "Do you have one person you think of as your doctor or nurse?" is a gate question. Respondents who answer "Yes" are directed to a series of questions regarding their doctor or nurse. Respondents who answer "No" are directed to skip this series of questions.

If some of these gate items are almost always answered in a way that does *not* require a skip to later questions (usually the answer "Yes"), then they would be potential candidates for omission. An analysis of data from the M+C and MFFS CAHPS surveys (see Table 1) provides an indication of which items least often require a skip pattern.

Eight items, indicated in bold in Table 1, are least likely to generate a skip instruction. Based on this criterion, these items would be potential candidates for elimination from the survey.

TABLE 1
PERCENTAGE OF RESPONSES *NOT* REQUIRING A SKIP PATTERN FOR CAHPS "GATE QUESTIONS"

Survey Item	*M+C CAHPS Survey*	*MFFS CAHPS Survey*
A personal doctor or nurse is… When you joined Medicare or any time since then, did you get a new doctor or nurse?	N/A	43%
Do you have one person you think of as your personal doctor or nurse?	**92%**	**90%**
Do you have a physical or medical condition that seriously interferes with your ability to work or manage your day-to-day activities?	30%	38%
Did you have the same personal doctor or nurse before you joined your Medicare health plan?	44%	N/A
In the last 6 months, did you or your doctor think you needed to see a specialist?	54%	50%
In the last 6 months, how many times did you go to specialists for care for yourself?	**56%**	**76%**
In the last 6 months, did you call a doctor's office or clinic during regular office hours to get advice or help for yourself?	50%	54%
In the last 6 months, did you make any appointments with a doctor or other health provider for regular or routine health care?	**69%**	**72%**
In the last 6 months, did you have an illness or injury that needed care right away from a doctor's office, clinic, or emergency room?	29%	33%

TABLE 1
PERCENTAGE OF RESPONSES *NOT* REQUIRING A SKIP PATTERN FOR CAHPS "GATE QUESTIONS" (CONTINUED)

Survey Item	*M+C CAHPS Survey*	*MFFS CAHPS Survey*
In the last 6 months, did you or a doctor believe you needed any care, tests or treatment?	**72%**	**80%**
In the last 6 months, did you need approval from [HEALTH PLAN NAME/Medicare] for any care, tests, or treatment?	37%	11%
In the last 6 months, how many times did you go to a doctor's office or clinic to get care for yourself?	**78%**	**80%**
In the last 6 months, did you have a health problem for which you needed special medical equipment such as…?	13%	17%
In the last 6 months, did you have any health problems that needed special therapy, such as physical, occupational, or speech therapy?	11%	13%
In the last 6 months, did you need someone to come into your home to give you home health care or assistance?	5%	8%
In the last 6 months, did you need any new prescription medicines or need to refill a prescription?	**81%**	N/A
In the last 6 months, did you look for any information in written materials from Medicare?	26%	18%
In the last 6 months, did you call Medicare customer service to get information or help?	29%	9%
In the last 6 months, have you called or written your Medicare health plan with a complaint or a problem?	10%	N/A
In the last 6 months, did you have any experiences with paperwork for Medicare?	10%	11%
Was there ever a time when you strongly believed that you needed and should have received health care…doctor decided not to give you?	7%	N/A

TABLE 1
PERCENTAGE OF RESPONSES *NOT* REQUIRING A SKIP PATTERN FOR CAHPS "GATE QUESTIONS" (CONTINUED)

Survey Item	*M+C CAHPS Survey*	*MFFS CAHPS Survey*
Was there ever a time when you strongly believed that you needed and should have received health care…[PLAN] decided not to give you?	8%	N/A
Do you now have any physical or medical conditions that have lasted for at least three months?	**66%**	**72%**
Do you have a physical or medical condition that seriously interferes with your independence, participation…quality of life?	20%	32%
Did you get a flu shot last year, at any time from September to December [year]?	**72%**	N/A
Have you ever smoked at least 100 cigarettes in your life?	53%	53%
Do you now smoke every day, some days, or not at all?	17%	23%
How long has it been since you quit smoking cigarettes?	2%	2%
Did someone help you complete this survey?	13%	19%

Source: 2000, 2001, and 2002 MFFS and M+C CAHPS Surveys
Note 1: Some items were not present on the questionnaires during all three years.
Note 2: Due to variations in the handling of skip patterns for telephone interviews, the above figures are approximate.
Note 3: The items least likely to generate a skip instruction are indicated in bold type.
Note 4: The flu shot item is not a gate question in the MFFS version of the survey.

Table 1 suggests that eight gate questions are potential candidates for elimination. However, it may be possible to eliminate many more of the gate questions without losing information. This is because, in most cases, the groups of questions that follow the great majority of gate questions contain a redundant response alternative. Here is an example from the 2003 M+C CAHPS survey:

13. In the last six months, did you or a doctor think you needed to see a specialist?

☐ Yes
☐ No → **If no, go to Question 15**

14. In the last 6 months, how much of a problem, if any, was it to see a specialist that you needed to see?

☐ **A big problem**
☐ **A small problem**
☐ **Not a problem**
☐ **I didn't need to see a specialist in the last 6 months.**

If the last response alternative to Question 14 is checked, this provides the same information as a "No" response to Question 13. This suggests that Question 13 can be eliminated without loss of information. Eliminating such redundancies could greatly reduce the length of the M+C CAHPS questionnaire. Before eliminating any of the gate questions, two points of caution should be considered. First, the responses to the gate questions have been useful in validating the responses to the item blocks that follow the gate questions. Therefore, eliminating the gate questions may reduce the accuracy of the survey results. Secondly, some of the gate questions are substantive in nature, such as "Did you get a flu shot last year at any time from September to December?", and "Do you now have any physical or medical conditions that have lasted for at least three months?" Eliminating these particular questions will result in a loss of information.

The HOS questionnaire contains 13 items that ask the respondents about chronic medical conditions that they have experienced (for example, "Has a doctor ever told you that you have diabetes, high blood sugar, or sugar in the urine?"). The value of theses items has been questioned, based on the fact that a "Yes" response to any of these items does not distinguish among a wide range of severity levels, from mild to extremely disabling. However, other stakeholders, representing plans and QIOs, report that these questions provide a useful "snapshot" of the distribution of illnesses in their beneficiary populations, and find these data much easier to work with than the plans' encounter data. But, do self-reports of chronic medical conditions accurately reflect actual diagnoses found in encounter data? Kwon et al. (2003) found that self-reports of antidepressant use agreed with claims data in 85% of 422 cases examined. Miller et al. (2004) compared self-reports of 12 chronic medical conditions in the HOS to diagnostic codes from encounter data for the same beneficiaries in a VA population. A total of 17,089 beneficiaries completed the HOS and had encounters with the VA health care system in the two years prior to the survey. Sensitivities (the probabilities that the patient would report the disease given its indication in the medical record) were moderate to good (65% to 85%) for most of the conditions. The exceptions were chronic low back pain, congestive heart failure, and lung and colon cancer.

The existence of both baseline and follow up data on the same HOS respondents provides an opportunity for an additional test of the validity of self-reports of chronic medical conditions. The preamble to the list of questions regarding chronic conditions reads as follows: "Has a doctor ever told you that you had:". Therefore, logically, respondents who report a chronic medical condition on the baseline survey should also report the condition on the follow up survey two years later. Table 2 shows that, for each condition, a notable proportion of those who reported the condition at baseline did *not* report the condition at follow up.

TABLE 2
PERCENTAGE OF ILLOGICAL REPORTS FOR 13 CHRONIC MEDICAL CONDITIONS

Condition	***Total Reporting Condition at Baseline***	***Percentage of Total at Baseline Not Reporting Condition at Follow Up***
A chronic gastrointestinal problem	6,882	47%
Sciatica	31,382	34%
Other heart problems	27,945	27%
Arthritis of the hand	46,355	25%
Congestive heart failure	7,061	23%
Angina/coronary artery disease	20,195	21%
Arthritis of the hip	52,041	19%
Chronic obstructive pulmonary disease	16,947	19%
Acute myocardial infarction	12,894	18%
Stroke	7,726	17%
Any cancer (other than skin cancer)	18,050	12%
Diabetes	20,632	11%
High blood pressure	72,938	7%

Source: Medicare Health Outcomes Survey, *Cohorts I, II* and *III Merged Baseline* and *Follow Up*

The percentage of illogical reports was particularly high for chronic gastrointestinal illness, sciatica, other heart problems, and arthritis of the hand. These results suggest that these particular items are potential candidates for elimination from an integrated survey.

Currently the HOS contains items that reflect the severity of five chronic medical conditions: heart disease (Items 16a and 16b), pulmonary disease (Items 17a through 17d), diabetes (Items 18a through 18d), stroke (Items 19a and 19b), and arthritis (Item 36). These items can prove helpful in distinguishing between mild and severe cases of these five conditions.

Legal Requirements

The current M+C CAHPS contains 10 items regarding the appeals and grievances process. The inclusion of these items was motivated by the ruling of the Ninth Circuit Court in the Grijalva versus Shalala case (1998). In its ruling, the Ninth Circuit Court had decided that Medicare managed care providers must extend due process protections to Medicare beneficiaries when their benefits are denied, reduced, or terminated (Lock, 1999). The questions regarding the appeals and grievances process provide Medicare program decision makers with information on the extent to which the program's beneficiaries understand and/or have accessed these protections. This information may be useful in documenting the Medicare program's responsiveness to the Grijalva versus Shalala lawsuit. However, the court's ruling does not mandate the use of any specific survey questions. Therefore, at least some of the items in the Appeals and Complaints section of the M+C CAHPS may be considered for elimination.

The recent Medicare Prescription Drug, Improvement, and Modernization Act of 2003 (S. Res. 1, 2003) states that "The Secretary shall not collect under subparagraph (A) data on quality, outcomes, and beneficiary satisfaction to facilitate consumer choice and program administration other than the types of data that were collected by the Secretary as of November 1, 2003" (Section 722(a)(3)(B)(i), page 283.) While this provision appears to bar the development of methods for collecting entirely new types of data, it does not appear to impede the modification of instruments for collecting *current* types of data.

Predictive Power

Certain items are useful to retain on the basis of their relationship to other items. For example, the item on self-reported mental health status has proven very useful as a case mix adjustment variable on the M+C CAHPS. Zaslavsky et al. (2001) found that the ADL items added little predictive power to the case mix adjustment model for the M+C CAHPS. Among the chronic medical conditions items on the HOS, pulmonary diseases, arthritis of the hip or knee, and arthritis of the hand or wrist are most predictive of two-year declines in PCS scores (Ellis et al., 2004). The question on beneficiary health compared to one's peers, currently found on the HOS survey, has proven to be a powerful predictor of subsequent health care utilization (Bierman, Bubolz, Fisher, and Wasson, 1999).

Availability from Other Data Sources

Demographic information, such as age, gender, and race, is available from CMS' Medicare Enrollment Data Base (EDB). This means that the questions capturing this information can potentially be dropped from the questionnaire. However, not all of the demographic information of interest is available from the EDB (e.g., educational level), and the accuracy of the data that are available has been questioned. Also, it has proven useful to collect demographic information from the survey for validation purposes.

In addition, CMS is now collecting some diagnostic information (both inpatient and outpatient) from M+COs. As a result, the questions on chronic medical conditions could potentially be eliminated from the survey. However, linking the survey information to this database would add a significant level of complexity to the data collection process.

Payment Adjustment

Items that may prove useful for the determination of payments, such as the ADL items, should be strongly considered for inclusion in an integrated survey. At the present time, the ADL items are used to adjust PACE plan reimbursements to reflect the frailty of the plans' members. An ADL-based frailty adjustment mechanism is also currently being tested for the M+C beneficiaries. A sample of about 50,000 MFFS beneficiaries will be surveyed. MFFS respondents, rather than M+C respondents, are being used because the M+C payments have been calibrated on data from MFFS beneficiaries (M+C beneficiaries do not have claims data that can be used for calibration).

Usefulness for Reporting

Survey questions that are used for reports to health care decision makers and/or the public are a high priority for retention in an integrated survey. For example, the two-year PCS and MCS change scores from the HOS, after adjustment for non-performance related differences among plans, provide "bottom line" measures of each plan's impact on health outcomes. The current HOS Performance Measurement Reports focus primarily on these measures. Also of interest to plan administrators are the distributions of chronic medical conditions and ADLs among the beneficiary population. At the present time, HOS results are not being publicly reported.

The overall rating questions and the composites from the CAHPS surveys are viewed as useful measures of plan performance for both plan managers and consumers. Dropping any of the items contained in the CAHPS composites would result in the loss of this information.

Survey questions that are seldom or never used in reports of survey results should be considered for elimination from an integrated survey.

Usefulness for Quality Improvement (QI) Activities

A number of plans have used the HOS results to develop statistical profiles of beneficiaries at risk for depression. Also, the chronic medical conditions items from the HOS have been used by some plans and QIOs to determine the medical conditions that their disease management programs should focus on. The HOS results can reveal if beneficiaries with certain chronic medical conditions have particularly pronounced declines in PCS or MCS scores. Such subgroups would be promising targets for disease management.

Items that have proven useful in the past for QI activities are strong candidates for retention in an integrated survey. Plan managers have found CAHPS items related to overall plan operations, such as the customer service composite, to be useful in this regard.

The SF-12 as an Alternative to the SF-36

The SF-12 is a shorter version of the SF-36 designed to produce the same PCS and MCS scores as the longer instrument. QualityMetric/MOT, RAND, and the VA have all created their own versions of both the SF-12 and the SF-36[8]. For clarity, this report will refer to each specific version as shown in Table 3.

TABLE 3
VERSIONS OF THE SF-12 AND SF-36

Agency	*SF-12*	*SF-36*
QualityMetric/MOT	SF-12®, SF-12v2™[9]	SF-36®, SF-36v2™[9]
RAND	RAND SF-12	RAND SF-36
Veterans Administration	Veterans SF-12	Veterans SF-36

If the results of the SF-36 can be largely reproduced by the SF-12, then replacement of the SF-36 with the SF-12 would be feasible and would reduce the length of an integrated survey through the elimination of 24 items. As does the SF-36, the SF-12 yields PCS and MCS scores. Ware, Kosinski, Turner-Bowker, and Gandek (2002) demonstrated that the SF-12® reproduces the results of the SF-36® with little loss of precision. They found that the correlations between the SF-12® and SF-36® versions of the PCS and MSC scores were .951 and .969, respectively. Equally important, the SF-36® contains all of the items in the SF-12®, suggesting that the SF-12® portion of the SF-36® in the current HOS could be trended with SF-12® results from a future integrated survey. Scores on the eight individual scales are also obtainable from the SF-12v2™ survey (Ware, Kosinski, Turner-Bowker, and Gandek, 2002).

[8] In this report, "SF-12" and "SF-36" will be used to refer to these three versions collectively.
[9] SF-12v2™ and SF-36v2™ are trademarks of QualityMetric Incorporated.

Choice of SF-12 Instrument

If the SF-36 is replaced by the SF-12, the four versions of this instrument shown in Table 3 would be available to choose from. These four versions are very similar in content, and differ primarily in the systems used to score the results. A key difference between the SF-12® and the Veterans SF-12 is that the Veterans version uses five-point ordinal ratings for the Role-Physical (RP) and Role-Emotional (RE) subscales, whereas the SF-12® uses dichotomous responses. However, the SF-12v2™ also uses five point ordinal ratings for the RP and RE subscales. The SF-12v2™ is currently in wider use than either the Veterans or the RAND SF-12, and has been adopted for inclusion in the Medical Expenditure Panel Survey or MEPS (SF Community Web Site, 2003). The results from a controlled experiment by Kazis, et al. (2004) make it possible to rescore the results of the Veterans versions so that they are in the same metric as the QualityMetric versions. This facilitates comparisons between results from the Veterans and QualityMetric versions.

All four versions of the SF-12 are currently available for use. According to QualityMetric's Web site, The SF-12v2™ instrument is "available royalty free to individuals and organizations for un-funded scholarly research" (SF Community Web Site, 2004). It is not clear from the Web site whether or not the scoring system is also royalty free to non-commercial users. The RAND and Veterans versions of the SF-12, along with their scoring systems, are in the public domain.

Formatting Issues

Format of Questionnaires

The existing HOS and CAHPS surveys are formatted differently. The HOS is formatted in a single column layout, whereas the CAHPS surveys are formatted in a double column layout.

On the HOS survey, groups of similar items are introduced by an introductory (or "stem") question that is not repeated for each question in a group. In the CAHPS surveys, the stem questions are repeated for each question in a group of questions. Repetition of the stem questions was based on findings from the literature (Harris-Kojetin, et al., 1999), as well as cognitive testing of the CAHPS items.

If the two surveys are to be successfully integrated, it will be necessary to decide whether each survey's current format should be retained, or a uniform format adopted. A decision to reformat a specific item could lead to a loss of comparability to previous data collected using the item's original format. On the other hand, an integrated survey that retains the differences in formatting styles may be confusing to the respondent.

Context Effects and Order Effects

Whenever the content of a questionnaire is altered, users of the questionnaire must be alert to the possible impact of these changes on the results obtained. Two major types of impact are possible. First, the results from separate health status and experience of care measures may differ from the results that would be obtained from an integrated instrument. This is an example of a *context effect*. Second, it is possible that results from functional status and experience of care measures will vary depending upon the order in which these two types of items appear in the questionnaire. This is an example of an *order effect*. No controlled studies of context and order effects for health status and satisfaction measures were found in the literature. Several of the stakeholders have suggested that the more subjective (i.e., experience of care) items be placed before the health status items so that responding to the health status items will not influence the satisfaction responses. This is, in fact, the ordering used in the current MFFS CAHPS questionnaire. One study suggests that the rating of overall health status is lower if this rating follows the SF-12®, compared to when it precedes the SF-12® (Shimada, Cioffi, Zaslavsky, and Cleary, 2003).

To the extent that context effects and/or order effects exist, this may complicate attempts to trend items across the current HOS and CAHPS surveys to their counterparts on an integrated instrument. Comparability of the integrated survey items to the items on the MFFS CAHPS questionnaire may also be affected. Even if an item is retained in the integrated survey with the exact wording and format as before, order effects and context effects may undermine the validity of the comparison. As discussed previously, one way to ensure valid comparisons between an item on one of the existing questionnaires and its counterpart on an integrated questionnaire is to compare responses to the two items in a randomized experiment. That is, a group of respondents would be randomly assigned to receive either the existing questionnaire or the integrated questionnaire, and their responses to the items at issue would be compared. The results of this experiment could then be used to "calibrate" comparisons between the two items.

Results from the current HOS, MFFS CAHPS, and M+C CAHPS surveys can help to determine the importance of context and order effects. If identical items from these three surveys are answered very similarly (after appropriate case mix adjustment of the respondent populations), then this would be a strong indication that the responses to these items are fairly insensitive to context and order effects. If, on the other hand, these items are answered quite differently from survey to survey, this would indicate the importance of calibrating item comparisons by means of controlled experiments. Table 4 shows the results of comparisons for six questions that are worded exactly the same in both the HOS and MFFS CAHPS surveys. To enhance the validity of these comparisons, the respondents to the two surveys were matched for age, education, general health perception, and proxy status. Only respondents that could be exactly matched (on a 1:1 basis between the two groups of respondents) were retained for this analysis. It is important to note that, even after matching the respondents as described above, significant differences may remain between the M+C and MFFS populations being compared.

TABLE 4
COMPARISON OF RESPONSES TO IDENTICAL ITEMS: HOS AND MFFS CAHPS SURVEYS

Does your health now limit you in climbing several flights of stairs?

Response Category	HOS	MFFS CAHPS
Yes, limited a lot	33%	33%
Yes, limited a little	35%	32%
No, not limited at all	32%	35%
N	*87,780*	*87,780*
Effect Size	*.03*	

During the past four weeks, how much did pain interfere with your normal work…?

Response Category	HOS	MFFS CAHPS
Not at all	37%	30%
A little bit	24%	24%
Moderately	20%	20%
Quite a bit	15%	18%
Extremely	4%	7%
N	*87,686*	*87,686*
Effect Size	*.09*	

How much of the time, during the past 4 weeks, have you felt calm and peaceful?

Response Category	HOS	MFFS CAHPS
All of the time	12%	11%
Most of the time	40%	43%
A good bit of the time	15%	17%
Some of the time	18%	19%
A little of the time	10%	8%
None of the time	5%	2%
N	*87,731*	*87,731*
Effect Size	*.09*	

How much of the time, during the past 4 weeks, did you have a lot of energy?

Response Category	HOS	MFFS CAHPS
All of the time	5%	4%
Most of the time	23%	23%
A good bit of the time	17%	17%
Some of the time	24%	28%
A little of the time	17%	17%
None of the time	14%	11%
N	*87,844*	*87,844*
Effect Size	*.07*	

TABLE 4
COMPARISON OF RESPONSES TO IDENTICAL ITEMS:
HOS AND MFFS CAHPS SURVEYS (CONTINUED)

How much of the time, during the past 4 weeks, have you felt downhearted and blue?

Response Category	HOS	MFFS CAHPS
All of the time	1%	2%
Most of the time	3%	4%
A good bit of the time	4%	6%
Some of the time	11%	22%
A little of the time	16%	32%
None of the time	65%	35%
N	*87,811*	*87,811*
Effect Size	*.30*	

During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities…?

