Estimation and Analysis of Non-response Bias in Medicare Surveys

Final Report

Prepared by

Nancy McCall, Sc.D.
Arthur Bonito, Ph.D.
Lily Trofimovich, M.S.
RTI International
Health, Social, and Economics Research
Research Triangle Park, NC 27709

ESTIMATION AND ANALYSIS OF NON-RESPONSE
BIAS IN MEDICARE SURVEYS

Nancy McCall, Sc.D.
Arthur Bonito, Ph.D.
Lily Trofimovich, M.S.

Scientific Reviewer: Gregory Pope, M.S.

Federal Project Officer: S. Chris Haffer

RTI International[*]

CMS Contract No. 500-00-0055

August 2004

---

[*] RTI International is a trade name of Research Triangle Institute.

CONTENTS

List of Tables

vii

EXECUTIVE SUMMARY

E.1   Organization of This Executive Summary

This executive summary presents the major findings from a study intended to assess the impact of survey non-response on estimates of important health status and consumer satisfaction measures derived from six surveys of Medicare beneficiaries conducted by the Centers for Medicare & Medicaid Services (CMS). In the next section of this executive summary, we discuss the purpose of this study and identify the individual surveys included in it. In the third section, we review the overall methodology used in this study. The same basic approach is taken to the analysis of each of the six surveys, although the absence or availability of selected items in the surveys accounted for slight differences in approach. In the fourth section, we discuss the overall results of the study. The final section highlights results from each of the six surveys of Medicare beneficiaries.

E.2   Purpose and Background

The goal of this study was to examine the potential degree of non-response bias in two major survey efforts that collect information from five different Medicare beneficiary populations: the Medicare Health Outcomes Survey (HOS) and the Consumer Assessment of Health Plans Survey (CAHPS®).[1] Both surveys are important instruments that have been designed and administered as a part of a larger CMS initiative to monitor and improve the quality of care provided to Medicare beneficiaries. The HOS is a HEDIS®[2] effectiveness-of-care survey that monitors the quality of care provided to Medicare beneficiaries by measuring changes in health status over time. The CAHPS® Survey is actually a family of surveys designed to collect information that may help beneficiaries make informed Medicare enrollment choices. The CAHPS® Survey provides a set of meaningful and reliable consumer-oriented measures on beneficiaries' experiences and satisfaction with health care. A variation of the CAHPS® is used to ascertain reasons why Medicare beneficiaries voluntarily disenroll from a Medicare+Choice (M+C) managed care plan.

Non-response may be a major threat to the validity of survey sample estimates obtained from surveys of Medicare beneficiaries. There are two possible types of non-response in surveys. One type occurs when a selected sample member does not respond at all to the survey. The second occurs when a selected sample member responds to some items but fails to answer all of them. Typically, the first type is referred to as survey non-response and the second as missing data or item non-response. Non-response bias is the systematic difference between the outcome scores for survey respondents and the (unknown) scores that would have been obtained if all subjects had completed the entire survey. The degree of bias is determined by two factors: (1) the difference in characteristics of interest (e.g., health status) between respondents and non-respondents, and (2) the non-response rate.

---

[1] CAHPS® is a registered trademark of the Agency for Healthcare Research and Quality (AHRQ).

[2] HEDIS® is a registered trademark of the National Committee for Quality Assurance (NCQA).

In this study, we examine the degree of potential non-response bias to six surveys of Medicare beneficiaries:

- 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey

- 2000 Cohort 3 Baseline Medicare Health Outcomes Survey

- 2000 CAHPS® M+C Enrollee Survey

- 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey

- 2000 CAHPS® M+C Disenrollment Assessment Survey

- 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only)

The unique contribution of this project to non-response bias research is the inclusion in our analyses of a claims-based measure of health status, the Principal Inpatient Diagnostic Cost Group (PIP-DCG) risk score, available for both survey respondents and non-respondents. To the extent that a claims-based measure of health status is a reasonable proxy for self-reported health status obtained in the HOS and the CAHPS® FFS Survey, we will be able to directly assess the probable degree of non-response bias for estimates of health status derived from survey respondents only. Similarly, if health status is correlated with measures of satisfaction and experiences with care that are ascertained in the CAHPS® Surveys, we will also be able to assess the probable degree of non-response bias for satisfaction estimates from the CAHPS® Surveys.

E.3   Methods and Approach

We conduct our assessment of probable degree of non-response bias in two steps. First, we evaluate how respondents differ from non-respondents by demographic, entitlement, and health status characteristics. If non-respondents are drastically different from respondents but represent only a negligibly small fraction of survey eligibles, then overall bias might not be significant.

Second, to understand how biased overall survey estimates become by using data from respondents only, we compare the PIP-DCG health status scores between eligibles and respondents, rather than between respondents and non-respondents. In doing so, we are assuming that the average health status estimate for the eligibles closely approximates the average health status estimate for the population from which the sample was drawn. Because the number of eligible beneficiaries is very large in most of the surveys, we consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the probable degree of bias that is present.

For each survey, we calculate response rates in total, by plan (state for FFS CAHPS[®][3]), and by demographic and enrollment characteristics of the beneficiaries. We estimate the scope of potential non-response bias by examining the differences in demographic characteristics, enrollment, health status, and service utilization between respondents and non-respondents. Demographic and enrollment information is obtained from CMS' Denominator file and Group Health Plan file. We use M+C inpatient encounter data and Medicare FFS claims data that are available for both respondents and non-respondents to assess differences in health status and service utilization. The claims and M+C inpatient encounter data provide information on the hospitalization rates, inpatient days, and diagnoses used in calculating the PIP-DCG risk adjustment score.

We also explore differences in survey-specific outcome measures by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. For the Medicare Health Outcomes Surveys and the CAHPS[®] FFS Survey, we report mean physical component summary (PCS) and mental component summary (MCS) health status scores as outcome measures. For the CAHPS[®] M+C Enrollee and Disenrollment Assessment Surveys and the CAHPS[®] FFS Survey, we report estimates of plan satisfaction and a "getting needed care" composite as the outcome measures. If the outcome measures vary by these characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias is increased.

While useful, univariate and bivariate analysis of response rates by beneficiary demographic, enrollment, health status, and utilization characteristics alone may result in misleading conclusions, because the characteristics are often highly correlated. Therefore, we conduct a multivariate logistic regression analysis of the likelihood of response that is estimated as a function of beneficiary characteristics to provide estimates of the *independent* effect of beneficiary characteristics on response. We estimate separate logistic regression models for each of the surveys and present the results as odds ratios.

We directly explore the degree of bias that may be present in estimates of health status and medical care usage by comparing means of these variables for respondents to those obtained for eligible beneficiaries, including non-respondents. We also report differences in mean PIP-DCG scores between respondents and survey eligibles, stratified by sociodemographic and medical care usage characteristics.

And last, we examine the differences between eligibles and respondents by plan response rate deciles to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles. Between eligibles and respondents, we compare average age; the proportion that are female, White, enrolled in Medicaid, and aged without end-stage renal disease (ESRD); average PIP-DCG risk score; number of hospitalizations; and number of inpatient days. Pairwise comparisons of differences in mean estimates between eligibles and respondents are made within deciles of response. Because the number of eligible

---

[3] Note that because it does not include M+C plans, the CAHPS® FFS Survey reports estimates of experience and satisfaction with care by state rather than by plan. Thus, in CAHPS® FFS, state is analogous to health plan, which is used to report information for the other five surveys.

beneficiaries is very large in each survey, we consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

E.4 Overall Findings

To produce comparability of results across the six surveys studied, we impose a uniform criterion for determining a response across these surveys. To do this, we use the response to a general health status question that appears in all six surveys as the definition of response. That question reads: *In general, would you say your health is excellent, very good, good, fair, or poor?* In addition, to keep the plans' or states' effects on the analysis in relative proportion, we apply a selection probability weight to the cases so that their contribution to the analysis is proportionate to the size of the plan (or state).