Response Category	HOS	MFFS CAHPS
All of the time	4%	5%
Most of the time	11%	8%
Some of the time	15%	18%
A little of the time	18%	15%
None of the time	52%	53%
N	*87,978*	*87,978*
Effect Size	*.08*	

Source: HOS *Cohort V Baseline*, 2002 and MFFS CAHPS 2002
Note: Observations for which a survey item had missing data were dropped from the analysis.

For the six items compared, the frequency distributions were quite similar across the two surveys, with effect sizes well below the conventional level of .20 for a small effect size. The one exception was the question regarding feeling "downhearted and blue." This result suggests that certain items may be sensitive to order and context effects, and that therefore, comparisons of the items in an integrated survey to their counterparts in the current surveys may be difficult to interpret.

The items compared above are all from the SF-36® and SF-12® instruments embedded in the current HOS and MFFS CAHPS questionnaires, respectively. Since the SF-12® items are present in both questionnaires, the SF-12®'s performance in the two surveys can be compared from two standpoints: number of missing responses and overall reliability. Once again, to enhance the validity of these comparisons, the respondents to the two surveys were matched for age, education, and proxy status. (Because general health perception is itself one of the SF-12® items, it was not used as a matching variable.) Table 5 compares each item on the SF-12® in terms of the percentage of missing responses.

TABLE 5
PERCENTAGE OF MISSING RESPONSES TO SF-12® ITEMS: HOS AND MFFS CAHPS SURVEYS

Item	*HOS*	*MFFS CAHPS*
General health rating	1.0%	1.3%
Health limitations on moderate activities	1.5%	1.5%
Health limitations on climbing several flights of stairs	1.8%	1.7%
Accomplished less than would like due to physical health	2.8%	2.1%
Limited in work or other activities due to physical health	3.9%	2.0%
Accomplished less than would like due to emotional problems	3.0%	2.0%
Didn't do work as carefully due to emotional problems	3.8%	2.4%
Interference of pain with normal work	1.5%	1.8%
Felt calm and peaceful during last 4 weeks	2.0%	1.4%
Had a lot of energy during last 4 weeks	1.9%	1.3%
Felt downhearted and blue during last 4 weeks	1.8%	1.7%
Physical or emotional health interfered with social activities	1.8%	1.2%

Source: HOS *Cohort V Baseline*, 2002 and MFFS CAHPS 2002

The rates of missing responses are uniformly low and very similar across the two surveys. This would suggest that placing the SF-12® instrument near the end of the survey does not increase the amount of missing responses. If anything, the missing response rates are slightly higher for the HOS questionnaire, in which the SF-36® items are positioned near the beginning of the survey.

Internal consistency reliabilities, as measured by Cronbach's alpha, were computed for the HOS respondents and MFFS CAHPS respondents, matched once again for age, education, and proxy status. The SF-12® reliabilities were very high and nearly equal for the two surveys: 0.91 for the HOS and 0.92 for the MFFS CAHPS. The placement of the SF-12® items near the end of the MFFS CAHPS survey does not appear to affect their performance.

Supplementary Items

Periodically, CMS adds specific question sets to the HOS questionnaire to address issues of current interest. Recently, question sets regarding urinary incontinence and the number of days of good health experienced during the last 30 days have been added to the HOS. An earlier question set regarding smoking behaviors was added in 1998; this was removed in 2003. When adopted, these question sets are included as standard for all of the HOS questionnaires administered.

To help plans more effectively use the CAHPS results, a *modular approach* to assessing the quality of care delivered by health plans is under development by AHRQ. The proposed Ambulatory CAHPS (A-CAHPS; CAHPS Survey Users Network, 2004b) would assess the quality of ambulatory care at different levels of the health care system while still retaining many features of the current CAHPS health plan surveys.

With the A-CAHPS, users would be able to assess beneficiary perceptions of care at one or more of the following levels:

- Health plans, including FFS plans, Preferred Provider Organizations (PPOs), or managed care plans
- Group practices
- Sites of care
- Individual clinicians

The above levels may exist within the same organization (for example, in a staff model HMO), or they may be fairly separate entities. In the modular approach, each plan would utilize a core set of questionnaire items, but could also choose to add optional sets of questions, or modules, to the questionnaire to assess experiences of care at a specific level. The results from the modules would provide plans with additional information to guide QI efforts.

The A-CAHPS questionnaire for group practices (known as the G-CAHPS) was released in March 2004. Users will be able to obtain benchmark comparisons for their results (CAHPS® Survey Users Network, 2004c).

Users of the Medicaid CAHPS survey have the option of adding supplemental questions of their own design to the survey. These questions must be approved by NCQA before the survey is administered.

If the Medicare CAHPS survey questionnaires evolve toward a modular format, then integration of the HOS and CAHPS surveys will require integration with each of the modular versions available. The presence of one or more CAHPS modules on an integrated questionnaire will increase its overall length, which may further reduce the space remaining for other items.

Preliminary Recommendations

1. *From the HOS questionnaire, replace the SF-36® with the Veterans SF-12, and delete a number of other items*. Table 6 shows the survey items that appear to be the best candidates for deletion from the HOS, with corresponding rationales.

TABLE 6
PRELIMINARY RECOMMENDED DELETIONS FROM THE MEDICARE HEALTH OUTCOMES SURVEY

Item Numbers[1]	*Description*	*Rationale*
3a,c, e-j; 4a,d; 5a; 6; 7; 9a-c; 9g-i; 11a-d	SF-36® items not contained in the SF-12®	SF-36® scores are well-reproduced by the SF-12®
25; 27; 30; 31; 32; 33 (two new items created)	Reports of chronic medical conditions for: gastrointestinal problems, sciatica, other heart problems, congestive heart failure, arthritis of the hip, arthritis of the hand	Respondents often report the first four of these items in an illogical fashion; other heart problems and congestive heart failure can be merged into a single item concerning "heart disease"; arthritis of hip and hand can be combined into one item concerning "arthritis"
38	Low back pain	Not used in case mix adjustment nor in HOS reports
43	Current smoking frequency	Already included in the CAHPS survey; will be retained in integrated survey
48; 49; 50; 51; 53	Age, gender, Hispanic descent, race, and education questions	Already included in the CAHPS survey; will be retained in integrated survey
55	Proxy item	Already included in the CAHPS survey; will be retained in integrated survey

[1] 2004 version of the HOS

If the recommendations outlined in Table 6 were to be adopted, a total of 36 items would be eliminated from the HOS.

2. *From the M+C CAHPS questionnaire, delete most of the gate questions and most of the grievances and appeals items, as well as a number of other items.* Table 7 shows the survey items that appear to be the best candidates for deletion from the M+C CAHPS, with corresponding rationales.

TABLE 7
PRELIMINARY RECOMMENDED DELETIONS FROM THE M+C CAHPS SURVEY

Item Numbers[1]	*Description*	*Rationale*
9; 11; 13; 15; 18; 20; 22; 25; 26; 28; 44; 51; 53; 56; 59	Gate questions for: condition that interferes with work, have same doctor or nurse before joining plan, needed to see specialist, number of times specialist seen in last six months; called doctor's office or clinic for advice, had condition that needed care right away, made appointment with a provider for health care, needed care/tests/treatment, needed approval from health plan, needed prescription medicines, looked for written information, called customer service, reported complaint or problem, needed to fill out paperwork	If item does not apply to respondent, they can indicate this by checking the last response category of the following item; gate items regarding personal doctor or nurse and flu shot are recommended for retention
38-43	Special equipment, special therapy, and home health questions	Most respondents do not report using these services
62; 64; 65; 68-71	Seven of the ten appeals and grievances items	The remaining three items are sufficient to document plan's efforts to comply with the Grijalva ruling; this ruling does not dictate any particular survey content
51-60	Finding information about health plan, contacting customer service, getting help from customer service; filing and handling of complaints; paperwork	Most beneficiaries do not use these services
72; 74; 77-81; 85; 87	General health perception; inpatient stay; questions about treatment for a chronic medical condition and prescription medicines; help needed due to an impairment	The general health perception question is on the HOS and should be retained in an integrated survey; other questions can be used to assess chronicity

[1] 2003 version of the M+C CAHPS

If the recommendations outlined in Table 7 were to be adopted, a total of 47 items would be eliminated from the M+C CAHPS. However, it should be noted that the gate questions have played an important role in validating the responses to the survey. Elimination of these items would therefore result in the loss of some information. An alternative strategy would be to eliminate some of the items used to form the composite scores. Zaslavsky and Cleary (2002) found a substantial amount of redundancy in the traditional reporting composites.

3. *Organize the remaining items with the majority of the CAHPS items first, followed by the Veterans SF-12, the remaining HOS items, and the demographic items that are common to both surveys*. Table 8 shows the items that appear to be the best candidates for inclusion in an integrated survey instrument, and a suggested order in which they might appear.

TABLE 8
PRELIMINARY RECOMMENDED CONTENT FOR AN INTEGRATED SURVEY

Item Numbers[1,2]	***Description***	***Rationale***
CAHPS: 4-6; 8; 10; 12	Section on personal doctor or nurse, minus two gate questions	Needed for calculation of composite scores and for CAHPS reporting
CAHPS: 7	Global rating of personal doctor or nurse	Needed for CAHPS reporting
CAHPS: 14; 17	Section on getting health care from specialist, minus two gate questions	Needed for calculation of composite scores
CAHPS: 16	Global rating of specialist	Needed for CAHPS reporting
CAHPS: 19; 21; 23-25; 27; 29-36	Section on health care in the last six months, minus five gate questions	Needed for calculation of composite scores
CAHPS: 37	Global rating of health care in the past six months	Needed for CAHPS reporting
CAHPS: 45-50	Section on other health services, minus special equipment, special therapy, and home health questions, and one other gate question	Retain prescription items; increasing importance due to Medicare Modernization Act
CAHPS: 61	Global rating of health plan	Needed for CAHPS reporting
CAHPS: 63; 66; 67	Appeals and complaints section, items on ability to ask plan to reconsider a doctor's decision, ever speak to someone regarding reconsideration, and information received when requesting an appeal	These items are sufficient to document a plan's attempts to comply with the Grijalva decision
HOS: 1; 3b,d; 4b,c; 5b,c; 8; 9d-f; 10	The items from the SF-36® that correspond to the 12 items of the SF-12.	Principal measure of health outcomes; SF-12V2™ and VA SF-12 yield the eight scale scores, for the VA version, the scoring algorithm and missing data imputation methods are known to be in the public domain
HOS: 12a-f	Activities of Daily Living	Useful for frailty adjustment
HOS: 13-15	Healthy Days questions	Can compare to the same items in the BRFSS survey
HOS: 23; 28; 29; 34; 35; two new items	Chronic medical conditions items for "heart disease" (new category), "arthritis" (new category), chronic obstructive pulmonary disease (COPD), stroke, cancer, diabetes, and high blood pressure	Most of these items produce relatively few illogical reports; useful to plans in assessing their beneficiary populations; can be used to develop disease-specific plan comparisons; each question has an accompanying symptom severity question (see below)

[1] 2004 version of the HOS
[2] 2003 version of the M+C CAHPS

TABLE 8
PRELIMINARY RECOMMENDED CONTENT FOR AN INTEGRATED SURVEY (CONTINUED)

Item Numbers[1,2]	*Description*	*Rationale*
HOS: 16a,b; 17a-f; 18a-d; 19a,b; 36	Chest pain, shortness of breath, foot symptom, stroke symptoms, and arthritis pain questions	Provide severity information for the heart disease, shortness of breath, diabetes, and stroke chronic medical conditions items
HOS: 20-22	Vision and hearing problems, urinary incontinence	Will be used for validation of new urinary incontinence HEDIS measure
HOS: 37a-d	Cancer treatment questions	HOS is only large scale database on outcomes and health-related quality of life data for persons with cancer; has been useful to American Cancer Society researchers
HOS: 39-41	Depression screener	Used in HOS reports
HOS: 44-47	Urinary incontinence	Will be used for validation of new urinary incontinence HEDIS measure
CAHPS: 73; 75; 76; 82; 83; 84- 88	Health compared to one year ago, overall mental health perception, condition lasting more than three months, flu shot, pneumonia shot; ever smoked at least 100 cigarettes, current smoker, how long since quit smoking, advice to quit smoking	Mental health perception question useful for case mix adjustment; question on condition lasting more than three months is an indicator of chronicity; flu and pneumonia shot questions required for HEDIS, current smoker question is strongly related to depression; advice to quit smoking question allows plans to monitor an important intervention; all four smoking items needed to calculate the HEDIS measure.
CAHPS: 89-93	CAHPS items on age, gender, Hispanic descent, race, education, homeownership, and marital status questions	Useful for case mix adjustment and studies of disadvantaged populations
CAHPS: 94; 95	Proxy questions	Results for self- and proxy respondents show strong differences; useful in case mix adjustment

[1] 2004 version of the HOS
[2] 2003 version of the M+C CAHPS

Research by Shimada et al. (2003) appears to indicate that the SF-12 is less sensitive to changes in ordering than the CAHPS questions, so it seems advisable to place the CAHPS questions in the first part of the questionnaire.

The integrated survey described above would contain 106 items. If modular or supplementary question sets were to be added to the questionnaire, then the number of items would increase further. Therefore, fewer items would be available for deletion. One possible strategy for accommodating supplementary questions or modular question sets, while still maintaining an acceptable survey length, would be to rotate these items in and out of the survey on an alternating year basis.

IV. SURVEY ADMINISTRATIVE PROTOCOLS

Comparison of Current HOS and CAHPS Survey Administrative Protocols

Integrating the HOS and CAHPS surveys requires not only a merging of the survey content, but also an integration of the administrative protocols of the two surveys. The current HOS and CAHPS survey protocols differ in several significant ways.

Research Design

There is a fundamental difference between the research designs of the two surveys. The HOS is a *longitudinal* survey; after the initial baseline survey, the original respondents to the baseline survey are surveyed again two years later. In contrast with the HOS, the CAHPS surveys are *cross-sectional* surveys. Each year a new sample of current beneficiaries is taken, and these beneficiaries are not sent a follow up survey at a later point in time. Although some beneficiaries receive the CAHPS survey again in subsequent years, this is not an intentional feature of the research design. To field a longitudinal survey, it is necessary to resurvey prior cohorts of respondents while at the same time surveying new cohorts of individuals. This requires a more complex protocol than is needed for cross-sectional surveys.

Eligibility

Currently, the types of organizations required to participate in the HOS and M+C CAHPS survey vary somewhat. The HOS is administered to M+COs, Social Health Maintenance Organizations (SHMOs), Continuing Cost Contracts, Private Fee For Service (PFFS) plans, and Medicare Alternative Payment Demonstration Plans (NCQA, 2004). The M+C CAHPS survey is administered to members of the first three groups (CMS, 2003).

Data Collection Procedures

The HOS, M+C CAHPS, and MFFS CAHPS surveys use a mixed-mode methodology, consisting of mailed questionnaires and telephone interviews of beneficiaries who do not respond to the mailed questionnaire. Table 9 compares the current survey administrative procedures for these three surveys.

TABLE 9
COMPARISON OF HOS, M+C CAHPS, AND MFFS CAHPS DATA COLLECTION PROCEDURES

HOS Survey Procedure	***M+C CAHPS Survey Procedure***	***MFFS CAHPS Survey Procedure***
Prenotification postcard	Prenotification letter	Prenotification letter (with option to complete by telephone)
First survey mailing	First survey mailing	First survey mailing
First reminder/thank you postcard	First reminder/thank you postcard	First reminder/thank you postcard
Second survey mailing	Second survey mailing	Second survey mailing
Second reminder/thank you postcard		
Telephone follow up*	Telephone follow up	Telephone follow up
	Third survey mailing**	Third survey mailing**

* Telephone follow up is also done for partially completed mail surveys.
** This mailing is sent to non-respondents without telephone numbers, via priority mail.
Note: For each survey, Spanish versions are available and proxy responses are allowed.

As the above table makes clear, the current data collection procedures are fairly similar. The most significant differences are the number of survey mailings and the fact that for the HOS, partially completed mail surveys can be completed by a follow up survey mailing or by telephone interview, whereas for the M+C and MFFS CAHPS surveys, a given survey is completed entirely by mail or entirely by telephone.

The response rate to the HOS baseline survey conducted in 2002 was 64%, and the response rate to the HOS follow up survey conducted in 2002 was 76%. The 2002 response rate to the M+C CAHPS survey was 83%. These response rates meet the acceptable threshold for a "good" response rate as described by a majority of the stakeholders. The HOS baseline survey has yielded a lower response rate than the M+C CAHPS survey, but this may be due to the fact that the definition of a completed questionnaire for the M+C CAHPS survey (one or more of the items completed), is less stringent than the definition of a completed HOS questionnaire (the presence of enough

data to calculate both a PCS and an MCS score). A portion of the difference in response rates may also be due to the different content of the surveys.

Management of Vendors

Currently, each M+CO can choose from one of five vendors, certified by NCQA, to administer the HOS to the plan's beneficiaries. There is a single vendor for the M+C CAHPS, chosen by CMS.

Timing of Data Collection

Both the HOS and M+C CAHPS surveys are administered annually during a fixed time period. The HOS surveys are administered during April through July. The M+C CAHPS, on the other hand, is administered during the period from September through December.

Review of Key Issues

Research Design

Should an integrated survey be longitudinal or cross-sectional? A longitudinal survey requires a larger sample at baseline in order to obtain a reasonable sample size at follow up. Also, the need to survey the same respondents at a later time increases the complexity of the survey protocol.

A hybrid version of the two approaches is also possible. In this scenario, a yearly baseline survey would be administered, and the baseline sample would be followed up again in a later year. A supplementary group of beneficiaries who were not previously sampled would be added to each follow up sample to obtain a representative cross-sectional sample as well. Another possibility is to administer both the HOS and CAHPS items at baseline, but only the HOS items at follow up. Each of these scenarios has different implications for design of an administrative protocol for an integrated survey.

Data Collection Procedures

The data collection for an integrated survey could be greatly simplified if the telephone component of data collection were dropped. However, respondents who are older, poorer, less educated, have more health problems or who use a proxy to respond to the survey are more likely to respond by telephone (NCQA, 2002b). Consequently, it is important to retain the telephone component of the survey protocol.

Management of Vendors

Use of one vendor is more efficient, but multiple vendors introduce an element of competition and provide a frame of reference for the quality of vendor performance. One potential disadvantage of using multiple vendors is that the results may not be strictly comparable across plans. An analysis of 1998 HOS survey results from six certified vendors (Haffer, 2003) suggests that comparability is not compromised by the use of multiple vendors. Haffer found that overall rates of agreement with a "gold standard" survey result ranged from 97.0% to 99.8% across the six vendors. When multiple vendors are used, a comparison of the response rates, the demographic composition of respondents, and question response patterns across vendors can be very useful in identifying a vendor that is not following all required data collection procedures. When a single vendor is used, it is much harder to spot these types of problems without an intensive audit of every aspect of the vendor's data handling operations.

When multiple vendors are involved, any changes in the survey protocol need to be discussed and clarified with each vendor in the group, which is more inefficient. Each additional vendor that is involved in the process has the potential to increase the effort required for coordination of the process. However, the viewpoints expressed by different vendors may lead to fresh insights into how to improve survey administrative procedures. Finally, the use of multiple vendors reduces the chances that some vendors will feel "shut out" of a major health care survey initiative.

Timing of Data Collection

Because the two surveys are fielded at different times of the year, integrating the two surveys would require a consensus regarding the best time of the year for survey administration.

Other Issues

In addition to the above considerations, the design of an administrative protocol for an integrated survey also has significant implications for survey sampling procedures, survey costs, and beneficiary burden. These implications will be discussed later in this report.

Preliminary Recommendations

1. *Employ a longitudinal research design*. A research design that provides cross-sectional as well as longitudinal information for both the HOS and CAHPS items has a number of analytic advantages, which will be discussed in the section on Analytic Strategies.

2. *Use at least three, but no more than four, vendors to manage the data collection process*. This will encourage competition among the vendors, will provide CMS with the ability to compare vendor performance and identify data collection

issues, and at the same time will reduce the coordination problems and increased costs associated with managing a larger number of vendors.

3. *Collect data on vendor performance (data quality issues, response rates, etc.), and make this information available to the M+COs each year*. This will provide vendors with an additional incentive to do their best work.

4. *Collect data by both mail and telephone*. Although it increases the complexity of the administrative protocol, telephone data collection is crucial to obtaining adequate response rates from certain beneficiary subgroups.

V. ANALYTIC STRATEGIES

Overview of Current Analytic Strategies

Research Design

The HOS and CAHPS surveys use fundamentally different research designs. The HOS is a *longitudinal* survey; the same respondents are resurveyed two years after the initial survey. The availability of both baseline and follow up scores for the same individuals allows functional status change scores to be calculated for individual beneficiaries.

In contrast with the HOS, the MFFS, M+C, and Medicaid CAHPS surveys are *cross-sectional* surveys. Each year a new sample of current beneficiaries is taken, and this sample of beneficiaries is *not* sent a second survey at a later point in time. Although some beneficiaries may receive a CAHPS survey in multiple years (this is more likely for those plans with small beneficiary populations), this is not an intentional feature of the CAHPS survey research designs.