Weighted response rates based on the general health status question across the six surveys ranged from a low of 39 percent for the 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter) to a high of 86 percent for the 2000 Cohort 1 Follow-up HOS.[4] For five of the six surveys, the response rates based on the general health status question are quite similar to those observed for survey-specific definitions of response, differing by no more than 4 percentage points. The exception is the 2000 CAHPS® M+C Disenrollment Reasons Survey. It exhibits a 19 percentage point reduction in response rate from the survey-specific definition of a respondent (answering affirmatively to one of the preprinted reasons for disenrolling from the plan) when the general heath status question is used to define a respondent.

Despite differences in response rates, we find similar response patterns across the surveys for most key stratifying variables. The response rates of beneficiaries under age 65 and above age 84 are significantly lower than response rates for beneficiaries 65 to 74 years of age. The response rates for Blacks are significantly lower than those for Whites. Beneficiaries of Hispanic and American Indian race/ethnicity have response rates that are significantly higher than those for Whites in some but not all of the surveys. Asians have response rates quite close to those reported for Whites, with the exception of the 2000 CAHPS® FFS Survey. Beneficiaries dually enrolled in Medicare and Medicaid have significantly lower response rates for all six surveys. With the exception of the 2000 CAHPS® M+C Disenrollment Reasons Survey, beneficiaries entitled to Medicare because they are disabled (without ESRD) respond at a significantly lower rate than aged beneficiaries without ESRD. For all the surveys except the two CAHPS® M+C Disenrollment Surveys, we find that beneficiaries in poorer health status have a significantly lower response rate than those with average health status; response rates increase as health status improves, and response rates decline as number of hospitalizations in the year prior to survey increases.

We predict the likelihood of response as a function of sociodemographic and health status characteristics of all sampled beneficiaries. We estimate the logistic model weighted by the

---

[4] It should be noted that the response rate for the Cohort 1 Medicare HOS Follow-up Survey is higher than the other surveys because it represents the rate of response to a survey by the respondents to an earlier survey (the Baseline). The actual response rate of the sample of persons in the HOS Survey couplet would be the product of the response rates for the Baseline and Follow-up Surveys and lower than reported for the Follow-up alone.

inverse of the likelihood of the beneficiary being selected for survey. There are a number of general patterns that emerge across the six surveys. With the exception of the 2000 CAHPS® M+C Disenrollment Reasons Survey, beneficiaries under the age of 65 and age 85 and older are less likely to respond than beneficiaries age 65 to 74. Blacks are consistently less likely than Whites to respond to any of the surveys. There is no consistent pattern of response relative to Whites across the surveys observed for the other racial minorities/ethnicities. Beneficiaries dually enrolled in Medicare and Medicaid are consistently less likely to respond than beneficiaries not enrolled in Medicaid. Compared to beneficiaries with average health status, beneficiaries with a high level of health status are generally more likely to respond, while beneficiaries in poor health status are less likely to respond. These latter findings do not hold for the 2000 CAHPS® M+C Disenrollment Reasons Survey. And, for the two Medicare HOS cohorts, beneficiaries residing in long-term institutionalized settings are significantly less likely to respond relative to community-based beneficiaries.

To the extent that a claims-based measure of health status is a reasonable proxy for self-reported health status (a measure obtained in the Medicare HOS and the CAHPS® FFS Survey), we directly assess the degree of non-response bias in estimates of health status when they are based on respondents only. Similarly, if health status is correlated with measures of satisfaction and experiences with care, we indirectly assess the degree of non-response bias for satisfaction estimates from the CAHPS® Surveys based on respondents alone. We compare means of three variables—PIP-DCG risk score, number of hospitalizations, and number of inpatient days—for respondents to those obtained for eligible beneficiaries, including non-respondents. Differences in mean PIP-DCG risk scores between respondents and the entire eligible sample range from 1 percent to 4 percent, a very small difference overall. However, across the six studies, the eligibles always have higher risk scores than respondents, indicating that respondents are healthier than the sample as a whole. Further, there are virtually no differences in mean number of hospitalizations between eligibles and respondents and only modest differences in mean number of inpatient days. The pattern is very similar across all six surveys.

In summary, the degree of non-response bias at the survey level, for the range of response rates observed across the six surveys, is relatively modest. Mean PIP-DCG risk scores are 2 percent to 4 percent lower for respondents than for survey eligibles. However, we observe a general pattern that certain subpopulations consistently have low response rates and poor health status. Beneficiaries under the age of 65 and age 85 and older are less likely to respond than beneficiaries age 65 to 74. Blacks are consistently less likely than Whites to respond to any of the surveys. Beneficiaries dually enrolled in Medicare and Medicaid are consistently less likely to respond than beneficiaries not enrolled in Medicaid. And, for the two Medicare HOS cohorts, beneficiaries residing in long-term institutionalized settings are significantly less likely to respond relative to community-based beneficiaries. Further, beneficiaries without these characteristics but in poor health are also less likely to respond than beneficiaries of average health status.

Because many of these special populations represent a small proportion of all sampled beneficiaries within each of the surveys, the influence of their significantly lower rate of response and health status on the overall response rate and mean health status estimate at the survey level is muted. Of more concern would be within subpopulation analyses as well as analyses focusing on health plans with large proportions of these special populations. Using the

5

Cohort 3 Baseline Medicare HOS as an example, dual Medicare and Medicaid enrollees who responded had an average PIP-DCG risk score of 1.31 as compared to a PIP-DCG risk score of 1.40 for dual enrolled beneficiaries who did not respond. Because 40 percent of dual Medicare and Medicaid enrollees are non-respondents, the mean PIP-DCG risk score for dual enrollees would be underestimated by 4 percentage points.

The four Medicare CAHPS® surveys analyzed in this study adjust all survey-derived estimates for non-response, taking the general approach of using predicted response propensities to adjust initial design-based weights (the inverse of the selection probability) upward for respondents so that they represent both respondents and non-respondents. Sampling weights enable design-consistent estimation of population parameters by scaling the disproportionalities between the sample and the population using available demographic information for all sampled beneficiaries. Sampling weights are not constructed for the Medicare HOS.

Given the modest degree of non-response bias observed in this study among the Medicare surveys, efforts to enhance the current Medicare CAHPS® sampling weights by including measures of health status or medical service use as a proxy for health status do not appear warranted. Consideration could be given to the construction of selection probability weights to scale the disproportionalities between the sample and the population for the Medicare Health Outcomes Survey. As with the Medicare CAHPS® sampling weights, demographic information readily available would appear to be reasonable weighting variables. Care should be exercised when conducting analyses within subpopulations that experience high rates of non-response and exhibit significant differences between respondents and non-respondents in the analytic variable of interest. One needs to recognize that significant non-response bias could exist.

E.5   Highlights of Survey-Specific Results

E.5.1 2000 Cohort 3 Baseline Medicare HOS Survey

The Medicare Health Outcomes Survey (HOS) is a survey of Medicare beneficiaries enrolled in Medicare+Choice (M+C) managed care organizations. It is a self-administered mail survey with telephone follow-up. The sample is drawn in early spring of each year, and the survey concludes in early summer. The HOS instrument is based on the SF-36®[5] Health Survey, which asks respondents to rate their general health, ability to perform certain physical tasks, level of pain, and emotional state. Summary scales of physical and mental health, denoted as physical component summary (PCS) and mental component summary (MCS), respectively, are calculated using eight scales based on 36 questions. Both components are normed such that the mean score is 50 with a standard deviation of 10 points in the general U.S. population. It also includes items on health status, activities of daily living, specific medical conditions, and demographics. The HOS has been conducted since 1998.

Medicare beneficiaries defined as eligible for this survey include those who had been continuously enrolled in the same health plan for at least 6 months at the time of sample selection. One thousand eligible beneficiaries were sampled from each participating plan.