CMS also gathers information from disenrollees for the M+C CAHPS survey. CMS conducts two different surveys of disenrollees who voluntarily leave an M+C plan. The Medicare CAHPS Disenrollment Assessment Survey is conducted annually, and measures beneficiaries' experiences with and ratings of their former health plan. The results from the Assessment Survey are combined and reported with data from the M+C CAHPS survey of current enrollees in order to provide a more complete picture of all beneficiaries' experiences with a given plan. The Medicare CAHPS Disenrollment Reasons Survey is conducted quarterly. This survey asks beneficiaries about their reasons for leaving their plans. The Disenrollment Reasons Survey contains the SF-12® instrument. The HOS is not currently administered to disenrollees.

Statistical Adjustments

The goal of statistical adjustments is to eliminate any differences among sampling units that are not related to the quality of care received. The distribution of different subgroups of beneficiaries may differ from plan to plan, and these different subgroups may rate their care differently. When different subgroups rate their care differently, this can be the result of actual differences in the care these subgroups experience, and it can also be due to differences in how subgroups rate the exact same experiences (*response bias*). Case mix adjustment is an analytic strategy that can be used to measure and correct for these differences. *Non-response bias* refers to the propensity of different beneficiary subgroups to respond to the survey questionnaire or to participate in the telephone interview. Statistical weighting of the responses from the various subgroups can be used to adjust for non-response bias.

Adjustment for Case Mix

Case mix adjustment is used to adjust comparisons of survey ratings among sampling units. For the HOS, the sampling units are the plan contract areas. For the M+C CAHPS, the sampling units are the plan service areas, and for the MFFS CAHPS, the sampling units are geographic areas. The goal of case mix adjustment is to correct for response bias due to beneficiary characteristics such as demographics, socioeconomic status, and general health status, all of which may vary across sampling units. Performing these adjustments allows us to answer the question: What ratings would a plan or a collection of FFS providers have received if all plans or providers treated exactly the same types of patients?

Approaches to case mix adjustment vary across the different beneficiary surveys. The CAHPS® Survey Users Network (CAHPS®-SUN) recommends that the Medicaid CAHPS results be case mix adjusted for health status, age, and educational level. The results reported in the National CAHPS® Benchmarking Database (NCBD) reflect these adjustments (CAHPS® Survey Users Network, 2004a). On the other hand, NCQA does not recommend case mix adjustment for the Medicaid CAHPS surveys (NCBD, 2002a).

For both the M+C and MFFS CAHPS surveys, the same variables are used to case mix adjust the plan comparisons and geounit comparisons, respectively. The case mix model coefficients are estimated separately for the M+C and MFFS data (Elliott et al., 2003). The adjuster variables are age, education, general health perception, mental health perception, proxy status, a set of CMS region-by-age variables, and a set of CMS region-by-general health perception variables. For age, education, general health perception, and mental health perception, the individual response categories are recoded into 1/0 dummy variables[10] to avoid assuming a linear relationship with the dependent variable. Proxy status is represented by two dummy variables, one of which indicates whether or not a proxy assisted in filling out the survey, and the other of which indicates whether or not the assistance specifically took the form of answering the questionnaire for the beneficiary. The CMS region-by-age variables and CMS region-by-general health perception variables are incorporated into the case mix model to capture differences in the coefficients for age and general health perception across CMS regions.

A key goal of the Medicare CAHPS survey efforts is to permit the comparison of MFFS and M+C results from the two surveys within those local areas where both FFS and managed care are available. The use of the same case mix adjustment model for both surveys facilitates this type of comparison.

For the HOS, a more complex case mix adjustment methodology is employed. The SF-36® produces two summary measures of health status: the Physical Component Summary (PCS) score and the Mental Component Summary (MCS) score. First, the beneficiaries are classified as to whether their PCS and MCS scores are better, the same, or worse over the two-year period. A classification of better or worse is assigned to

[10] For each of these variables, one category is omitted to serve as a reference category.

beneficiaries whose amount of change falls above or below, respectively, the 95% confidence interval for an individual beneficiary, defined as 1.96 standard errors of measurement (HSAG, 2002a). Death within two years of the baseline survey (obtained from CMS files) is classified as a worse than expected physical outcome. Classification of death as a worse than expected physical outcome is advantageous because it combines mortality and health status into one physical health measure, without the need to assign a scale value to the outcome of death. This classification also reflects the fact that PCS has a strong statistical relationship to death. Death is not included in the calculation of mental health (MCS) outcomes because it has a much weaker relationship to MCS.

Separate case mix models have been developed for death, change in PCS, and change in MCS. A series of eight different death models, three different PCS models, and three different MCS models are used, since not all beneficiaries have data for all of the adjuster variables that could be used to calculate an expected score. The most comprehensive model possible is used for each beneficiary.

The models used to predict the probability of death include adjuster variables for demographic and socioeconomic characteristics, chronic medical conditions, functional status, and mode of survey administration. The models used to predict expected change in the PCS and MCS scores include adjuster variables for demographic and socioeconomic characteristics, and for mode of survey administration. A summary of the variables used as case mix adjusters for each of the current surveys can be found on Page 6 of Attachment D.

The results of the death and PCS models are combined to estimate the *expected* "alive and PCS same or better" rate for each plan. This rate is then compared to the *actual* "alive and PCS same or better" rate for each plan using a *t* test. In a similar fashion, the results of the MCS model are used to estimate the expected "MCS same or better" rate for each plan, which is then compared to the actual rate using a *t* test.

Adjustment for Non-Response Bias

It is well known that different subgroups of individuals vary in their propensity to respond to surveys. McCall (2004) examined non-response biases for the baseline HOS, follow up HOS, M+C CAHPS, MFFS CAHPS, and the two CAHPS disenrollment surveys for the year 2000. For each of these surveys, the impacts of age, race, gender, Medicaid status, reason for Medicare entitlement, number of hospitalizations, and health status upon the probability of responding was examined. CMS collects encounter data from M+C plans, and these data were the source of the number of hospitalizations. The encounter data were also used to calculate the Principal Inpatient Diagnostic Cost Group (PIP-DCG) risk adjustment score. The PIP-DCG combines principal inpatient diagnoses with demographic information into an index of predicted future health care expenditures. Since this measure is based on claims data, it provides a measure of health status that is available for both respondents and non-respondents. This permits the assessment of non-response bias associated with health status.

To facilitate comparisons of response rates among the surveys, a uniform definition of a completed survey was established: completion of the question regarding general health status. The comparative analysis revealed similar patterns of non-response biases across the six surveys. In general, beneficiaries aged 65 to 74 are more likely to respond than either younger or older beneficiaries; White and Asian beneficiaries are more likely to respond than Black beneficiaries; non-Medicaid-eligible beneficiaries are more likely to respond than Medicaid-eligible beneficiaries; and beneficiaries with higher health status are more likely to respond than beneficiaries with lower health status. The overall degree of response bias was modest.

The MFFS CAHPS analytic team has attempted to account for non-response bias by utilizing post-stratification weights when calculating geounit means. These weights are based on gender, age, race, and Medicaid status, and are used to adjust the survey results to reflect what would be found if all subgroups were equally likely to respond. Similar weights are generated for the M+C CAHPS survey, but to date, these weights have not been used in the calculation of plan means. Cioffi, Cleary, and Zaslavsky (2004) have pointed out that non-response weights are designed to estimate the plan scores that would have been observed if the respondents at each plan had the same characteristics as the non-respondents in that plan. If certain subgroups that are less likely to respond are concentrated in certain plans, then this type of adjustment could actually make the scores less comparable *across* plans. Using the 2002 M+C CAHPS survey data, these authors found that weighting for non-response biases had little impact on case mix adjusted plan comparisons.

To date, weighting for non-response bias has not been incorporated into the analytic strategy for the HOS. Hwang et al. (2002) found that age, race, gender and Medicaid status were most strongly related to the propensity to respond to the 1999 HOS. However, the overall impact of weighting the results for differential non-response was small.

Adjustment for Mode Effects

The tendency for survey responses collected by telephone to be more positive than survey responses collected by mail is well documented in the literature on survey research (e.g., Dillman, Sangster, and Tarnai, 1996). This mode effect has been observed for the Medicaid CAHPS (Fowler, Gallagher, and Nederend, 1999), the Medicare CAHPS (Iannacchione and Campbell, 2003), the SF-36® (Rogers, et al., 2000), and the Veterans versions of the SF-12 and SF-36 (Jones et al., 2001). Telephone responses can be more positive than mail survey responses for two different reasons: 1) individuals who are more likely to answer by telephone than by mail may be more likely to have positive opinions; and 2) among those individuals who are equally likely to answer by telephone or by mail, the answers given may be influenced by the mode used, with the telephone mode producing more positive responses. The first of these two effects is an example of non-response bias, and the second is referred to as response bias, or the pure mode effect. The only way to accurately estimate the separate contributions of these two effects to the overall mode effect is by means of a controlled experiment. Ideally, the pure mode

effect should be incorporated into the case mix adjustment process. For the HOS, MFFS CAHPS, and M+C CAHPS, mail surveys are distributed first, followed by telephone interviews if the mail survey effort is not successful. This means that the non-response bias is confounded with the pure mode effect, and this suggests that mode should not be used as a case mix adjustment variable for these surveys.

Handling of Missing Values

Frequently, respondents to health care surveys neglect to answer some of the questions. One simple solution to the missing data problem is to exclude from analysis any cases that have missing data on one or more variables. However, this may result in the loss of substantial amounts of data. The problem of missing data is even more pronounced for the elderly. In one study, 21% of the elderly had missing responses on the SF-36®, compared to 6% of the non-elderly (Rogers, Ware, Gandek, and Bayliss, 2000). Within the elderly population, the rates of missing data may vary by beneficiary characteristics. Therefore, exclusion of missing data may not only result in a smaller sample, but also an unrepresentative sample (Colsher & Wallace, 1989). Another solution is to use one of several statistical methods to impute, or estimate, the missing values, thereby achieving complete data for cases that would otherwise be dropped from analysis. Rogers, Qian, and Kazis (2004) note that either approach (omitting missing data or imputing missing data) may introduce bias into the results.

Missing Data Estimation in the SF-12 and the SF-36

The scoring algorithms for the SF-36® allow for the estimation of scores even when some data are missing. The original algorithm (referred to as the "half scale" rule) requires responses for at least one half of the items in each of the eight scales. In cases where at least one half of the items are present for a scale, the values for the missing items are estimated by substituting the average of the items that are present. If one or more of the scales are less than half complete, then estimation of the PCS and MCS scores is not possible (Ware et al., 1980).

More recently, the developers of the SF-36® created a Missing Data Estimation (MDE) utility. The MDE utility, derived from a psychometric approach known as Item Response Theory (IRT), can calculate a score as long as one item is answered within each scale. Furthermore, the MDE utility uses regression methods to estimate PCS and MCS scores in cases where one scale is missing. This proprietary utility is available exclusively through QualityMetric's scoring service. The utility allows the recovery of more missing data than the original half scale rule (Kosinski, Bayliss, Bjorner, and Ware, 2000).

Researchers at the VA have developed a regression-based imputation method for the Veterans SF-36, called the Modified Regression Estimate (MRE). The MRE produces results that are highly comparable to results from Version 1.0 of the SF-36®. Furthermore, the MRE approach recovers more cases than either the half scale rule or the MDE approach (Rogers, Qian, and Kazis, 2004). This methodology is in the public domain.

Missing Data Estimation in the CAHPS Surveys

The protocol established by NCQA for the calculation of HEDIS/CAHPS 3.0H composite scores deals with missing values by excluding them from the composite calculations. For each question in a composite, the mean of the responses is calculated across all respondents who answered the question. Therefore, the number of responses included in the calculation may vary from one question to another. To derive a composite score, the mean of the question means is calculated (NCQA, 2002a).

For the NCBD reports, values are imputed when the case mix adjuster variables have missing values. These missing values are replaced with the plan mean for the variable in question (CAHPS® Survey Users Network, 2003).

In the Medicare CAHPS surveys, the procedure for handling missing values is similar to that used by NCQA. However, when calculating the Medicare CAHPS composites, the weight given to an item within a given composite is based on the proportion of all respondents who answer that survey item.

Review of Key Issues

Research Design

Currently, the HOS is a longitudinal survey and the M+C CAHPS survey is a cross-sectional survey. There are several possible research designs for an integrated survey. These designs and a summary of their analytic advantages and disadvantages are described below.

Option 1: Maintain the Current Research Designs for HOS and M+C CAHPS

It is possible to maintain the current HOS and M+C CAHPS research designs even if the two surveys are integrated. To accomplish this, both the HOS and M+C CAHPS items would be administered to the baseline sample, with only the HOS items administered to the same sample at follow up.

This design would preserve the ability to measure changes in health status at the individual beneficiary level. However, it would not allow us to track changes in CAHPS scores over time at the individual beneficiary level. And, as long as a follow up survey is being fielded, it requires little additional effort to include the CAHPS items on the follow up questionnaire.

Option 2: Cross-Sectional Design for Both HOS and M+C CAHPS

Another option is to employ a cross-sectional approach for both the HOS and M+C CAHPS items. Under this scenario, the integrated HOS/CAHPS survey would be conducted on new samples each year, with no attempt to obtain either HOS or M+C CAHPS scores on the same respondents over time.

Whenever a longitudinal design is used, as with the current HOS, selection effects restrict the inferences that can be made about the change scores obtained. Over a two-year period, there is significant attrition due to mortality, voluntary disenrollment, and involuntary disenrollment. The loss of less healthy beneficiaries due to mortality and the loss of less satisfied beneficiaries due to voluntary disenrollment over time means that plan ratings obtained from the follow up survey may appear higher than they actually are. With a cross-sectional design, the drawing of a new sample of current beneficiaries means that more of the beneficiaries who are less healthy and less satisfied will be included in the survey sample, thereby mitigating these selection effects.

With a purely cross-sectional design for the integrated survey, it would not be possible to measure changes in either health status or CAHPS scores at the level of the individual beneficiary. This is the current situation for the CAHPS surveys, but would represent a fundamental change for the HOS, which has been developed from the premise of obtaining health status *change* scores. However, with a purely cross-sectional design, it would still be possible to calculate change scores at the plan level. And, useful information can be gained from relating variables that are measured at the plan level. For example, one recent study uncovered some positive relationships between plan MCS scores from the HOS and plan HEDIS/CAHPS scores obtained from NCQA's database. A statistically significant relationship was found between several of the HEDIS measures of effectiveness of care and CAHPS composite scores on the one hand, and MCS scores on the other (HSAG, 2003). Similarly, Schneider et al. (2001) have demonstrated a relationship between HEDIS measures and CAHPS composite scores.

If change scores based on cross-sectional plan samples do not differ very much from change scores based on longitudinal samples, a case can be made that the longitudinal design feature contributes little to the findings and can be eliminated. However, the sampling variance of the estimated difference based on a panel of repeated respondents is generally much smaller than the variance based on comparing independent samples (Cleary, Zaslavsky, and Cioffi, 2004). In other words, longitudinal comparisons provide much more reliable estimates of change than do cross-sectional comparisons, for the same sample size. This is an important consideration for the assessment of the impacts of QI activities. Further, cross-sectional designs do not allow us to identify causal relationships between variables, and also do not allow us to distinguish cohort effects from the effects of aging (Gayle, 2003). One way to increase the statistical power of cross-sectional comparisons over time is to compare demographically matched groups of beneficiaries across time periods. This matching design would avoid the attrition problems inherent in a longitudinal design. However, this approach would not provide as

much power as a longitudinal comparison, and the matching process could reduce the available sample size, further limiting power.

Option 3: Longitudinal Design for Both HOS and M+C CAHPS

A third option is to employ a longitudinal approach to *both* the HOS and M+C CAHPS items. This would mean conducting the complete integrated HOS/CAHPS survey at baseline and again at follow up on the same sample.

Longitudinal designs allow us to make inferences about the direction of causality. For example, an integrated survey would allow us to explore the relationship between health status and experiences of care. Do positive experiences of care lead to improved health status, or does health status lead to more positive experiences of care, or do both relationships exist? If both health status and satisfaction information are collected from the same individuals at two different points in time, we can answer this question (Gayle, 2003). For example, if the direction of causality is primarily from health status to experiences of care, then the correlation between health status at baseline and experiences of care at follow up will be relatively large, whereas the correlation between experiences of care at baseline and health status at follow up will be relatively weak.

Longitudinal surveys also allow us to detect selection effects by relating beneficiary satisfaction at baseline to the propensity to respond at follow up. For example, if the beneficiaries who are dissatisfied with their care at baseline are less likely to respond to the follow up survey than the satisfied beneficiaries, this is evidence for a selection effect. There is, in fact, evidence for this specific relationship in the CAHPS literature (Allen, 1998).

It should also be pointed out that an individual beneficiary is much more likely to be sampled in consecutive years if that beneficiary belongs to a plan with a small enrollment (Cleary, Zaslavsky, and Cioffi, 2004). Therefore, for small plans, the M+C CAHPS can be regarded as a partially longitudinal survey, although it was not designed for that purpose.

This design provides the advantages of longitudinal designs (the measurement of change at the beneficiary level, the elucidation of causality, and the measurement of selection effects), with a key advantage of cross-sectional designs (more accurate measurement of overall plan ratings).

The Allen study (1998) cited previously showed that cross-sectional and longitudinal survey results derived from the same group of beneficiaries differed dramatically. Allen administered the CAHPS survey to managed care and indemnity plan members in 1993 and 1995. Data were collected from 14,587 members in 1993 and 9,018 members in 1995. A total of 5,729 members completed both surveys and remained in the same plan during the study period; these formed the longitudinal sample. Allen found a pronounced tendency for dissatisfied managed care plan members to switch to indemnity plans between measurement periods, a finding that was totally masked by examining the cross-

sectional results only. The cross-sectional comparison for the indemnity plans supported the conclusion that indemnity plan members increased their overall satisfaction by six percentage points over two years. In contrast, the longitudinal results showed that indemnity plan member satisfaction *decreased* by two percentage points over two years.

This study demonstrates that cross-sectional and longitudinal results together provide a more accurate picture of plan ratings than does either type of result separately. In the HOS design, it has been the practice to report out the cross-sectional ("Baseline") and longitudinal ("Performance Measurement") results separately. However, there is no reason in principle why the results for the baseline respondents and follow up respondents for a given year could not be combined to provide a more accurate picture of plan ratings for that year's survey. To produce accurate findings, the results from these two populations would need to be weighted to reflect their proportions in the overall plan population.

Surveying Disenrollees

Although a longitudinal survey can be used to measure the magnitude of selection effects due to disenrollment, it cannot by itself provide an estimate of their impact on the survey results. Only by surveying a sample of disenrollees can we know how their ratings impact overall plan results. Lied et al. (2003) found that voluntary disenrollment rates are strongly related to CAHPS measures of experiences of care. Bender, Lance, and Guess (2003) demonstrated that deriving plan ratings from both enrollees and disenrollees resulted in significantly lower CAHPS scores compared to the CAHPS scores derived from enrollees only. A failure to include disenrollees not only overstates overall plan performance, but will also bias the results in favor of plans with high rates of disenrollment due to dissatisfaction. Change scores calculated for disenrollees do not provide a valid measure of plan impact on disenrollee health or on ratings of health care experiences, because by definition they will have left the plan at some point before the follow up measure is obtained. For the purpose of detecting selection effects, it will suffice to assess disenrollees at baseline only.

In addition, a survey of disenrollees can be used to query respondents about their specific reasons for disenrollment. This is being done currently for the M+C CAHPS survey. This information can help plan administrators design beneficiary retention strategies.

The choice of survey research design and the decision to include or exclude disenrollees also have implications for sampling and for survey cost. These implications will be discussed further in the sections on Sampling Methods, and Survey Costs and Respondent Burden later in this report.

Analysis

Adjustment for Case Mix

Integrating the HOS and M+C CAHPS surveys does not necessarily require an integrated case mix adjustment model. The HOS and M+C CAHPS results can continue to be adjusted using separate models. However, integration of the two surveys would mean that new adjuster variables would be available that might improve the current case mix adjustment models. For example, the current CAHPS survey case mix adjustment model includes the general health perception and mental health perception items as adjuster variables. In an integrated survey, the more sensitive PCS and MCS measures could be substituted for these single rating items to obtain a possible improvement in the adjustment of plan scores. Research on the MFFS CAHPS data, which contain responses to the SF-12®, indicates that both the MCS and the mental health perception items substantially improve the predictive power of a base model built with age, general health perception, education, and proxy status. The MCS performed better than the mental health perception item, but only slightly. The PCS score added virtually no improvement to the base model (Elliott et al., 2003).

Since some questions would have to be dropped from both the HOS and M+C CAHPS surveys in order to create an integrated survey of reasonable length, new case mix adjustment models may have to be built and tested. These tests are important because they will cast light on the usefulness and validity of the case mix adjustment of an integrated survey's results.

This testing may reveal that the adjusted plan scores differ very little from unadjusted plan scores. This would not be an unexpected finding. For the M+C CAHPS, Zaslavsky et al. (2001) found that case mix adjustments were mostly small and did not greatly change the ordering of plans in terms of the ratings they received. For Cohort I of the HOS, the most complete case mix adjustment model for death explained 17% of the variation in plan scores, and the most complete case mix adjustment models for the PCS and MCS scores explained less than one half of one percent of the variation in plan scores (Rogers, Gandek, and Sinclair, 2004). However, even if this testing reveals that the case mix adjustments are mostly minor, it may also reveal that there are a few plans for which the case mix adjustment procedure produces a substantial adjustment.

Testing of the new case mix adjustment models is also important in order to check for possible violations of model assumptions. If, for example, the effect of self-reported health status on the quality of care varies across plans, then entering self-reported health status into the case mix adjustment model will produce estimates of plan differences that are not completely accurate. In statistical terms, this situation would result in a model coefficient for self-reported health status that varies across plans. In a study of 54 commercial and 31 Medicaid plans, Elliott et al. (2001) found some heterogeneity of coefficients for age, education, and self-reported health status when these variables were used to case mix adjust CAHPS scores. Zaslavsky, Zaborski, and Cleary (2000) also found some evidence for heterogeneity of coefficients for case mix adjuster variables in a

study of CAHPS results for plans in the state of Washington. The case mix adjustment models for an integrated survey should be checked for this potential violation of model assumptions.

Adjustment for Non-Response

Unlike the MFFS CAHPS, the HOS does not use post-stratification weighting to adjust the survey results. If an integrated survey is developed, post-stratification weights could be calculated, and their impact on plan comparisons assessed. If the impact is notable, then post-stratification weighting should be considered for the integrated survey.

Handling of Missing Values

Methods of imputation that depend on substituting mean values for missing values are relatively simple to implement. However, imputation by substitution of mean values will produce variances that are understated, and the resulting correlations among variables can also be misleading (Levy and Lemeshow, 1999). There are other methods available that largely avoid the underestimation of variances, for example hot deck imputation and regression based imputation (Levy and Lemeshow, 1999). Multiple imputation methods (Rubin, 1987), replace each missing value with multiple imputed values, which allows the researcher to estimate the degree of uncertainty associated with the missing values.

Integrating the two questionnaires does not necessarily require the adoption of a single strategy for handling missing values. Given the variety of approaches that are available for dealing with missing values, a thorough review of these approaches and their pros and cons for use with an integrated instrument is warranted.

Additional Analyses

An integrated instrument would permit a number of analyses that cannot be performed with the currently separate HOS and M+C CAHPS surveys. These additional analyses would make the data more useful to decision makers. Below are descriptions of some of the additional analyses that an integrated instrument would make possible.

Relating HOS and CAHPS Scores

If the integrated instrument were cross-sectional only (Option 2), we would be able to determine which specific CAHPS measures are most related to specific measures of health status. As discussed earlier, a longitudinal implementation of an integrated survey (Option 3) can help to pinpoint the CAHPS measures that lead to improvements in health status, and/or which CAHPS measures are most likely to be impacted by improvements in health status.

A longitudinal version of an integrated survey would also allow us to examine the relationship between experiences of care and health status *change* scores.

Currently, a number of beneficiaries complete both the HOS and the M+C CAHPS surveys, and all of the data files for these two surveys contain a unique identifier in the form of the Health Insurance Claim Number (HICN). These two facts permit a preliminary examination of the relationship between health status and experiences of care at the beneficiary level. HOS survey respondents for the years 1998 through 2002, and M+C CAHPS survey respondents for the years 2000 through 2002 were merged by HICN in order to determine the number of beneficiaries who responded to both surveys. Two merged databases were created:

- Linked Database Number One contained 53,560 beneficiaries who completed at least one HOS and one M+C CAHPS survey within a six-month time period. This database allows us to examine the relationship between PCS/MCS scores and M+C CAHPS scores at the beneficiary level.

- Linked Database Number Two contained 2,424 beneficiaries who completed both a baseline and follow up HOS, as well as two M+C CAHPS surveys, within a two-year window. This database allows us to examine the relationship between PCS/MCS change scores and M+C CAHPS change scores at the beneficiary level.

Using Linked Database Number One, the relationships between PCS and MCS scores on the one hand, and the four global CAHPS ratings on the other, were examined at the beneficiary level. The results are shown in Table 10.

TABLE 10
RELATIONSHIPS BETWEEN HEALTH STATUS MEASURES AND CAHPS GLOBAL RATINGS
(*N* = 53,560)

Global Rating	*Correlation with PCS Score*	*Correlation with MCS Score*
Personal Doctor	.03	.08
Specialist	.05	**.10**
Health Care	.08	**.14**
Health Plan	.07	**.12**

Source: Linked Database Number One (HOS results for 1998-2002 and M+C CAHPS results for 2000-2002).
Note: Bold type indicates a small effect size.

For correlation coefficients, the conventional level for a small effect size is .10 (Cohen, 1988). The relationships between the PCS scores and the global ratings were modest; none of the four correlations met this criterion for a small effect size. On the other hand, three of the four correlations between the MCS score and the global ratings did meet this criterion.

The relative impacts of experiences of care and health status upon one another at different time periods have practical implications for the design of QI activities. For example, if CAHPS ratings of access to care are strongly related to subsequent ratings of physical health status, this relationship would provide health care managers with a focus for a specific QI initiative. Assessing these relative impacts is not possible with cross-sectional surveys.

Several researchers have examined these relationships using data from longitudinal surveys. Ren et al. (2001) analyzed data from a longitudinal survey of VA patients, and found that health status was strongly correlated with satisfaction with care measured at a later time. They also found a strong correlation between satisfaction and health status measured at a later time. Marshall et al. (1996) studied SF-36® results over time. They found that baseline satisfaction was correlated with subsequent mental health status and that baseline mental health was also correlated with subsequent satisfaction. No such relationships were observed for physical health.

Linked Database Number One allows us to study longitudinal relationships among HOS and CAHPS scores. PCS and MCS scores were correlated with the CAHPS global ratings measured at a subsequent point in time (within six months; $N = 36{,}659$), and the CAHPS global ratings were also correlated with PCS scores measured at a subsequent time ($N = 16{,}901$). The results are shown in Tables 11 and 12.

TABLE 11
LONGITUDINAL RELATIONSHIPS BETWEEN PCS SCORES AND CAHPS GLOBAL RATINGS
($N = 53{,}560$)

Measurement Sequence ($N = 36{,}659$)	*Correlation*	*Measurement Sequence ($N = 16{,}901$)*	*Correlation*
PCS → Rating of personal doctor	.03	Rating of personal doctor → PCS	.03
PCS → Rating of specialist	.05	Rating of specialist → PCS	.04
PCS → Rating of health care	.09	Rating of health care → PCS	.09
PCS → Rating of health plan	.07	Rating of health plan → PCS	.06

Source: Linked Database Number One (HOS results for 1998-2002 and M+C CAHPS results for 2000-2002.
Note: Bold type indicates a small effect size.

The relationships between PCS scores and the global ratings were modest, and the correlations obtained differed little according to the sequence of measurement.

TABLE 12
LONGITUDINAL RELATIONSHIPS BETWEEN MCS SCORES AND CAHPS GLOBAL RATINGS
(*N* = 36,659)

Measurement Sequence (N = 36,659)	*Correl-ation*	*Measurement Sequence (N = 16,901)*	*Correl-ation*
MCS → Rating of personal doctor	.08	Rating of personal doctor → MCS	.06
MCS → Rating of specialist	**.11**	Rating of specialist → MCS	.08
MCS → Rating of health care	**.15**	Rating of health care → MCS	**.12**
MCS → Rating of health plan	**.12**	Rating of health plan → MCS	**.12**

Source: Linked Database Number One (HOS results for 1998-2002 and M+C CAHPS results for 2000-2002.
Note: Bold type indicates a small effect size.

The relationships between MCS scores and the four global ratings were stronger than for the PCS scores, with several measures reaching the conventional level for a small effect size. However, once again the sequence of measurement had little impact on the size of the correlations obtained.

Another beneficiary-level analysis that an integrated survey would make possible is an examination of the relationship between PCS/MCS *change* scores and M+C CAHPS *change* scores. Table 13 shows the results of such an analysis using Linked Database Number Two.

TABLE 13
RELATIONSHIPS BETWEEN CHANGES IN HEALTH STATUS AND CHANGES IN THE CAHPS GLOBAL RATINGS
(*N* = 2,424)

Global Rating Change Score	*Correlation with PCS Change Score*	*Correlation with MCS Change Score*
Personal Doctor	-.03	.03
Specialist	**-.11**	.06
Health Care	-.04	.02
Health Plan	-.02	.02

Source: Linked Database Number Two (HOS results for 1998-2002 and M+C CAHPS results for 2000-2002).
Note: Bold type indicates a small effect size.

Only one of the correlations met the criterion for a small effect size: the correlation between the change in the specialist rating and the change in the PCS score. Declines in PCS were associated with improvements in the rating of the specialist. This may be due to the fact that beneficiaries with worsening health tend to have more frequent and intensive contact with their specialists. Although the correlations are small, declines in physical health seem to be related to improvements in CAHPS ratings, whereas declines in mental health seem to be related to deteriorations in CAHPS ratings.

The analyses described above are only possible for those beneficiaries who happen to complete both the HOS and M+C CAHPS within a given time period. An integrated survey would permit such analyses for all of the beneficiaries completing a baseline and a follow up questionnaire, and would have the additional advantage of providing health status and experience of care ratings at the same points in time.

Medical Condition-Specific Norms

Currently the HOS contains self-report items for 13 chronic medical conditions, and the M+C CAHPS survey contains items that measure experiences of care received from physicians, specialists, nurses, etc. If an integrated survey were to contain both types of items, then it would be possible to develop norms for specific types of beneficiaries. It would be very useful for diabetic beneficiaries, for example, to see plan rankings based on the CAHPS ratings of diabetic beneficiaries only. As it happens, a test of this type of analysis is possible with current data. The 2000 M+C CAHPS survey contained self-reports of five chronic medical conditions as well as the standard CAHPS items and global ratings. The five chronic medical conditions were heart disease, cancer, stroke, chronic obstructive pulmonary disease (COPD), and diabetes.

Using standard Statistical Process Control (SPC) techniques (Longo and Bohr, 1991), the 2000 M+C CAHPS data were analyzed separately for individuals reporting each of the five chronic medical conditions. For each condition, the CAHPS global rating of satisfaction with the health plan was compared across plans, and both positive and negative plan outliers were identified. Specifically, the proportion of 8, 9 and 10 ratings was compared across plans using a *p* chart. Plans with less than 25 individuals reporting the condition in question were excluded from the analysis. For each of the five conditions, large numbers of both positive and negative outlier plans were found, indicating a high degree of variability of plan ratings among each of these medical condition subgroups. For each condition, there were also a fair number of positive outlier plans that were outliers for that condition only (see Table 14 below). This suggests that some plans may have developed unique expertise in serving beneficiaries with specific conditions.

TABLE 14
SUMMARY OF SPC ANALYSIS FOR OVERALL RATING OF PLAN
(PROPORTION OF RATINGS EQUAL TO 8, 9, OR 10)

	CHRONIC MEDICAL CONDITION				
	Heart Disease	***Cancer***	***Stroke***	***COPD***	***Diabetes***
Overall Proportion of 8s, 9s, and 10s	.78	.79	.77	.75	.78
Range of Plan Proportions	.52 to .95	.40 to .98	.46 to .99	.47 to .97	.54 to 1.00
Total Plans	276	266	204	116	270
Total Positive Outlier Plans	60	40	23	20	38
Total Negative Outlier Plans	57	54	29	18	42
Total Unique Positive Outlier Plans*	18	6	5	5	6

* Plans that were positive outliers for this chronic medical condition only
Source: 2000 M+C CAHPS survey data
Note 1: Plans with fewer than 25 cases for a given chronic medical condition were excluded from the analysis.
Note 2: The data were not case mix adjusted.

Functional Status-Specific Norms

In addition to developing norms for specific types of beneficiaries, an integrated survey would allow us to develop norms for beneficiaries at particular levels of functioning, as measured by the SF-12 or SF-36. It is possible that beneficiaries with low levels of health status may experience better care in specific areas of the country, or in specific plans. These areas or plans may not be the same for beneficiaries with high levels of health status. Since the MFFS CAHPS survey is administered to beneficiaries in 276 geounits across the country, and contains both CAHPS items and the SF-12®, data from this survey can be used to identify small area variations in the delivery of care to the lowest functioning and highest functioning members of the Medicare population.

To test the above hypothesis, data from the 2002 MFFS CAHPS survey were analyzed. First, beneficiaries who fell into the highest quartile of PCS scores were considered to have high physical health status, and were placed into a separate subgroup for analysis. In a similar fashion, three other subgroups were created: low physical health status, high mental health status, and low mental health status. For each of these four groups, the global rating of health care received was calculated across geounits, and geounit outliers were identified. Table 15 shows that outlier geounits were identified for each of the four

subgroups. Furthermore, the positive outliers (high performing geounits) identified for each subgroup tended to be very distinct. For example, the geounits with the highest health care ratings from the low physical health status beneficiaries are not the geounits with the highest health care ratings from the *high* physical health status beneficiaries.

TABLE 15
GEOUNIT OUTLIERS FOR HIGH-FUNCTIONING[1] AND LOW-FUNCTIONING[2] BENEFICIARIES: OVERALL RATING OF HEALTH CARE

	SUBGROUP			
	Beneficiaries in Highest Quartile of PCS Scores (PCS Scores from 50.7 to 70.0)	***Beneficiaries in Lowest Quartile of PCS Scores (PCS Scores from 10.4 to 27.1)***	***Beneficiaries in Highest Quartile of MCS Scores (MCS Scores from 59.2 to 78.3)***	***Beneficiaries in Lowest Quartile of MCS Scores (PCS Scores from 7.8 to 48.5)***
Overall Rating of Health Care Received[3]	2.44	2.24	2.48	2.18
Range of Geounit Means	2.13 to 2.68	1.84 to 2.54	2.25 to 2.76	1.78 to 2.48
Total Geounits	276	276	276	276
Total Positive Outlier Geounits	5	5	9	19
Total Negative Outlier Geounits	10	12	8	11
Total Unique Positive Outlier Geounits[4]	4	3	8	17

[1] Defined as beneficiaries falling into the top 25 percent of scores.
[2] Defined as beneficiaries falling into the bottom 25 percent of scores.
[3] The original 0 to 10 scale was compressed as follows: 0 to 7 = 1; 8 and 9 = 2; 10 = 3.
[4] Plans that were positive outliers for this group only.
Source: 2002 MFFS CAHPS survey data
Note: The data were case mix adjusted using the standard set of adjuster variables for Medicare CAHPS surveys.

It is interesting to note that only one geounit was a positive outlier for all four subgroups. Only one geounit was a positive outlier for both the low physical health status beneficiaries and the low mental health status beneficiaries. The 19 geounits that were positive outliers for the low mental health status beneficiaries tended to cluster in the northern part of the Midwest. If an integrated instrument is launched in the M+C

population, a similar analysis can be used to identify *plans* that perform particularly well at serving low functioning beneficiaries.

Health Status and Compliance with Treatment Regimens

An integrated survey would also provide the opportunity to explore the relationship between health status, as measured by the PCS and MCS scores, and compliance with certain preventive health regimens. The current MFFS CAHPS survey contains both the SF-12® and five preventive health questions that ask if the beneficiary has obtained a mammogram, a pap smear, a prostate screening, a flu shot, and a pneumonia shot in the past 12 months.

It is well documented that depressed individuals are less likely to follow through with obtaining recommended tests or treatments (e.g., Polonsky et al., 2003). And, Ware et al. (1994) have established that individuals with SF-36® MCS scores less than or equal to 42 are much more likely to be diagnosed with depression. Furthermore, Ware et al. (2002) have demonstrated that the SF-36® and SF-12® perform very similarly in identifying depression. For this report, data from the 2002 MFFS CAHPS survey were used in a preliminary test of the relationship between SF-12® MCS scores and compliance with five preventive treatments in the MFFS population. The MCS scores were dichotomized into two groups: individuals at risk for depression (MCS score less than or equal to 42) and individuals not at risk for depression (MCS score greater than 42). Table 16 summarizes the results.

TABLE 16
PERCENTAGE OF ELIGIBLE BENEFICIARIES COMPLYING WITH VARIOUS PREVENTIVE TREATMENTS

	PERCENTAGE OF BENEFICIARIES REPORTING:				
	***Mammo-gram*[1]**	***Pap Smear*[1]**	***PSA Test*[2]**	***Flu Shot*[3]**	***Pneumonia Shot*[3]**
Beneficiaries at Risk for Depression (MCS <= 42)	51%	37%	51%	66%	59%
Beneficiaries not at Risk for Depression (MCS > 42)	59%	41%	63%	70%	63%
Total Beneficiaries	134,234	132,984	103,042	330,285	311,188
***p* Value**	< .0001	< .0001	< .0001	< .0001	< .0001
Effect Size	.07	.03	.10	.04	.03

[1] Mammogram and pap smear rates calculated for female beneficiaries only. Source: 2001 and 2002 MFFS CAHPS data.
[2] PSA test rates calculated for male beneficiaries only. Source: 2001 and 2002 MFFS CAHPS data.
[3] Source: 2000, 2001 and 2002 MFFS CAHPS data.

For all five preventive health behaviors, the beneficiaries with an MCS score less than or equal to 42 were less likely to have performed the behavior than beneficiaries with an MCS score greater than 42. In each case, the chi-square test yields a highly significant result. Effect sizes provide a more accurate measure of statistical importance in cases, such as this one, where sample sizes are very large (Cohen, 1988). The effect size calculations suggest more modest effects that fall below the conventional level for a small effect size of .20.

Adverse Selection

There is strong evidence that more healthy beneficiaries migrate out of FFS and into an M+C plan in areas where this choice is available, leaving the FFS program with a less healthy beneficiary population. This evidence comes from analysis of self-ratings of overall health (Murgulo, 2002), inpatient encounter data (Greenwald, Levy, and Ingber, 2000), and cost and market share data (Cao and McGuire, 2002). This phenomenon is referred to as *adverse selection*.

Of the 276 geounits surveyed with the MFFS CAHPS instrument in 2000, 160 included one or more M+C plans (Iannacchione, Bernard, and Elliott, 2002). This means that the MFFS CAHPS survey can be used to study adverse selection in depth using PCS and MCS scores rather than simpler measures such as global self ratings of health. If MFFS beneficiaries in the geounits with one or more M+C plans have lower health status scores than the MFFS beneficiaries in the geounits with no M+C plans, this would be evidence for adverse selection. Shimada et al. (2004) examined this relationship at the county level and confirmed that M+C plans continue to experience favorable selection. Beneficiaries who rated their health as good or better were more likely to be enrolled in M+C plans than their less well peers. The amount of favorable selection observed for M+C plans was weaker in areas with high M+C penetration, more M+C options, and more competition among plans.

The MFFS CAHPS survey is currently cross-sectional. If an integrated survey with longitudinal measures of both functional status and experiences of care were to be implemented, even stronger inferences about the nature and magnitude of adverse selection would be possible.

Use of Modular Questionnaires

Some health care survey researchers have recommended the use of *modular surveys* to improve the ability of healthcare managers to act on the survey results. For example, the A-CAHPS survey will assess the quality of ambulatory care at different levels of the health care system while still retaining many features of the current CAHPS health plan surveys (CAHPS® Survey Users Network, 2004b). The A-CAHPS would contain a core set of questionnaire items, but the core questionnaire would also permit the insertion of modules containing questions targeted at specific departments, providers, or health care processes. The results from the modules would provide plans with additional information to guide their QI efforts.

From time to time, supplemental question sets have been added to the HOS, including supplements for smoking, urinary incontinence, and healthy days. These supplements differ from the modules described above in that, for a given survey year, all of the questionnaires distributed contain the exact same supplement or supplements. The future addition of other supplements to the HOS has also been discussed.

National Research Corporation (NRC), a major vendor for health care surveys, has taken the modular concept a step further. Rather than fielding distinct surveys for distinct service modalities, such as hospital inpatients and ambulatory surgery patients, NRC has the capability to generate patient-specific surveys based on the specific modality or modalities the patient has used. For example, if a patient was hospitalized, visited a physician's office, and had an outpatient radiology exam in a given period of time, that patient would receive a questionnaire that contains three modules: the inpatient module, the physician's office module (with the name of the physician seen by the patient printed on the questionnaire), and the outpatient radiology module. An automated questionnaire generation system links to the appropriate data files and creates questionnaires containing the appropriate modules for each patient (NRC, 2004).

If an integrated survey with a modular design were to be fielded, this would allow plan administrators to incorporate modules that target the issues and quality initiatives of greatest interest. The use of a modular questionnaire would complicate the process of sampling, analyzing the results and preparing comparative reports, however.

Preliminary Recommendations

1. *Administer the integrated survey longitudinally, in a manner similar to the administrative protocol currently used for the HOS*. The baseline and follow up versions of this survey should each contain both HOS and CAHPS items. Before making a firm decision to proceed with a longitudinal design, compare the performance of longitudinal, cross-sectional, and matched cross-sectional designs using current HOS data.

2. *Implement a similar survey of disenrollee beneficiaries who have disenrolled shortly after baseline*. Otherwise, plan ratings will appear higher than they actually are.

3. *Derive plan scores for any given year from a combination of baseline survey respondents, follow up survey respondents, and disenrolled respondents*. Combining results from these three groups of respondents will allow administrators and researchers to understand the impact of attrition on plan results. The results from each of the three subgroups will need to be weighted to reflect their proportions of the plan population.

4. *Each year, report plan comparisons for both the three-group composite described above and for HOS and CAHPS change scores*. Both sets of plan comparisons are needed for a full understanding of plan performance.

5. *On an annual basis, develop a formal plan for implementing an analytic strategy for producing results that are of use to individual plans in designing QI interventions*. Recommended elements for this analytic plan include:

a) An assessment of plan performance on disease-specific and level of function-specific norms (where sample sizes are adequate).

b) A comparison of plan performance on the extent to which compliance with treatment protocols is occurring.

c) A plan-level identification of those aspects of the care experience at baseline that are most strongly associated with PCS and MCS change scores at follow up.

6. *Develop a case mix adjustment model that includes only those variables that significantly impact the variation among plan scores.* An additional adjuster variable should not be added to the model unless this addition results in a statistically significant improvement in the performance of the model. A parsimonious case mix adjustment model will also be easier to communicate to users of the survey results.

7. *Test the resulting case mix adjustment model to determine if the regression coefficients of the adjuster variables vary significantly across plans*. If significant heterogeneity of coefficients is found, then additional steps may need to be taken to achieve a valid model. One such solution is to incorporate a question into the model that is intended to be a pure measure of response bias and is unlikely to vary across plans. Such an item is currently being tested in the MFFS CAHPS ("We want to know how you feel about your life overall. How would you rate your life now?"). Initial results indicate that this item performs well when added to the standard case mix adjustment model for the MFFS CAHPS (Elliott, 2004). The addition of a similar item to an integrated survey should be considered.

8. *Test the effects of post-stratification weights carefully before utilizing them in the reporting of results*. Post-stratification weighting can work at cross-purposes with case mix adjustment.

VI. SAMPLING METHODS

Overview of Current Sampling Methods

The current HOS and CAHPS sampling designs differ in several significant ways.

Eligibility

Eligibility requirements for participation in the HOS and M+C CAHPS are similar but not identical. To receive the HOS, a beneficiary must be continuously enrolled for six months prior to March 1 of the measurement year, and *may* be institutionalized. To participate in the M+C CAHPS survey, a beneficiary must be continuously enrolled for six months prior to July 1 of the measurement year, and may *not* be institutionalized.

The HOS is administered to members of M+COs, SHMOs, Continuing Cost Contracts, Private FFS plans, and Medicare Alternative Payment Demonstration Plans. The M+C CAHPS survey is administered to members of the first three groups.

Sampling Units

For the HOS, the sampling units are the contract areas. For the M+C CAHPS, the sampling units are plan service areas, each consisting of the members of a plan that can be assigned to that geographic area. Some of the more heavily populated plan service areas are subdivided further (Zaslavsky et al., 2000). In the FFS setting, the beneficiaries are not restricted to use of a particular plan. Therefore, CAHPS researchers implemented a geographic approach in the creation of sampling units for this survey. In 2000, the nation was divided into 276[11] geographic areas, termed "geounits," by the MFFS CAHPS survey team. Each geounit consisted of one or more counties. Several factors were simultaneously considered in grouping the counties into geounits, including a desire to roughly equalize the number of respondents per geounit; geographical contiguity; and the creation of geounits that would not cross the boundaries of managed care contract areas, metropolitan statistical areas (MSAs), or state boundaries.

Sample Size

The two goals of the MFFS CAHPS sampling process are to 1) achieve a total of 300 respondents in each geounit and 2) facilitate the comparison of CAHPS results between the M+C and MFFS populations at the state level. Within each geounit, a simple random sample by county is drawn. The allocation is approximately proportional to the size of the MFFS population within that geounit, but has been altered somewhat to more closely resemble the distribution of M+C beneficiaries. In some cases, the total geounit allocations were increased to more closely resemble state-level M+C beneficiary

[11] An additional geounit has since been added, bringing the total to 277.

distributions. Finally, the geounit sample size has since been adjusted upward for geounits with relatively low response rates (generally geounits in large urban areas) and downward for geounits with relatively high response rates.

In order to generate accurate estimates for the MFFS population as a whole, a sampling weight is assigned to each selected beneficiary. This weight is the inverse of the selection probability, so that the differential selection probabilities for beneficiaries in each geounit would be reflected. These weights are further adjusted to reflect the potential bias caused by differential non-response. The non-response adjustments are based upon age, gender, race, and Medicaid status (Iannacchione and Campbell, 2003).

For the M+C CAHPS, 600 beneficiaries are drawn from each sampling unit, with a goal of obtaining 300 completed surveys. Post-stratification weights are calculated but so far are not used in reporting the M+C CAHPS results. For the Disenrollment Assessment Survey, the target number of completed surveys is 300, and for the Disenrollment Reasons Survey, the target number of completed surveys is 388 across four quarters.

Because the HOS is a longitudinal survey, the possibility of non-response to the baseline and/or follow up surveys must be factored into the calculation of the target number for the baseline survey. For the HOS, the initial target for the follow up survey sample size was 500 completes. This target allowed for a two-point difference in PCS or MCS scores between plans to be detectable with 99% power at the 95% significance level. Based on an analysis of *Cohorts I* and *II* attrition rates and response rates, a baseline sample size of 1,420 would be needed to achieve this target (NCQA, 2002b). Currently the HOS is averaging approximately 400 completes from the follow up survey, and the current baseline sample size has been set at 1,000. For the *Cohort III Follow Up* survey, the average number of follow up respondents per plan was 408. Seventy-five percent of the plans had at least 274 respondents (HSAG, 2003).

Results for the first three baseline and follow up cohorts of the HOS were examined in order to estimate the attrition rates that can be expected for the average plan. To obtain these estimates, the total numbers of eligible members, members lost through death, members lost through disenrollment and members who responded to the survey were summed across the three cohorts. These totals were then divided by the sum of the numbers of plans that participated in each cohort.

Table 17 shows the results of this analysis.

TABLE 17
SOURCES OF ATTRITION FOR MEDICARE HEALTH OUTCOMES SURVEY

Category	*Average Number per Plan*	
Survey distributed at baseline		1,161
Ineligible to participate in baseline survey	minus	32
Survey returned; insufficient data for calculation of a PCS and MCS score at baseline	minus	381
Survey returned; sufficient data for calculation of a PCS and MCS score at baseline	equals	748
Involuntarily disenrolled	minus	205
Deceased	minus	53
Resurveyed	equals	490
Ineligible to participate in follow up survey	minus	5
Insufficient data for calculation of a PCS and MCS score at follow up	minus	79
Sufficient data for calculation of a PCS and MCS score at follow up	equals	404
Response rate: Based on baseline sample		66%
Response rate: Based on follow up sample		83%

Source: Medicare Health Outcomes Survey, *Cohorts I, II and III (Baseline and Follow Up)*, 1998 – 2002
Note 1: Limited to beneficiaries aged 65 or older.
Note 2: The number of surveys distributed at baseline (1,161) exceeds the starting sample size of 1,000 due to the presence of plans with multiple market areas.
Note 3: Figures are based on plans with contracts still in place at the time of follow up.

The above results show that, for the average plan, an initial sample size of 1,161 generated a final sample size of 404. This implies that an initial sample size of 1,336 would be required to achieve a final sample size of 500. If the sample were to also include beneficiaries younger than 65, the initial sample size would have to be increased by approximately 100.

Review of Key Issues

Eligibility

Eligibility requirements for participation in the HOS and M+C CAHPS currently differ. First, institutionalized beneficiaries are eligible to participate in the HOS; institutionalized beneficiaries are not surveyed with the M+C CAHPS survey. Second, the dates of the eligibility periods differ between the two surveys. An integrated survey would require a uniform definition of eligibility.

Sample Size

The choice of research design will influence the sample sizes needed for statistical reliability of the results. A longitudinal survey requires a larger sample at baseline in order to obtain a reasonable sample size at follow up. Cross-sectional survey sample sizes can be smaller, but these surveys only allow calculation of change scores at the plan level. If the integrated survey were to be a longitudinal survey, CAHPS change scores would be available for the first time.

The three survey research design options discussed earlier lead to different recommendations for how sampling should be conducted. If the research design for an integrated survey has a longitudinal component (Options 1 or 3), then the current baseline sample of roughly 1,160 per plan will achieve a reasonable sample size (400) for the follow up survey. If Option 2 (the purely cross-sectional integrated survey) is selected, then the current sample size of 600 used for the M+C CAHPS survey should suffice.

If disenrollees are surveyed as well, then this would increase the needed sample sizes even further.

Non-Response Bias

The fact that certain beneficiary subgroups are less likely to respond to surveys means that sample sizes may need to be increased for these subgroups. Post-stratification weights can help to adjust the overall results for non-response bias, but the only way to increase the precision of the estimates for these subgroups is to increase their sample sizes.

Modular Surveys

The use of survey modules may require a major increase in sample sizes, particularly if modules designed to compare results for specific group practices, care sites, or individual providers are included in the questionnaire.

Preliminary Recommendations

1. *Adopt the current longitudinal design and current sample sizes of the HOS for the integrated survey*. These sample sizes should be sufficient to generate a follow up sample size of approximately 400 per plan.

2. *Add an annual disenrollee survey with a target final sample size of approximately 400*.

VII. COST AND BURDEN

Overview of Current Survey Costs and Respondent Burden

Any survey of Medicare patients imposes costs on the survey administrators, and a time burden on the beneficiaries who respond to the survey. One obvious advantage of combining the HOS and M+C CAHPS surveys is the reduced administrative costs and respondent burden that would result. Table 18 summarizes estimated costs for several current Medicare surveys as well as the SHEP.

TABLE 18
ESTIMATED SURVEY COSTS

Survey	***MFFS CAHPS Survey***	***M+C CAHPS Survey***	***Adult Medicaid CAHPS Survey***[1]	***Medicare CAHPS Disenroll-ment Reasons Survey***	***Medicare CAHPS Disenroll-ment Assessment Survey***	***M+CO HOS***	***SHEP***
Total annual cost	$3,300,000	$5,000,000	$1,549,820	$3,000,000	$370,250	$5,787,540	$4,500,000
Approximate total surveys fielded	185,000	184,780	119,220	101,140	20,760	237,700	576,000
Cost per fielded survey	$17.84	$27.06	$13.00	$29.66	$15.18	$24.42	$7.81
Cost to plan per fielded survey	$0.00	$0.00	$13.00	$0.00	$0.00	$15.00	N/A
Approximate total surveys returned	131,400	144,800	35,855	45,000	11,485	173,500	388,800
Cost per completed survey	$25.12	$34.53	$39.89	$66.67	$0.00	$33.14	$11.57
Cost to plan per complet-ed survey	$0.00	$0.00	$39.89	$0.00	$0.00	$20.24	N/A

Source: Discussions with project leaders from CMS, NCQA, and the VA.
[1] Medicaid CAHPS costs vary considerably due to the variety of vendors and survey options selected by the different plans; therefore the estimated costs for these surveys are only an approximation.

For all of the above surveys except the HOS, the Medicaid CAHPS, and the SHEP, these costs are absorbed by CMS. The HOS is unique in that the M+C plans share some of the cost with CMS. The overall costs per survey are roughly comparable for each of the surveys listed above. For the Medicaid CAHPS, the cost per completed survey is much higher than the cost per fielded survey due to the much lower response rates obtained from this population.

The MFFS CAHPS, M+C CAHPS, and HOS surveys all use a mixed mode approach to survey administration (mail with telephone follow up of initial non-responders). The telephone component of these surveys adds significantly to their costs. Jones et al. (2001) found that overall, telephone administration of the Veterans SF-12 and SF-36 instruments was about 30% more expensive than mail administration, primarily due to labor costs. For the M+C CAHPS, the cost of data collection is about three times higher when telephone follow up is needed.

Table 19 summarizes the estimated respondent burden for several current Medicare surveys as well as the SHEP.

TABLE 19
ESTIMATED RESPONDENT BURDEN

Survey	***MFFS CAHPS Survey***	***M+C CAHPS Survey***	***Adult Medicaid CAHPS Survey***[1]	***Medicare CAHPS Disenroll-ment Reasons Survey***	***Medicare CAHPS Disenroll-ment Assessment Survey***	***M+CO HOS***	***SHEP***
Approximate total number of surveys returned	131,400	144,800	38,855	90,000	11,485	173,500	388,800
Approximate time needed to complete	20 minutes	20 minutes	20 minutes	23 minutes	27 minutes	20 minutes	35 minutes
Total respondent burden	43,800 hours	48,270 hours	12,950 hours	34,500 hours	5,170 hours	57,250 hours	176,400 hours

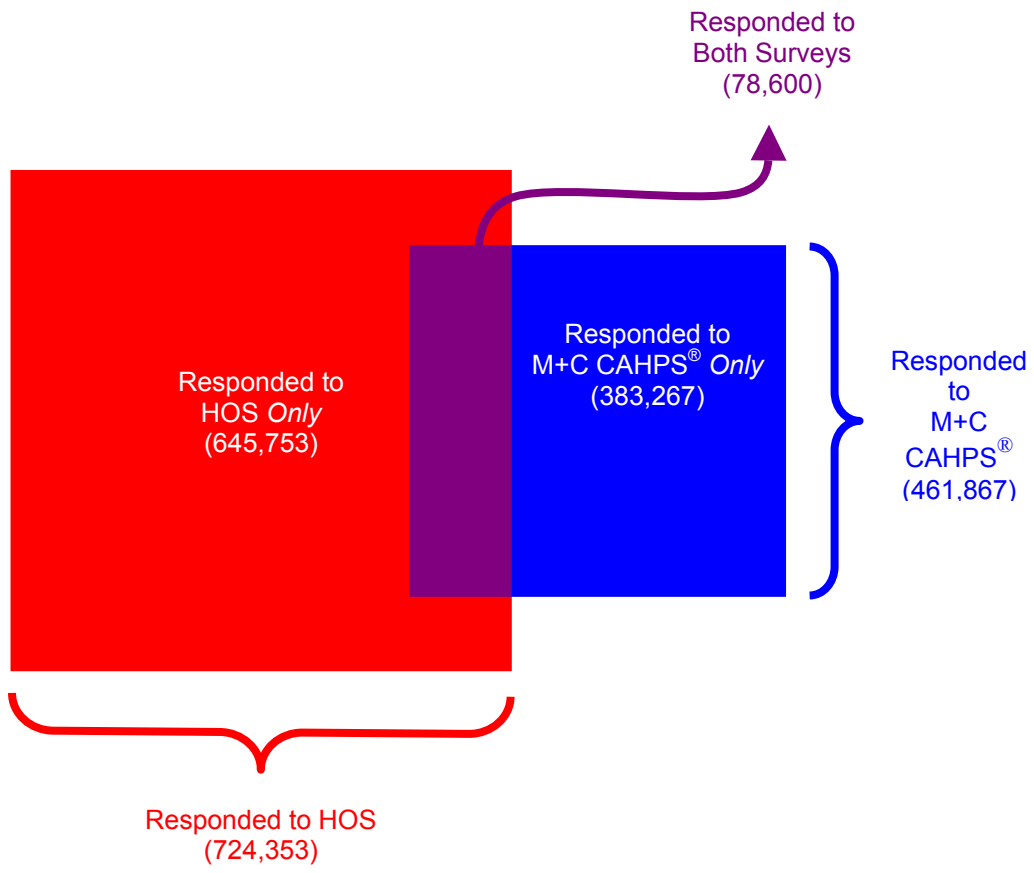
Sources: CMS Task Leaders, Federal Register Notices

[1] Medicaid CAHPS respondent burden varies considerably due to the variety of vendors and survey options selected by the different plans; therefore the estimated burden for these surveys is only an approximation.

Tables 18 and 19 suggest that an integrated survey will reduce annual survey costs by approximately $5,000,000 (the cost of the current M+C CAHPS survey), and will reduce annual respondent burden by approximately 48,270 hours (144,800 returned surveys multiplied by 20 minutes, or one third of an hour).

Many respondents, particularly those from smaller M+COs, complete more than one survey in a given year. This imposes an additional burden on some respondents. As discussed earlier, HOS survey respondents for the years 1998 through 2002, and M+C CAHPS® survey respondents for the years 2000 through 2002 were merged by HICN in order to determine the number of beneficiaries who responded to both surveys. Figure 1 shows the results of this analysis.

FIGURE 1
PATTERNS OF RESPONSE TO HOS AND M+C CAHPS® SURVEYS



Source: HOS and M+C CAHPS data merged by HICN.

This analysis reveals a substantial amount of overlap between the two surveys. A total of 78,600 unique beneficiaries responded at least once to both the HOS and the M+C CAHPS. These 78,600 dual responders represented 11% of the total responders to the HOS, and 17% of the total responders to the M+C CAHPS. Roughly two-thirds of these beneficiaries (54,061) responded to both an HOS and an M+C CAHPS survey within a six month time period. Within a two year time period, 2,424 unique beneficiaries responded to *two* HOS surveys and *two* M+C CAHPS surveys. Furthermore, some of these respondents may also have completed one of the disenrollment surveys, and/or possibly the Medicare Current Beneficiary Survey. Integrating the HOS and M+C CAHPS surveys will significantly reduce respondent burden in terms of the total number of surveys distributed as well as the number of beneficiaries receiving more than one survey in a given year.

Review of Key Issues

Clearly, integration of the two surveys would eliminate the administrative costs associated with one entire survey. However, these cost savings would be somewhat offset by the costs incurred by requiring contractors and vendors to reconfigure their procedures for questionnaire production, data collection, data cleaning, and reporting.

An integrated survey would also relieve a substantial amount of respondent burden by eliminating the burden due to one entire survey, and the elimination of the possibility that a given beneficiary will participate in both the HOS and M+C CAHPS surveys.

Based on cost data from the HOS and Veterans SF-12 and SF-36 instruments, eliminating the telephone mode of data collection would significantly reduce the costs of administering an integrated instrument. However, an analysis of the HOS protocol by NCQA revealed that certain respondents show a clear preference for the telephone (NCQA, 2002b). Respondents who are older, poorer, less educated, have more health problems or use a proxy to respond to the survey are more likely to respond by telephone. Consequently, it is important to retain the telephone component of the survey protocol in order to maximize the response rates from the more disadvantaged beneficiaries.

Both the SF-12v2™ and the Veterans version of the same instrument are feasible for use in an integrated survey. Both instruments yield the eight scale scores. And, responses to one instrument can be rescored in terms of the other. However, from a cost perspective, the Veterans version may be preferable because both the instrument itself and the associated missing data estimation technique are in the public domain, while a licensing fee may be required for scoring and missing data estimation for the SF-12v2™.

Merging of the HOS and M+C CAHPS surveys would also allow consolidation of the TEPs of the two projects, thereby eliminating the costs associated with one of these panels. Further savings would result from merging the educational conferences and user group meetings for the two surveys.

Preliminary Recommendations

1. *Retain the telephone mode of survey administration*. Even though telephone data collection substantially increases survey costs, it is a necessary tool for obtaining a robust response from certain key subgroups of beneficiaries.

2. *If fees are required to score the SF-12v2™, use the Veterans version of the SF-12 rather than the SF-12v2™*. The Veterans SF-12 and associated scoring algorithms are in the public domain.

3. *Merge the TEPs for the HOS and CAHPS surveys into a single TEP*.

4. *Also merge the educational conferences and user group meetings for the two surveys*.

5. *Share the cost of the integrated survey with the plans*. Since the plans already pay for a portion of the current HOS survey, and do not pay for the current M+C CAHPS survey, their total survey costs are likely to remain approximately the same. Furthermore, CMS will be able to replace the current M+C Disenrollment Assessment Survey with a disenrollee assessment for the integrated survey, for roughly the same cost.

VIII. DISSEMINATION OF THE RESULTS

Overview of Current Dissemination of Survey Results

This section describes the current dissemination of results from the HOS and M+C CAHPS surveys. The dissemination process is described for each of four main audiences: M+C plans, QIOs, researchers, and Medicare beneficiaries.

Medicare Health Outcomes Survey

Dissemination to M+C Plans

After the administration of each baseline cohort, a cohort specific baseline report is produced for each M+CO participating in the Medicare HOS. The baseline reports present SF-36® PCS and MCS scores, which are case mix adjusted using demographics, chronic medical conditions, and HOS study design variables. The baseline reports also provide plan, state, and HOS national information on health status indicators and demographics (including a comparison between respondents and non-respondents). Administrators at each participating M+CO have access to a plan specific baseline report that presents results for their plan, the state total, and the HOS national total.

A plan-specific performance measurement report is also produced for each M+CO participating in the Medicare HOS. The performance measurement reports document a health plan's ability to maintain or improve the physical and mental heath functioning of its Medicare beneficiaries over a two-year period of time. In the performance measurement report, change scores for both physical and mental health are calculated, and presented in terms of the percentage of the plan's beneficiaries who are better, the same, or worse than expected. These percentages are then case mix adjusted to correct for plan variations due to differences in the populations they serve, as described earlier. In addition to the performance measurement results, the reports also contain supplemental information summarizing response rates, health status indicators, and demographics.

Although the baseline and performance measurement reports were initially distributed to the plans in hard copy format, M+CO administrators now have access to the reports electronically via CMS' Health Plan Management System (HPMS). In addition to the reports, beneficiary level performance measurement data are also made available and distributed to the plans upon request.

HOS findings and strategies for their use have also been presented at national conferences. CMS has sponsored conferences for M+CO and QIO staff on this topic. In addition, HOS research findings have been presented at conferences sponsored by the American Health Quality Association, the American Association of Health Plans, the Medicaid Managed Care Congress, and the National Association for Healthcare Quality, as well as several other organizations.

The HOS Web site (http://www.cms.hhs.gov/surveys/hos/) is designed to provide current information on the progress of the HOS project, as well as house the full spectrum of HOS related data and reports.

Dissemination to QIOs

Each state's QIO receives a state specific baseline report and a state specific performance measurement report. These reports present results for all plans in the QIO's state(s), the state totals, and the HOS national total. Formerly distributed in hard copy format, these reports are now distributed electronically to the QIOs via the secure QNE application.

After baseline reports have been distributed to the QIOs, beneficiary level data are disseminated electronically to all participating QIOs. The electronic dissemination occurs through the QNE. Each QIO receives a beneficiary level SAS®[12] data set comprised of data from all beneficiaries surveyed in the QIO's state(s). In addition to the SAS® data set, each QIO is provided with a Data User's Guide that describes each of the data elements in the HOS baseline data files.

Similarly, after performance measurement reports have been distributed to the participating QIOs, beneficiary level performance measurement data are disseminated electronically to all participating QIOs via the QNE application, in the form of a SAS® data set. In addition to the data set, each QIO is provided with a Data User's Guide that describes each of the data elements in the HOS performance measurement data files.

As described above, HOS results and strategies for their use are presented at conferences designed for QIO and M+CO staff, as well as on the HOS Web site.

Dissemination to Health Care Administrators and Researchers

Included on the HOS Web site are the survey questionnaires used in each year of the HOS; an HOS Reports section that discusses the reporting process and provides links to sample reports; a Published Reports section that provides an overview of HOS research published in peer-reviewed journals; a section that provides answers to questions encountered on the HOS Information and Technical Support Telephone Line and e-mail address; and a section describing the HOS TEP that provides recommendations for advancing the science of the HOS instrument.

HOS data files are also available for download from the HOS Web site at no cost. HOS Public Use Files (PUFs) contain the majority of the survey items collected on the HOS instrument as well as selected additional administrative variables. Beneficiary identifying information is excluded. HOS PUFs are constructed to prevent the identification of any single beneficiary or plan, and only respondents to the survey are

[12] SAS® is a registered trademark of SAS Institute Inc.

included in the files. The PUF downloads can be accessed through the HOS Web site at (http://www.cms.hhs.gov/surveys/hos/hosdata.asp).

Research based on the HOS data has been published in academic peer-reviewed journals such as the *International Journal of Geriatric Psychiatry, Health Services Research, Health and Quality of Life Outcomes, International Journal for Quality in Health Care, Health Care Financing Review, Journal of Aging and Health*, and *Cancer*. A list of papers and publications based on the HOS data is available on the HOS Web site.

The Research Data Assistance Center (ResDAC) at the University of Minnesota is a CMS contractor that provides assistance to academic, government and non-profit researchers interested in using Medicare and/or Medicaid data. The mission of ResDAC is to assist researchers in locating and using Medicare and Medicaid data sources, including data from the HOS. ResDAC maintains a Web page (available at http://www.resdac.umn.edu) that provides descriptions of the various data sets available and instructions on how to obtain them.

Researchers interested in working with HOS data can obtain two types of data files from CMS by contacting ResDAC. These files contain more data elements than are available from PUFs, but the requestor must fulfill certain requirements in order to obtain the more sensitive information that they contain. Limited Data Sets (LDSs) are comprised of the entire national sample for a given cohort (including both respondents and non-respondents), and contain all of the HOS survey items. LDSs include plan identifiers as well as several additional variables describing plan characteristics. LDSs contain protected beneficiary-level health information such as date of birth; however, specific direct person identifiers (i.e. name and health insurance claim number) are not included. Research Identifiable Files (RIFs) contain all of the information contained in the LDSs, and also contain specific direct person identifiers (i.e., name and health insurance claim number). The requesting researcher must execute a signed Data Use Agreement with CMS in order to obtain LDS or RIF data.

Dissemination to Medicare Beneficiaries

An initial goal of the Medicare HOS project was to assist Medicare beneficiaries to choose a health plan by providing them with plan comparisons. The intent was to publicly report HOS plan results to consumers via the Medicare Web site's Medicare Compare tool (now called the Medicare Personal Plan Finder [MPPF]). In preparation for public reporting, an HOS Web-based Reporting Research Project was conducted by the CMS Center for Beneficiary Choices and QualityMetric in the years 2000 and 2001. This project included several focus groups and in-depth interviews with both beneficiaries and caregivers.

Results of the HOS Web-Based Reporting Research Project were mixed. Although the introduction of the project was met with some uncertainty, beneficiaries were able to interpret the meaning of HOS and accepted the results after they were shown graphs and data. Beneficiaries considered "same" health over a two-year period to be a positive

outcome for seniors, since health is more likely to decline as one ages. However, in general, beneficiaries did not think of the M+CO as accountable or responsible for the physical and emotional health outcomes tracked by the HOS. The patient and the physician were considered most responsible for physical and emotional health status and outcomes (Bayliss and Koepke, 2001).

As was standard for other first year HEDIS measures, the *Cohort I Performance Measurement* results were not publicly reported. The results of the Web-based reporting project, the lack of distinct variation among M+C plans on physical health in the first Performance Measurement results, and industry resistance all contributed to the postponement of the public reporting of data. Public reporting of HOS results still has not occurred as of the date of this report.

CAHPS Survey

Dissemination to M+C Plans

Each year a hard copy report is distributed to each participating M+CO. This report contains a high level summary of the M+CO's results, compares the plan in question to the average performance of plans in the same state, and outlines the plan's strengths as well as opportunities to improve care. A CD-ROM included with the hard copy print report provides additional detailed information for each of the CAHPS composite scores and CAHPS global ratings, as well as plan-level frequency distributions of each survey item.

The HPMS is also used to distribute results. Plans can use the HPMS to download copies of any of their CAHPS reports. In addition, HPMS provides an interactive method of viewing a plan's performance compared to other plans, as well as viewing survey results by age, race, education and other demographics. The interactive HPMS tool also allows plans to track their performance over time and compare their results to the MFFS CAHPS results for their region and state.

Dissemination to QIOs

Hard copies of the CAHPS reports are distributed to the QIOs. Beneficiary level data are also disseminated electronically to all participating QIOs via the QNE application. Each QIO receives a beneficiary level SAS® data set comprised of data from all beneficiaries surveyed in the QIO's state. Due to the sensitivity of the data, beneficiary level CAHPS data are not returned to the M+CO.

Dissemination to Medicaid and Commercial Plans

The National CAHPS® Benchmarking Database or NCBD (http://www.ncbd.cahps.org) is an important resource for providers and plans because it enables them to assess their performance relative to local, regional, and national benchmarks. Another valuable resource is the *CAHPS® Improvement Guide* (Edgman-Levitan et al., 2003), which is

designed to assist health plans and medical practices to assess their CAHPS performance and use the results to identify practical strategies for improving patients' experiences with care.

Dissemination to Health Care Administrators and Researchers

Several Medicare CAHPS data files are available for research purposes. Several years of M+C CAHPS, MFFS CAHPS, and Disenrollment CAHPS data files are available to researchers as RIFs from CMS through ResDAC. A signed Data Use Agreement with CMS is required to obtain RIF data files. The NCBD also offers respondent-level data files for the commercial and Medicaid populations. To obtain access to these files, researchers must submit an application for review by the NCBD Advisory Group. A description of the application process is available at http://ncbd.cahps.org/Products/Products.asp.

Periodically, CMS and AHRQ sponsor a National CAHPS Survey User Group Meeting, where researchers and plan staff can learn about interpretation of the results and current QI initiatives.

The literature on the CAHPS surveys is very extensive. A bibliography is available on the CAHPS®-SUN Web site at http://www.cahps-sun.org/References/References.asp#rart.

Dissemination to Beneficiaries

Beneficiaries and consumers are the primary audience for CAHPS survey results. A major objective of the CAHPS survey program is to provide beneficiaries with standardized data presented in a way that is both easy to understand and to use in choosing a health plan. Organizations that serve as consumer or beneficiary advocates, such as the AARP, are also an important audience for CAHPS results because they can publicize and interpret the information for their constituencies.

The Medicare Web site is specifically designed to help beneficiaries find answers to complex questions (www.medicare.gov/Choices/Overview.asp). This Web site provides a link to the MPPF that aids beneficiaries in finding specific plans within their local areas. The following seven M+C CAHPS measures are currently reported on the MPPF: getting needed care; getting care quickly; doctors who communicate well; ease of getting referrals to specialists; flu shot rate; overall rating of health care; and overall rating of health plan. In addition, the Medicare Web site also displays results from the MFFS CAHPS and information about beneficiaries' reasons for leaving M+COs.

Review of Key Issues

Plan staff as well as beneficiaries will be more likely to use results from the integrated survey if these results are available from a user-friendly Web-based interface that permits the user to query the data.

Since the M+C CAHPS results are currently being publicly reported, an integrated survey that contains both HOS and M+C CAHPS items will renew interest in public reporting of the HOS as well as the CAHPS results. It will be difficult to explain to stakeholder groups why one type of results should be reported while the other should not.

Public reporting of plan comparisons for specific subgroups of beneficiaries (for example, diabetics) can provide very useful information for consumers as well as M+CO administrators. However, the reporting of plan comparisons specifically for high functioning patients should be avoided as this could potentially stimulate adverse selection.

If the research design proposed earlier is adopted, then both cross-sectional and longitudinal results will be available. To portray plan performance accurately, both types of data need to be reported to plans and beneficiaries. Communicating both types of results to plans and beneficiaries in a clear and understandable fashion may pose some challenges.

Preliminary Recommendations

1. *Utilize the current HPMS site to provide access to data and results from the integrated survey to the plans*.

2. *Publicly report the results for the health status items as well as the CAHPS items*.

3. *Establish an interactive Web site for beneficiaries, based upon the current MPPF, where they can easily access survey results for specific plans and compare these plans' performance to that of other plans*. This site should be periodically tested with consumers to be sure that it is easy to use and understand.

4. *Report plan performance for specific subgroups of beneficiaries, such as beneficiaries with certain chronic medical conditions*. Beneficiaries as well as plans should find such information very helpful. Caution is advised regarding the reporting of plan comparisons specifically for high functioning patients—this could help to foster adverse selection.

5. *Include summaries of both the cross-sectional and longitudinal results in the summary reports*.

6. *Publicize the availability of the survey data to researchers at appropriate venues such as the annual Managed Health Care Congress*.

IX. USES OF THE RESULTS

Current Uses of the HOS Results

A CMS memo of June 22, 1999 (CMS, 1999) states CMS' expectations for how the QIOs will use the HOS data:

1. To identify QI opportunities at M+COs, and to initiate QI projects to address these opportunities;

2. To collaborate with M+COs that approach QIOs with their own proposals for using HOS data to conduct QI activities; and

3. To provide technical assistance to M+COs.

A number of the QIOs have been active in working with the M+COs in understanding and using the HOS results. For example, Florida Medical Quality Assurance, Inc. (FMQAI, the Florida QIO) has created summaries of the results from the self-reports of chronic medical conditions for each of the M+COs in the state. The M+COs find these data useful as a snapshot of the comorbidities in their beneficiary populations, and have used the information to pinpoint the conditions that would most benefit from disease management programs. The American Cancer Society and the Florida Department of Health have also requested summaries of FMQAI's results. OMPRO, the Oregon QIO, produces targeted summaries of the HOS data for the M+COs in the state. One analysis looked at the relationship between depression and chronic medical conditions in the nursing home setting. OMPRO used these results to help recruit several of the plans into an intervention project centered on increasing screenings and referrals for depression in nursing homes.

CMS' Quality Assessment and Performance Improvement (QAPI) program was created to help M+COs conduct QI projects and measure the impact of these projects. Past QAPI projects have used HOS data to measure the success of QI initiatives. The availability of baseline and follow up data on the same cohort of beneficiaries facilitates the ability to determine if the intervention(s) have had the intended effect. For example, McDonald, Ma, and Dulabone (2004) documented how an M+CO in Florida was able to improve the PCS scores of its CHF patients with a series of interventions over a four-year period. It should be noted that, for projects such as these, the individual beneficiaries are not identifiable to the plans.

Beginning in 1999, six QIOs and 16 M+COs participated in a National Pilot Project on Depression to improve the management of depression in the primary care setting. Depression was chosen as the study focus because it is a prevalent condition that responds well to treatment, but is often overlooked by the Primary Care Provider (PCP). The QIOs obtained utilization data from each participating M+CO in their respective

states. QIO staff then linked these data to demographic information, reports of chronic medical conditions, and mental status scores from the HOS. The results were used to generate a statistical profile of beneficiaries at high risk for depression. A separate risk profile was developed for each plan. The most commonly identified risk factors for depression were diabetes, heart disease, stroke, Medicaid status, and female gender age 75 or over.

For each M+CO, the key risk factors identified were used to classify all of the beneficiaries in the M+CO's population into low and high risk groups. Each M+CO then provided each of their PCPs with a list of the high risk beneficiaries in his or her caseload, along with clinical guidelines and treatment protocols for depression management. Using the HOS data to identify the high risk beneficiaries relieved the PCPs of the burden of screening all of their caseloads, and allowed them to focus their screening and management efforts on their high risk populations. The project increased the number of depression management activities deployed by the plans, and encouraged a number of the plans to adopt such best practices as depression registries and use of depression screening instruments (HSAG, 2002b).

In 2003, AHRQ contracted with the Center for Health Care Policy and Evaluation at UnitedHealthcare, to build on the statistical profiling strategy pioneered by the National Pilot Project on Depression. UnitedHealthcare researchers are using the HOS data and utilization data from 13 UnitedHealthcare plans to develop a claims-based statistical profile of beneficiaries at high risk for two-year declines in either PCS or MCS scores.

The HOS includes several questions that address current cancer symptoms and treatments. As a result, the HOS yields nationwide data on a large cohort of cancer survivors. The HOS has enabled researchers at the American Cancer Society to monitor the quality of life of four specific groups of cancer patients over time. They found that cancer survivors had lower PCS and MCS scores when compared to beneficiaries without cancer, and this relationship was independent of the effects of age. Of the types of cancer patients surveyed, lung carcinoma patients showed the lowest quality of life (Baker, Haffer, and Denniston, 2003).

Current Uses of the CAHPS Results

Results from the Medicare CAHPS surveys have been used to address a number of policy issues. As mentioned previously, the sampling units (geounits) for the MFFS CAHPS survey have been designed to correspond closely with the M+C plan market areas (in those parts of the country where M+C plans exist). This has facilitated comparisons between MFFS and M+C beneficiaries to identify differences in the care these two types of beneficiaries experience. This has also allowed CAHPS researchers to assess the extent of adverse selection in areas where both FFS and M+C provider arrangements exist.

As with the HOS data, some plans have used CAHPS results to demonstrate the impact of their QAPI projects to CMS. The *CAHPS® Improvement Guide* (Edgman-Levitan, 2003)

describes a number of strategies that plans have used to guide their QI efforts. Among these are:

1. The use of correlation and regression to pinpoint specific items that are most strongly related to global measures of satisfaction ("key driver analysis");
2. Identification of subgroups of beneficiaries that report unsatisfactory experiences of care;
3. Comparison of individual plan results to regional or national norms; and
4. Observing changes in scores over time.

The Medicare Managed Care Performance Assessment (PA) Project

As part of CMS' ongoing Strategy for Health Care Safety and Quality, the agency has developed a Performance Assessment (PA) Monitoring Tool (CMS, 2004) in order to provide the agency with a uniform approach to collecting outcomes-oriented measures for all of its Medicare managed care contractors. As part of this effort, CMS has created a Performance Assessment Summary Report that compares M+COs on a single combined score that aggregates four composite measures of M+CO performance: HEDIS®, CAHPS, HOS, and voluntary disenrollment rates. For each plan, each of these composites is converted to a percentile rank. The four percentile ranks are then averaged, and finally, this average percentile rank is also converted to a percentile. The effect of this calculation is to assign each of the four composite indicators equal weight in the final PA score. A major goal of the Strategy for Health Care Safety and Quality is to provide an incentive to M+COs who score highly on the current PA indicators by exempting them from certain review requirements.

Review of Key Issues

For both the HOS and CAHPS surveys, plan and QIO staff have repeatedly asked for guidance in how to translate their survey results into concrete QI action plans. If the two surveys are successfully integrated into a common instrument, but key decision makers do not use the results to create change, then the true value of an integrated survey will not have been realized.

To maximize the value that an integrated survey can provide, decision makers will need a clear "road map" for incorporating the results into their QI procedures. A good model for this road map is the *CAHPS® Improvement Guide* (Edgman-Levitan et al., 2003). This guide provides detailed examples and instructions for interpreting the survey results and using the results to design interventions. Numerous case studies provide concrete illustrations.

Periodic meetings of peer decision makers to share ideas, problems, and tactics are also needed. An emerging technology for involving large numbers of decision makers at minimal cost is "Webcasting." This technology allows individuals with Web access to participate in meetings from remote locations. The participants can engage in both audio communication and visual communication (by means of Web-based slide presentations).

Lumetra, the California QIO, has been designated as the Quality Improvement Organization Support Center (QIOSC) for M+C plans in the Seventh Scope of Work. As part of its portfolio of services to QIOs, Lumetra has begun sponsoring Webcasts on various topics of interest to the QIO community. Webcasting provides an ideal medium for transferring knowledge such as that contained in the *CAHPS® Improvement Guide* to large numbers of QIO and plan employees. And, it is financially feasible to conduct Webcasts at a much greater frequency than face-to-face meetings.

Due to the time and financial pressures under which M+COs operate, the development and initiation of QI activities often take a back seat to more short-term imperatives. The creation of incentives for using the results of an integrated survey is one strategy for moving this issue back on to the "front burner" for plans. For example, CMS has granted exemptions from certain review requirements to provide an incentive for high scores on the measures included in the PA Tool.

If, all of the current CAHPS global ratings and composite score items are retained in the integrated instrument, as recommended in this report, it will still be possible for NCQA to use these data elements for accreditation purposes.

The current M+C and MFFS CAHPS surveys are conducted in a parallel fashion, using similar administrative protocols and analytic strategies, in order to facilitate comparisons between the survey results for managed care and FFS beneficiaries. The availability of such data aids CMS in addressing important policy questions. Integrating the M+C CAHPS and HOS into a single questionnaire needs to be done in such a way that these M+C/FFS comparisons can still occur. Since this integration requires modifications to the M+C CAHPS, similar modifications to the MFFS CAHPS are indicated almost as a matter of course. And, by extension, the implementation of an integrated HOS/M+C CAHPS questionnaire points the way to the eventual implementation of an integrated HOS/MFFS CAHPS questionnaire.

Preliminary Recommendations

1. *Develop a User's Guide for the integrated survey, similar in scope and format to the CAHPS® Improvement Guide*. A comprehensive collection of strategies and actual case studies needs to be readily available as a response to health care decision makers requesting guidance in using the survey results for QI.

2. *Stress QI principles and activities at an annual conference for users of the integrated survey*. The conference should focus on practical QI applications rather than general research or policy issues.

3. *Develop and present Webcasts on the topic of using the integrated survey results for QI*. The M+C QIOSC can be a useful resource for this activity. Webcasting has the potential to greatly expand the community of integrated survey users.

4. *Tie incentives to performance on the integrated survey as a way to engender plan interest in using the survey results*. Two strategies can be used to accomplish this: 1) exempt plans from certain CMS requirements as a reward for good survey performance; and/or 2) make use of the survey results a specific CMS requirement (for example, a required task for the Eighth Scope of Work).

5. *Carefully review the pros and cons of integrating the HOS and MFFS CAHPS surveys at the same time that the HOS and M+C CAHPS surveys are integrated*.

X. CONCLUSIONS AND SUMMARY OF PRELIMINARY RECOMMENDATIONS

Feasibility of Integrating the HOS and M+C CAHPS Instruments

Based on the information reviewed in this report, integrating the HOS and M+C CAHPS surveys is clearly feasible. Indeed, several survey questionnaires that contain both experience of care questions and functional status questions are already in use: the MFFS CAHPS, the CAHPS Disenrollment Reasons Survey, and the SHEP survey. However, while feasible, integration of the HOS and M+C CAHPS surveys will require stakeholders to make a number of difficult decisions. These include:

- *Which HOS and CAHPS questions to retain and which to eliminate*. It is likely that different stakeholders will advocate for the retention of different groups of questions.

- *Whether to harmonize the two very different survey formats*. Attempting to develop a uniform format may interfere with comparisons to previously collected data, but not doing so may result in a less user-friendly instrument.

- *Which version of the SF-12 to utilize*. Both the SF-12®v2™and the Veterans SF-12 have advantages and disadvantages.

- *Whether to utilize a longitudinal or a cross-sectional survey research design*. The longitudinal design offers significant analytic advantages, but increases the complexity of the administrative protocols and associated costs.

- *Whether to publicly report the results*. Since CAHPS results are already publicly reported, it will be difficult to confine public reporting of integrated survey results to the CAHPS portion of the survey only. Also, if a longitudinal research design is used, it will be important to report comparisons of both baseline scores and change scores, which may prove confusing to the consumers of the information.

- *Whether to lobby CMS decision makers to incorporate use of the integrated survey results into one or more of the tasks for the Eighth Scope of Work, or some other CMS requirement*. The M+COs must contend with a very turbulent health care market, and are not likely to focus their attention on measures that are not tied to specific CMS initiatives.

Summary of Preliminary Recommendations

Below is a summary of our preliminary recommendations for addressing the key decisions described above:

1. *Substitute the SF-12 for the SF-36*. If there are licensing fees involved with scoring of the SF-12v2™, consider utilizing the Veterans Version in order to reduce costs.

2. *From the HOS, retain selected chronic medical conditions items*. These are useful to health care decision makers and can be used to develop plan norms for specific subgroups of beneficiaries. However, reduce their number by collapsing categories and eliminating the items that produce a high frequency of illogical response patterns. Also retain the ADL questions to provide data for frailty-based payment systems.

3. *From the M+C CAHPS survey, retain all items needed to create the CAHPS global rating and composite scores*. *From the M+C CAHPS survey, consider eliminating the gate questions*. *Also eliminate the majority of the appeals and grievances items*.

4. *Retain all demographic items needed for case mix adjustment*. When selecting adjuster variables for case mix adjustment, *use only those variables that make statistically important contributions to the overall adjustment model*.

5. *Use multiple vendors to manage the data collection process*. This will encourage market competition and help to lower vendor fees while insuring that CMS is able to compare differing vendor approaches to data collection and quality control.

6. *Continue to collect data by both telephone and mail*, to ensure representation in the survey results of those beneficiaries that are more likely to respond to a telephone survey.

7. *Administer the integrated survey on a longitudinal basis, and include disenrollees in the survey sample*. While more costly and analytically complex, this is the only way to address researcher and M+CO concerns about the biases that result from beneficiary attrition.

8. *Each year, develop a formal analytic plan for producing results that will be of use to individual M+COs and QIOs*.

9. *Merge the TEPs and the educational conferences for the HOS and M+C CAHPS surveys*. This will save money and foster the use of both experience of care and health outcome information by the same community of users.

10. *Share the cost of the integrated survey with the M+COs*. Since the plans already pay for a portion of the HOS, their total survey costs will remain approximately the same. And, the elimination of a separate M+C CAHPS Disenrollment Assessment survey means that CMS can support a disenrollee component for the integrated survey with little or no additional expenditure of funds.

11. *Add a page to the current HPMS Web site that provides data and results from the integrated survey to plans and QIOs* Provide the users with interactive tools for performing custom analyses of the data.

12. *Report results based on both cross-sectional and longitudinal samples*. There is a risk that both plan decision makers and consumers will find this confusing, but both types of results are crucial to a true understanding of plan performance.

13. *Publicly report results for both the health outcomes and experiences of care questions*. This will help motivate the M+COs to review and act upon the findings.

14. *Develop a User's Guide for plan administrators that clearly describes strategies for using the integrated survey results to drive QI initiatives*. This guide should include detailed case studies and contact persons for these case studies. The *CAHPS® Improvement Guide* provides a good model.

15. *Use Webcasts to further disseminate the strategies outlined in the guide described above*. This new technology will permit large numbers of M+CO and QIO staff to participate in a learning community that would not otherwise be cost effective to create.

16. *Incorporate use of the integrated survey results into a formal CMS requirement*. Due to limited resources, M+CO and QIO staff are strongly motivated to pursue only those measurement initiatives that are part of their contract obligations.

Next Steps

We recommend that CMS convene a task force of stakeholders and survey experts to decide on the feasibility of integrating the two surveys, review the preliminary recommendations set forth in this document, build a consensus regarding the decisions outlined above, and plan next steps.

If it is the task force's consensus that integration is feasible and desirable, then it should be the charge of the task force to develop a project timeline. This timeline should include such steps as 1) finalization of survey content and research design; 2) development of a sampling strategy; and 3) pilot testing of the instrument.

Version 2 of the HOS is currently under development. It may be some time before the HOS Version 2 and A-CAHPS survey questionnaires are ready for fielding. In light of

this, we suggest that the task force consider adding the SF-12 and the Activities of Daily Living (ADLs) questions to the current M+C CAHPS questionnaire for the next possible survey cycle. This would allow stakeholders to obtain at least some of the benefits of an integrated survey in the near future. One of these benefits would be the ability to compare health status information across the FFS and M+C populations.

XI. REFERENCES

Allen, H. (1998). Anticipating market demand: tracking enrollee satisfaction and health over time. *International Journal for Quality in Health Care*, 10(6), 521-530.

Baker, F., Haffer, S.C., & Denniston, M. (2003). Health-related quality of life of cancer and noncancer patients in Medicare managed care. *Cancer*, 97: 674-681.

Bayliss, M. & Koepke, C. (2001). HOS Consumer Reporting Project. Presentation at the Medicare Health Outcomes Survey Technical Expert Panel Meeting, May 8, 2001.

Bender, R.H., Lance, T.X., & Guess, L.L. (2003). Including disenrollees in CAHPS® managed health care plan assessment reporting. *Health Care Financing Review*, Fall, 25(1), 67-79.

Bierman, A.S., Bubolz, T.A., Fisher, E.S, & Wasson, J.H. (1999). How well does a single question about health predict the financial health of Medicare managed care plans? *Effective Clinical Practice*, 2(2), 56-62.

CAHPS® Survey Users Network. (2003). Instructions for Analyzing CAHPS® Data: Using the CAHPS Analysis Program Version 3.0. Available at http://www.cahps-sun.org/home/index.asp#kit.

CAHPS® Survey Users Network. (2004a). Frequently asked questions about the analysis of CAHPS® survey data. Internet address: http://www.cahps-sun.org/FAQs/analysis-rev.asp#2. (Last accessed March 8, 2004).

CAHPS® Survey Users Network. (2004b). Ambulatory CAHPS (A-CAHPS). Internet address: http://www.cahps-sun.org/Products/A-CAHPS/ACAHPSIntro.asp. (Last accessed June 15, 2004).

CAHPS® Survey Users Network. (2004c). The CAHPS Group Practice Survey. Internet address: http://www.cahps-sun.org/Products/GCAHPS/GCAHPSIntro.asp. (Last accessed June 15, 2004).

Cao, Z. & McGuire, T.G. (2003). Service-level selection by HMOs in Medicare. *Journal of Health Economics*, November, 22(6), 915-931.

Centers for Disease Control and Prevention (2003). About the BRFSS. Internet address: http://www.cdc.gov/brfss/about.htm. (Last accessed December 4, 2003.)

Centers for Medicare & Medicaid Services (CMS). (1999). OCSQ TOPS Control Number PRO-99-11. June 22, 1999. Internet address: http://qionet.sdps.org/sdps_memos/ sdps_memos_archive/SDPS_memos_99/99-11.pdf. (Last accessed July 6, 2004.)

Centers for Medicare & Medicaid Services (CMS). (2003). *Medicare Manual System Pub. 100-16 Managed Care.* Transmittal:32, October 3, 2003. Internet address: http://www.medicare.gov/publications/pubs/pdf/10050.pdf. (Last accessed October 23, 2003.)

Centers for Medicare & Medicaid Services (CMS). (2004). *Medicare Managed Care Organization Performance Assessment Reports: Industry Guide.* Internet address: http://hpmstest.fu.com/secure/pa/User_Guide.pdf?PI=060101050000. (Last accessed July 14, 2004.)

Cioffi, M.J., Cleary, P.D., and Zaslavsky, A.M. (2004). Impact of using individual level nonresponse weights with case-mix adjustment on the Medicare Managed Care CAHPS® 2002 Enrollee Survey Data. Internal memo, March 4, 2004.

Cleary, P.D., Zaslavsky, A.M., & Cioffi, M.J. (2004). Stability of reports about health care experiences: The Medicare Managed Care CAHPS survey. Manuscript submitted for publication.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences (second edition)*. Hillsdale, NJ: Erlbaum.

Colsher, P.L., & Wallace, R.B. (1989). Data quality and age: Health and psycho-behavioral correlates of item non-response and inconsistent responses. *Journal of Gerontology: Psychological Sciences*, 44:45-52.

Dillman, D.A., Sangster, R.L., Tarnai, J., et al. (1996). Understanding differences in people's answers to telephone and mail surveys. M.T. Braverman and J.K. Slater (Eds.), New Directions for Evaluation Series, 70 (Advances in Survey Research), 45-62. San Francisco: Jossey-Bass.

Edgman-Levitan, S., Shaller, D., McInnes, K., Joyce, R., Coltin, K.L., & Cleary, P.D. (2003). *The CAHPS® Improvement Guide*. Available at http://www.cahps-sun.org/References/CAHPSImprovementGuide.asp.

Elliott, M.N., Hambarsoomian, K., & Edwards, C. (2003). Analysis of Case mix Strategies and Recommendations for Medicare Fee-for-Service CAHPS. Unpublished report to CMS, April 2003.

Elliott, M.N., Swartz, R., Adams, J., Spritzer, K.L., & Hays, R.D. (2001). Case-mix adjustment of the National CAHPS® Benchmarking Data 1.0: A violation of model assumptions? *Health Services Research*, 36:3, 555-573.

Ellis, B.H., Shannon, E.D., Cox, J.C., Aiken, L., & Fowler, B.F. (2004). Chronic conditions: Results of the Medicare Health Outcomes Survey, 1998-2000. *Health Care Financing Review*, 25(4), Summer, 2004, 75-91.

Fowler, F.J., Gallagher, P.M., & Nederend, S. (1999). Comparing mail and telephone responses to the CAHPS® survey instrument. *Medical Care*, 37(3), MS41-MS49.

Gallagher, P.M., & Fowler, F.J. (1998). Size doesn't matter: Response rates of Medicaid enrollees to questionnaires of varying lengths. Presented at The Fourth CAHPS User Group Meeting, Baltimore, MD.

Gayle, V. (2003). Thinking about cross-sectional and longitudinal data. Internet address: http://www.lss.stir.ac.uk/longitudinal.pdf. (Last accessed December 2, 2003.)

Grace, S.C., Fowler, B.M., Li, F.F., & Shannon, E.D. (2003). Comparison of seniors in PACE and M+CO plans. Unpublished manuscript.

Greenwald, L.M., Levy, J.M., & Ingber, M.J. (2000). Favorable selection in the Medicare+Choice program: new evidence. *Health Care Financing Review*, Spring 2000, Volume 21, Number 3, 127-134.

Grijalva versus Shalala, 152 F.3d 1115 (Ninth Circuit, 1998).

Haffer, S.C. (2003). Using multiple survey vendors to collect health outcomes information: How accurate are the data? Health and Quality of Life Outcomes, 1, 1-6.

Harris-Kojetin, L.D., Fowler, F.J., Brown, J.A., Schnaier, J.A., & Sweeny, S.F. (1999). The use of cognitive testing to develop and evaluate CAHPS™ 1.0 core survey items. *Medical Care*, 37(3), 10-20.

Health Services Advisory Group (HSAG). (2002a). Performance Measurement Report. *Cohort III*, 2000-2002. Unpublished report to CMS.

Health Services Advisory Group (HSAG). (2002b). Medicare Health Outcomes Survey National Pilot Project on Depression: Final Report. September 30. Internet address: http://www.cms.hhs.gov/surveys/hos/download/HOS_NationalPilotProject_FinalReport.pdf. (Last accessed August 19, 2004.)

Health Services Advisory Group (HSAG). (2003). MCS Outlier Evaluation Study: *Cohort I* (1998-2000). Internet address: http://www.cms.hhs.gov/surveys/hos/download/HOSARC_MCSOutlier.pdf. (Last accessed, March 11, 2004).

Hwang, Y., Bierman, A.S., Haffer, S.C., & Wun, L.M. (2002). Weight adjustments in estimates for the 1999 Medicare Health Outcomes Survey. Proceedings of the American Statistical Association, Section on Health Policy Statistics, 1565-1570.

Iannacchione, V., Bernard, S., & Elliott, M. (2002). Medicare CAHPS® Fee-For-Service Survey. Presentation at Eighth National CAHPS® User Group Meeting. Nashville, Tennessee.

Iannacchione, V.G., & Campbell, L.N. (2003). Implementation of Medicare CAHPS Fee-For-Service Survey: Final Report for the 2003 Survey. Unpublished report, RTI International.

Jones, D., Kazis, L., Lee, A., Rogers, W., Skinner, K., Cassar, L., Wilson, N., & Hendricks, A. (2001). Health status assessments using the Veterans SF-12 and SF-36: Methods for evaluating outcomes in the Veterans Health Administration. *Journal of Ambulatory Care Management,* 2001, 24(3), 68-86.

Kane, R.L., Keckhafer, G., & Robst, J. (2002). Evaluation of the Evercare Demonstration Program: Final Report. Internal report to CMS, August, 2002. Internet address: http://www.cms.hhs.gov/researchers/demos/Evaluation%20of%20the%20Evercare%20Demonstration%20Program.pdf (last accessed on September 27, 2004).

Kazis, L.E. (2003). Personal communication.

Kazis, L.E., Miller, D.R., Clark, J.A., Skinner, K.M., Ren, X.S., Lee, A., Rogers, W., & Ware, J.E. (2004). Improving response choices of the SF-36 role functioning scales: results from the Veterans Health Study. *Journal of Ambulatory Care Management*, July, 27(3), 263-280.

Kosinski, M.K., Bayliss, M., Bjorner, J.B., & Ware, J.E. (2000). Improving estimates of SF-36® Health Survey scores for respondents with missing data. *Medical Outcomes Trust Monitor*, Fall; 5(1): 8-10.

Kwon, A., Bungay, K.M., Pei, Y., Rogers, W.H., Wilson, I.B., Zhou, Q., & Adler, D.A. (2003). Antidepressant use: Concordance between self-report and claims data. *Medical Care*, 41(3), 368-374.

Levy, P.S., & Lemeshow, S. (1999). *Sampling of populations: Methods and applications (third edition)*. John Wiley & Sons, Inc.

Lied, T.R., Sheingold, S.H., Landon, B.E., Shaul, J.A., & Cleary, P.D. (2003). Beneficiary reported experience and voluntary disenrollment in Medicare managed care. *Health Care Financing Review*, Fall, 25(1), 55-66.

Lock, S.L. (1999). Supreme Court state action ruling: Cause for worry. *Elder Law Forum*, Autumn, 1999, 5.

Longo, D.R., & Bohr, D. (1991). *Quantitative Methods in Quality Management: A Guide for Practitioners*. American Hospital Publishing, Inc.

Marshall, G.N., Hays, R.D., & Mazel, R. (1996). Health status and satisfaction with health care: Results from the Medical Outcomes Study. *Journal of Consulting and Clinical Psychology*. Apr; 64(2):380-390.

McCall, N., Khatutsky, G., Smith, K., & Pope, G.C. (2004). Estimation of Nonresponse Bias in Medicare Surveys. *Health Care Financing Review*, 25(4), Summer, 2004, 27-42.

McDonald, K., Ma, J., & Dulabone, E. (2004). Use of HOS Data in Florida. *Health Care Financing Review*, 25(4), Summer, 2004, 93-104.

Miller, L., Rogers, W.H., Spiro, A., & Kazis, L.E. (2004). Evaluation of disease status based on patient self-report in the Medicare Health Outcomes Survey: Using linked data from surveys and computerized medical data from the Veterans Health Administration. Internal report to CMS (in preparation).

Murgulo, M.S. (2002). Comparison of Medicare Risk HMO and FFS Enrollees. *Health Care Financing Review*, Fall 2002, Volume 24, Number 1, 177-185.

National CAHPS® Benchmarking Database (NCBD). (2002). CAHPS® Survey and Reporting Kit 2002. NCBD.

National Committee for Quality Assurance (NCQA). (2002a). *HEDIS® 2003, Volume 3: Specifications for Survey Measures.* Washington DC: NCQA Publication.

National Committee for Quality Assurance (NCQA). (2002b). HOS protocol analysis. Presented at the Medicare Health Outcomes Survey Technical Expert Panel Meeting. Washington, DC.

National Committee for Quality Assurance (NCQA). (2003). *HEDIS® 2003, Volume 6: Specifications for the Medicare Health Outcomes Survey*. Washington DC: NCQA Publication.

National Committee for Quality Assurance (NCQA). (2004). *HEDIS® 2004, Volume 6: Specifications for the Medicare Health Outcomes Survey*. Washington DC: NCQA Publication.

National Research Corporation (NRC). (2004). NRC's Dynamic Survey. Internet address: http://www.nationalresearch.com/dynamic.html. (Last accessed June 16, 2004).

Polonsky, W.H., Earles, J., Smith S, Pease, D.J., Macmillan, M., Christensen, R., Taylor T., Dickert, J. & Jackson, R.A. (2003.) Integrating medical management with diabetes self-management training. *Diabetes Care*, 26: 3048-3053.

Ren, X.S., Kazis, L., Lee, A., Rogers, W., & Pendergrass, S. (2001). Health status and satisfaction with health care: A longitudinal study among patients served by the Veterans Health Administration. *American Journal of Medical Quality*, Sep-Oct; 16(5): 166-173.

Rogers, W.H., Gandek, B., & Sinclair, S.J. (2004). Calculating Medicare Health Outcomes Survey Performance Measurement Results. Internet address: http://www.cms.hhs.gov/surveys/hos/hosresearch.asp#Calc.

Rogers, W.H., Qian, S., & Kazis, L. (2004). Imputing the physical and mental summary scores (PCS and MCS) for the MOS SF-36 and the Veterans SF-36 Health Survey in the presence of missing data. Internal report to CMS (in preparation).

Rogers, W.H., Ware, J.E., Gandek, B., & Bayliss, M.S. (2000). Examining longitudinal outcomes in the Medicare Health Outcomes Survey: Preliminary recommendations. Unpublished technical report.

Rubin, D.B. (1987). Multiple Imputation for Nonresponse in Surveys. Wiley, New York.

S. Res. 1, 108th Cong., 149 Cong. Rec. 8898 (2003) (enacted).

Schneider, E.C., Zaslavsky, A.M., Landon, B.E., Lied, T.R., Sheingold, S., & Cleary, P.D. (2001). National quality monitoring of Medicare health plans: The relationship between enrollees' reports and the quality of clinical care. *Medical Care*, 39(12): 1313-1325.

SF Community Web Site. (2003). AHRQ adopts use of SF-12v2™ Health Survey for National Medical Expenditure Panel Survey (MEPS). Internet address: http://www.sf-36.org/news/release_030915_AHRQ.shtml. (Last accessed December 5, 2003).

Shimada, S.L., Cioffi, M.J., Zaslavsky, A.M., & Cleary, P.D. (2003). Ordering of health assessment items in the Medicare CAHPS® Surveys. Manuscript submitted for publication.

Shimada, S.L., Zaslavsky, A.M., O'Malley, A.J., & Cleary, P.D. (2004). Market and beneficiary characteristics associated with enrollment in Medicare + Choice plans and fee for service. Manuscript submitted for publication.

Walsh, E.G., Nason, C.A., Moore, A., & Khatutsky, M.S. (2003). Pilot test of the Medicare Health Survey for PACE and Evercare: PACE pilot report. Internal Report to CMS.

Ware, J.E., Brook, R.H., Davies-Avery, A., Williams, K., Stewart, A.L., Rogers, W.H., et al. (1980). *Model of Health and Methodology*. Santa Monica, CA: RAND Corporation, R-1987/1-HEW.

Ware, J.E., Gandek, B., & Kosinski, M. (2002). Strategies for shortening the Medicare Health Outcomes Survey. Unpublished memorandum submitted to the HOS Technical Expert Panel, November 8, 2002.

Ware, J.E., & Kosinski, M. (2001). *SF-36® Physical and Mental Health Summary Scales: A Manual for Users of Version 1, Second Edition*. Lincoln, RI: QualityMetric, Inc.

Ware, J.E. Kosinski, M., & Keller, S.D.. (1994.) *SF-36® physical and mental health summary scales: A user's manual*. Boston: The Health Institute.

Ware, J.E., Kosinski, M., Turner-Bowker, D.M., & Gandek, B. (2002). *How to Score Version 2 of the SF-12® Health Survey*. QualityMetric, Inc.: Lincoln, RI & Health Assessment Lab: Boston, MA.

Ware, J.E., Snow, K.K., Kosinski, M., & Gandek, B. (1993). *SF-36® Health Survey Manual and Interpretation Guide*. Boston, MA: New England Medical Center, The Health Institute.

Zaslavsky, A.M., & Cleary, P.D. (2002). Dimensions of plan performance for sick and healthy members on the *Consumer Assessments of Health Plans Study 2.0 survey. Medical* Care, 40(10): 951-964.

Zaslavsky, A.M., Cleary, P.D., & Zaborski, L.B. (2000). Procedures for Design of MMC-CAHPS Sampling Areas. Unpublished report. Boston: Harvard Medical School.

Zaslavsky, A.M., Zaborski, M.S., & Cleary, P.D. (2000). Does the effect of respondent characteristics on consumer assessments vary across health plans? *Medical Care Research and Review*, September, 57(3), 379-394.

Zaslavsky, A.M., Zaborski, M.S., & Cleary, P.D. (2003). Plan, geographical, and temporal variation of consumer assessments of ambulatory health care. Unpublished manuscript.

Zaslavsky, A.M., Zaborski, M.S., Ding, L., Shaul, J.A., Cioffi, M.J., & Cleary, P.D. (2001). Adjusting performance measures to ensure equitable plan comparisons. *Health Care Financing Review*, 22(3), 109-126.

ATTACHMENT A

GLOSSARY OF ABBREVIATIONS

Abbreviation	*Reference*
A-CAHPS	Ambulatory CAHPS
AARP	Formerly the American Association of Retired Persons, now referred to solely by its initials
ADLs	Activities of Daily Living
AHRQ	Agency for Healthcare Research and Quality
BRFSS	Behavioral Risk Factor Surveillance System
CAHPS	Formerly, the Consumer Assessment of Health Plans survey, now referred to solely by its initials
CAHPS-SUN	CAHPS Survey Users Network
CCC	Children with Chronic Conditions (a version of the Medicaid CAHPS survey)
CMS	Centers for Medicare & Medicaid Services
COPD	Chronic Obstructive Pulmonary Disease
EDB	Enrollment Data Base
FMQAI	Florida Medical Quality Assurance, Inc.
G-CAHPS	CAHPS Group Practice Survey
HAL	Health Assessment Lab
HEDIS®	Health Plan Employer Data and Information Set
HICN	Health Insurance Claim Number
HOS	Medicare Health Outcomes Survey
HPMS	Health Plan Management System
HSAG	Health Services Advisory Group
IRT	Item Response Theory
LDS	Limited Data Set
M+C	Medicare + Choice
M+CO	Medicare + Choice Organization
MCO	Managed Care Organization
MCS	Mental Component Summary
MDE	Missing Data Estimation
MEPS	Medicare Expenditure Panel Survey
MFFS	Medicare Fee-For-Service
MHSPE	Medicare Health Survey for PACE and Evercare
MOT	Medical Outcomes Trust
MPPF	Medicare Personal Plan Finder
MRE	Modified Regression Estimate

ATTACHMENT A

GLOSSARY OF ABBREVIATIONS (CONTINUED)

Abbreviation	*Reference*
MSA	Metropolitan Statistical Area
NCBD	National CAHPS Benchmarking Database
NCQA	National Committee for Quality Assurance
NRC	National Research Corporation
OLAP	On Line Analytical Programming
PA Tool	CMS' Medicare Managed Care Performance Assessment Monitoring Tool
PACE	Program of All-Inclusive Care for the Elderly
PCP	Primary Care Provider
PCS	Physical Component Summary
PFFS	Private Fee-For-Service
PIP-DCG	Principal Inpatient Diagnostic Cost Group
PUF	Public Use File
QAPI	Quality Assessment and Performance Improvement
QI	Quality Improvement
QIO	Quality Improvement Organization
QIOSC	Quality Improvement Organization Support Center
QNE	QualityNet Exchange
ResDAC	Research Data Assistance Center
RIF	Research Identifiable File
SF-12®	Short-Form 12 Health Survey
SF-12v2™	Short-Form 12 Health Survey, Version 2
SF-36®	Short-Form 36 Health Survey
SF-36v2™	Short-Form 36 Health Survey, Version 2
SHEP	Survey of Health Experiences of Veterans
SHMO	Social Health Maintenance Organization
TEP	Technical Expert Panel
VA	Veterans Administration

ATTACHMENT B

LIST OF STAKEHOLDERS INTERVIEWED

Name	*Company*
Frank Baker, PhD	American Cancer Society
Shula Bernard, PhD	Research Triangle Institute
Arlene Bierman, MD	University of Toronto
Paul Cleary, PhD	Harvard Medical School
Kathy Coltin, MPH	Harvard Pilgrim Health Care
Chuck Darby, MA	Agency for Healthcare Research and Quality
Joyce Dubow, MUP	American Association of Retired Persons
Marc Elliott, PhD	RAND Corporation
Barb Gandek, MS	Health Assessment Lab
Liz Goldstein, PhD	Centers for Medicare & Medicaid Services
Chris Haffer, PhD	Centers for Medicare & Medicaid Services
Amy Heller, PhD	Centers for Medicare & Medicaid Services
Alan Hoffman, MHS	National Committee for Quality Assurance
Charles Humble, PhD	Veterans Administration
Vince Iannacchione, PhD	Research Triangle Institute
Lewis Kazis, ScD	Boston University; Veterans Administration
Nancy McCall, ScD	Research Triangle Institute
Kathie McDonald, RN, MPH	Florida Medical Quality Assurance, Inc.
Yelena Rosenfeld	OMPRO
Ted Sekscenski, MPH	Centers for Medicare & Medicaid Services
Carl Seratto, PhD	Kaiser Permanente
Dale Shaller, MPA	National CAHPS Benchmarking Database
Samantha Sheridan, MA	Westat
Chris Smith-Ritter, MPA	Centers for Medicare & Medicaid Services
Kris Spector, MA	National Committee for Quality Assurance
Edie Walsh, PhD	Research Triangle Institute
Alan Zaslavsky, PhD	Harvard Medical School

ATTACHMENT C
COMPARISON OF CONTENT BETWEEN THE 2003 HOS AND CAHPS SURVEYS

Brief Description	Exact Wording—M+C-CAHPS Survey	Exact Wording--MFFS CAHPS Survey	Exact Wording--HOS
General health	#72 In general, how would you rate your overall health now?	#56 In general, how would you rate your overall health now?	#1 In general would you say your health is?
Health compared to 1 year ago	#73 Compared to one year ago, how would you rate your health in general now?	#57 Compared to one year ago, how would you rate your health in general now?	#2 Compared to one year ago, how would you rate your health in general now?
Moderate activities	none	*#66 Does your health now limit you in doing moderate activities, such as moving a table, pushing a vacuum cleaner, bowling, or playing golf? If so, how much?*	#3b Does your health now limit you in these activities, If so, how much? Moderate activities, such as moving a table, pushing a vacuum cleaner, bowling or playing golf.
ADLs	none	#91 Because of a health or physical problem, do you have any difficulty doing the following activities? Bathing, Dressing, Eating, Getting in or out of chairs, Walking, Using the toilet	#12 Because of a health or physical problem, do you have any difficulty doing the following activities? Bathing, Dressing, Eating, Getting in or out of chairs, Walking, Using the toilet
Climb stairs	none	#67 Does your health now limit you in climbing several flights of stairs?	#3d Does your health now limit you in climbing several flights of stairs?
Accomplish less	none	#68 During the past 4 weeks, have you accomplished less than you would like as a result of your physical health?	#4b During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health? Accomplished less than you would like
Limited in kind of work	none	#69 During the past 4 weeks, were you limited in the kind of work or other regular daily activities you did as a result of your physical health?	#4c During the past 4 weeks, have you had any of the following problems with your work or other regular daily activities as a result of your physical health? Were limited in the kind of work or other activities
Pain limiting	none	#72 During the past 4 weeks, how much did pain interfere with your normal work, (including both work outside the home and housework)? *	#8 During the past 4 weeks, how much did pain interfere with your normal work (including both work outside the home and housework)?
Calm/peaceful	none	#73 How much of the time, during the past 4 weeks, have you felt calm and peaceful? *	#9d These questions are about how you feel and how things have been with you during the past 4 weeks. Have you felt calm and peaceful?
Energy	none	#74 How much of the time, during the past 4 weeks, did you have a lot of energy? *	#9e These questions are about how you feel and how things have been with you during the past 4 weeks. Did you have a lot of energy?
Downhearted/ blue	none	#75 How much of the time, during the past 4 weeks, have you felt downhearted and blue? *	#9f These questions are about how you feel and how things have been with you during the past 4 weeks. Have you felt downhearted and blue?
Physical health/emotional problems	#81 Do you have a physical or medical condition that seriously interferes with your independence, participation in the community, or quality of life?	#76 During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?	#10 During the past 4 weeks, how much of the time has your physical health or emotional problems interfered with your social activities (like visiting with friends, relatives, etc.)?
Smoking	#86 Do you now smoke every day, some days, or not at all?	#79 Do you smoke cigarettes every day, some days, or not at all?	#43 Do you now smoke every day, some days, or not at all?
Gender	#90 Are you male or female?	#82 Are you male or female?	#49 Are you male or female?
Hispanic	#92 Are you of Hispanic or Latino origin or descent?	#84 Are you of Hispanic or Latino origin or descent?	#50 Are you of Hispanic or Spanish family background?
Race	#93 What is your race? White, Black or African American, Asian, Native Hawaiian or other Pacific Islander, American Indian or Alaskan Native	#85 What is your race? White, Black or African American, Asian, Native Hawaiian or other Pacific Islander, American Indian or Alaskan Native	#51 How would you describe your race? American Indian or Alaskan Native, Asian or Pacific Islander, Black or African American, White, Another race or Multiracial)
Education	#91 What is the highest grade or level of school that you have completed?	#83 What is the highest grade or level of school that you have completed?	#53 What is the highest grade or level of school that you have completed?
Home	none	#90 Please mark the box that best describes your current living arrangement:	#54 Is the house or apartment you currently live in:
Proxy	#94 Did someone help you complete this survey? How did that person help you? How would you describe your relationship to the person who helped you complete this survey?	#89 Did someone help you complete this survey? How did that person help you? How would you describe your relationship to the person who helped you complete this survey?	#55 Who completed this survey form?

Green Exact match to HOS wording (* indicates different item format)
Red Similar to HOS wording

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

CONTENT	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Year of version examined	2003	2003	2003	2003	2003	2003	2003	2004
Basic content	Satisfaction, experiences of care, SF-12®, ASTQ, flu and pneumonia shots, demographics	Satisfaction, experiences of care, ASTQ, flu and pneumonia shots, demographics	Satisfaction, demographics	Reasons for disenrollment, experiences of care, SF-12®, demographics	Reasons for disenrollment, satisfaction, experiences with care, demographics	Comorbidities, ADLs, SF-36®, demographics, CDC healthy days, urinary incontinence	ADLs, SF-12® with minor revisions to questions	Satisfaction, experiences of care, SF-12 VA version, utilization, health behaviors
Are supplemental questions allowed?	No	No	Yes	No	No	No	No	No
Number of items	92	95	67 (adult) 74 (child)	80	92	99	34	99 (inpatient) 102 (outpatient)

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

ADMINIS-TRATIVE PROTOCOLS	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Administration period	Fall	Fall	Winter/Spring	Quarterly	Fall	Spring/Summer	Spring (PACE plans) Summer (others)	Monthly
Sampling units	Geounits	Plan sampling areas	Healthcare organizations	Plan sampling areas	Plan sampling areas	Contracts	Contracts	Inpatients: inpatient bed section; Outpatients: clinics
Administrative protocol (**Key**: PN=prenotifi-cation letter or postcard; SM=survey mailing; REM=reminder/ thank you; PH=phone follow up	PN (with option to complete by phone) 1st SM 1st REM 2nd SM PH (6-12 attempts); 3rd SM (to a sample of those without phone numbers, via priority mail)	PN 1st SM 1st REM 2nd SM PH 3rd SM (to non-respondents without phone numbers, via priority mail)	PN (optional) Mail only option: 1st SM 1st REM 2nd SM 2nd REM 3rd SM Mixed mode option: 1st SM 1st REM 2nd SM 2nd REM PH (3 attempts)	PN (with option to complete by phone) 1st SM 1st REM 2nd SM PH (12 attempts) 3rd SM (to non-respondents without phone numbers, via priority mail)	PN (with option to complete by phone) 1st SM 1st REM 2nd SM PH (12 attempts) 3rd SM (to non-respondents without phone numbers, via priority mail)	PN 1st SM 1st REM 2nd SM 2nd REM PH Phone follow up is also done for incomplete mail surveys	PN 1st SM 1st REM 2nd SM 2nd REM PH (6 attempts) Phone follow up of incomplete mail surveys to complete ADL questions	PN 1st SM REM 2nd SM

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

ADMINISTRATIVE PROTOCOLS (Continued)	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Dates plans are eligible to participate	Not applicable	Medicare contract in place no later than July 1 of previous year	The date the MCO becomes financially responsible for the enrollee or the date the MCO was notified	Medicare contract in place no later than January 1 of previous year	Medicare contract in place no later than July 1 of previous year	Medicare contract in place no later than January 1 of previous year (baseline); no later than January 1 three years ago (follow up)	Contract in place for one year	N/A
Member eligibility	Continuously enrolled for 6 months in MFFS; and resided in the 50 states, DC, or Puerto Rico; and did not have a representative payee.	Continuously enrolled for 6 months prior to July 1 of measurement year, and not institutionalized	Continuously enrolled the last 6 months of measurement year with no more than one 45 day gap	Members who disenrolled during previous quarter	Continuously enrolled for 6 months during the measurement year prior to disenrolling, and disenrolled from May thru July of the measurement year	Continuously enrolled for 6 months prior to the administration of the survey, may be institutionalized, ESRD patients are excluded	Enrolled for at least one month; the long term institutionalized are excluded from frailty calculation for PACE, and are excluded entirely for other plans	Inpatients: Must be admitted and discharged from same bed section (exception: Rehab). Outpatients: Receive outpatient care from either primary or specialty care clinic.

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

ADMINISTRATIVE PROTOCOLS (Continued)	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Participating organizations	Beneficiaries sampled from MFFS population in the CMS EDB	M+COs, PPOs, SHMOs, Continuing Cost Contracts.	MCOs with Medicaid contracts	M+COs, SHMOs, Continuing Cost Contracts Private FFS, Medicare Alternative Payment Demo I	M+COs, PPOs, SHMOs, Continuing Cost Contracts, Private FFS, Medicare Alternative Payment Demo I	M+COs, SHMOs, Continuing Cost Contracts, Private FFS, Medicare Alternative Payment Demo I	PACE plans, WI Partnership Program, MN Senior Health and Disability Health Options, MA Senior Care Organization	Inpatient facilities, outpatient clinics.
Vendors	University of Wisconsin (prime contractor) RTI (subcontractor for analysis), NCS Pierson (subcontractor for administration)	Westat (prime contractor); Harvard Medical School (subcontractor for case mix adjustment); DRC (subcontractor for administration); BearingPoint (subcontractor for report production).	27 HEDIS Vendors in 2003	University of Wisconsin (prime contractor), RTI (subcontractor)	Westat (prime contractor), CMS (subcontractor)	DataStat, DSS, GHS, Solucient, Synovate	RTI (prime contractor), New England Research Institute (NERI; subcontractor)	NRC (data collection)

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

ADMINISTRATIVE PROTOCOLS (Continued)	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Handling of mergers or acquisitions	N/A	Members who remain enrolled only	Plans may count members who switch due to a merger or acquisition as continuously enrolled	Members who remain enrolled only	Members who remain enrolled only	Members who remain enrolled only	Not available	N/A
Proxies Allowed?	Yes	Yes	No (for adults)	Yes	Yes	Yes	Yes	Yes
Foreign Language Versions	Spanish	Spanish	Spanish optional by organization and vendor. Other languages available: Vietnamese, Mandarin, Cambodian, and Korean	Spanish	Spanish	Spanish, Chinese	Spanish, Chinese; Initial survey sent out in the member's own language	Spanish (Puerto Rican)

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

ANALYTIC METHODS	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Case mix variables	Education, age, general health perception, mental health perception, proxy status (two variables: assisted with question-naire, completed questionnaire for respondent), and CMS region (as an interaction term with age and general health perception).	Education, age, general health perception, mental health perception, proxy status (two variables: assisted with question-naire, completed questionnaire for respondent), and CMS region (as an interaction term with age and general health perception).	Education, age, general health status (NCBD only)	None	Education, age, general health status, proxy status, mental health status	Performance Measurement *Death*: age, gender, race, education, marital status, income, home ownership, Medicaid status, 13 chronic cond-itions, 4 types of current cancer treatment, functional status, survey mode *MCS/PCS change scores*: age, gender, race, education, marital status, income, home ownership, Medicaid status Baseline age, gender, race, education, marital status, income, 13 chronic cond-itions, proxy status, survey mode, survey vendor, CMS region	None	Self-reported health status, age

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

ANALYTIC METHODS (Continued)	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Research design	Cross-sectional	Cross-sectional	Cross-sectional	Cross-sectional	Cross-sectional	Longitudinal	Longitudinal	Cross-sectional
Adjustment for non-response bias	Yes	Yes	No	Yes	Yes	No	No	Yes
Imputation of missing data	Yes	Yes	Yes	Yes	Yes	Yes	No	No

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

SAMPLING METHODS	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Sample size(s)	600	600	1,350 (adult) 1,650 (child)	388 across all four quarters, or all if less are available	600 max	1,000 (baseline) ≈500 (follow up)	All members, (100 min/plan, ≈7,000 total for PACE)	Variable (based on unit specific power calculations)
Target response rate	70%	60%	45% (adult) 55% (child)	68%	60%	70% (baseline) 80% (follow up)	65%	N/A
Most recent response rate	68%	83%	Approx. 30% (adult) Approx. 40% (child)	63%	60%	66% (baseline) 83% (follow up)	77%	60% (inpatient) 70% (outpatient)
Definition of a survey response	One or more items answered	One or more items answered	80% of items completed	At least one reason for leaving must be provided	One or more items answered	80% of items completed	All 6 ADL items answered	One or more items answered

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

COST AND BURDEN	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Approximate annual cost	$3,300,000	$5,000,000	$1,549,820	$5,500,000	$370,250	$5,787,540	Not Available	$4,500,000
Approximate total number of surveys fielded	185,000	184,870	119,220	202,300	20,760	237,700	Not available	576,000 (144,000 inpatient; 532,000 outpatient)
Approximate total number of surveys returned	131,400	144,800	35,855	90,000	11,485	173,500	Not available	388,800
Cost per fielded survey	$17.84	$27.06	$13.00	$27.35	$15.18	$24.42	Not available	$7.81
Cost per completed survey	$25.12	$34.53	$39.89	$61.46	$27.43	$33.14	Not available	$11.57
Cost to plan per fielded survey	N/A	$0.00	$13.00	$0.00	$0.00	$15.00	Not available	N/A
Cost to plan per completed survey	N/A	$0.00	$39.89	$0.00	$0.00	$20.24	Not available	N/A
Average time needed to complete survey	20 minutes	15 to 20 minutes	20 minutes	23 minutes	27 minutes	20 minutes	10 to 15 minutes	35 minutes

Note: Medicaid CAHPS costs vary considerably due to the variety of vendors and survey options selected by the different plans; therefore the estimated costs for these surveys are only an approximation.

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

DISSEMINATION OF RESULTS	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
M+COs	Comparison of FFS results with M+C results	Comparison of plan scores to statewide and national norms	Comparison of plan scores to regional, statewide, and national norms	Plan-specific reports of reasons reported for disenrollment	Plan scores based on both enrollees and disenrollees	Performance measurement data files, plan-specific reports of baseline and change scores, with comparisons to statewide and national norms; conferences; Web site.	PACE plans	N/A
QIOs	Comparison of FFS results with M+C results	Data files, comparison of plan scores to statewide and national norms	N/A	Plan-specific reports of reasons reported for disenrollment, with comparisons to statewide and CMS regional norms	Plan scores based on both enrollees and disenrollees	Performance measurement data files, state-specific reports of baseline and change scores, with comparisons to statewide and national norms; conferences; Web site.	N/A	N/A
Health Care Administrators and Researchers	Data files (through ResDAC); conferences, published research	Data files (through ResDAC); conferences, published research	Data files (through NCBD); conferences; published research	Data files (through ResDAC);	Data files (through ResDAC);	Data files (through ResDAC); published research	CMS administrators only	Report to Congress
Beneficiaries	Comparison of state to national averages, M+C to MFFS averages	Comparison of state to national averages, M+C to MFFS averages, reasons for disenrollment	Varies	Summary of reasons for disenrollment	Summary of reasons for disenrollment	Not currently reported to beneficiaries	N/A	Web site

ATTACHMENT D
COMPARISON OF EIGHT MAJOR SURVEYS

USES OF THE RESULTS	*MFFS CAHPS Survey*	*M+C CAHPS Survey*	*Medicaid CAHPS Survey*	*Medicare CAHPS Disenrollment Reasons Survey*	*Medicare CAHPS Disenrollment Assessment Survey*	*M+CO HOS*	*MHSPE*	*SHEP*
Policy Issues	Comparison of FFS and M+C responses; detection of adverse selection	Comparison of FFS and M+C responses; detection of adverse selection	Comparison of Medicaid with Medicare and commercial responses	Identification of key determinants of disenrollment	Comparison of enrollee and disenrollee responses	Determinants of changes in functional status; health status of cancer survivors	Adjustment of payments for frailty	Trends in satisfaction over time
Public Reporting	Comparison of FFS and M+C responses	Comparison of plan ratings to national and regional norms	Comparison of plan ratings to national and regional norms	Disenrollment rates and principal reasons given for disenrollment	Comparison of plan ratings to national and regional norms	Not currently reported to beneficiaries	Not currently reported to beneficiaries	Summary reports
Plan or Program Accountability	HPMS results	HPMS results	Various reports to state government health agencies	CMS MCO Performance Assessment Reports; HPMS results	CMS MCO Performance Assessment Reports; HPMS results	CMS MCO Performance Assessment Reports; HPMS results	Not currently used for performance monitoring	Performance monitoring
Quality Improvement	Analysis of key drivers; identification of beneficiary subgroups with low levels of satisfaction; benchmarking to national or regional norms; tracking of interventions over time	Analysis of key drivers; identification of beneficiary subgroups with low levels of satisfaction; benchmarking to national or regional norms; tracking of interventions over time	Analysis of key drivers; identification of beneficiary subgroups with low levels of satisfaction; benchmarking to national or regional norms; tracking of interventions over time	Guidance to QIOs in working with M+COs to reduce disenrollment	Guidance to QIOs in working with M+COs to reduce disenrollment	Statistical profiling of beneficiaries at risk for depression; claims-based profiling of beneficiaries at risk for declines in health status; targeting of disease manage-ment programs; identification of subgroups at high risk for poor health outcomes	Not currently used for quality improve-ment	Facility- and network-level QI initiatives