---

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

Beneficiaries with ESRD are excluded. Other individuals declared ineligible ex post include those reported deceased, unable to complete the survey because of language, with bad addresses or non-working telephone numbers, or not enrolled in the appropriate plan. There were 306 plans that participated in the 2000 Baseline Medicare M+C HOS and 291,221 eligible persons in the sample. A Spanish version of the survey was available. The definition of a completed survey, as specified by the HOS protocol, was 80 percent or more of the questions answered. The overall response rate was 72 percent.

Response rates vary by all enrollment and demographic characteristics and health status measures evaluated in this study. Compared to beneficiaries age 65 to 74, beneficiaries under the age of 65 and those 85 years of age or older have significantly lower rates of response. Aged beneficiaries with ESRD are significantly more likely to respond than aged beneficiaries without ESRD. Beneficiaries whose race/ethnicity is White or Asian are the most likely to respond to the HOS Baseline Survey, while Blacks and Hispanics are the least likely to respond. Females are modestly more likely than males to respond. Beneficiaries dually enrolled in Medicare and Medicaid are significantly less likely to respond. The response rate for beneficiaries residing in long-term institutionalized facilities is very low—only 28 percent. This is in contrast to the response rate of 71 percent for community-residing beneficiaries. Response rates decline as health status declines or numbers of hospitalizations increase.

Mean PCS scores differ substantially across respondents based on sociodemographic characteristics and levels of health status. Compared to beneficiaries age 65 to 74, beneficiaries of all other age groups have lower physical functioning. Compared to Whites, Blacks and Hispanics have lower mean PCS scores, while Asians have a higher mean PCS score. Men self-report a higher level of physical health than women. Beneficiaries dually enrolled in Medicare and Medicaid report a significantly lower level of physical health than Medicare-only beneficiaries. Beneficiaries residing in long-term care facilities and beneficiaries residing in the community who are deemed nursing home certifiable report significantly lower levels of physical functioning than community residents. Compared to beneficiaries with average health status, beneficiaries with better health status have higher average PCS scores and beneficiaries with worse health status have lower PCS scores. There is also an observed negative relationship with PCS scores and number of hospitalizations; as frequency of prior year hospitalizations increases, one observes declining average PIP-DCG scores.

Mean MCS scores also differ substantially across respondents based on sociodemographic characteristics and levels of health status. A similar pattern as that observed for PCS scores is observed for MCS scores for the demographic characteristics of age, race, gender, Medicaid enrollment, and institutionalized status and with number of hospitalizations.

Multivariate analysis of response propensity among the entire sample of eligible beneficiaries reveals that beneficiaries under the age of 65 and age 85 and older are less likely to respond than beneficiaries age 65 to 74. Beneficiaries of Asian descent are more likely to respond than White beneficiaries. In contrast, all other minority races are far less likely than White beneficiaries to respond to the M+C HOS. Men are less likely to respond than women. Beneficiaries dually enrolled in Medicare and Medicaid are less likely to respond than beneficiaries not enrolled in Medicaid. After controlling for health status, race, and age, beneficiaries with ESRD are significantly more likely than beneficiaries without ESRD to

respond to the HOS. The long-term institutionalized are less likely to respond than community-residing beneficiaries, while those that are nursing home certifiable are more likely to respond than community residents. Compared to beneficiaries with average health status, those with better health status are generally more likely to respond to the HOS. And, beneficiaries in poorer health status are less likely to respond. Last, the likelihood of response to the Medicare HOS declines as the number of hospitalizations experienced during the year prior to survey increases.

Differences in the mean health status between survey eligibles and respondents display a general trend in which health status estimates derived using the PIP-DCG risk score are often lower (better health) than those derived for survey eligibles across most major subpopulations of Medicare beneficiaries. This suggests that health status estimates derived from respondents only tend to modestly overestimate the health of M+C Medicare enrollees. Also, a comparison of the differences between eligibles and respondents by plan response rate deciles does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles.

E.5.2 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey

As discussed in Section E.5.1, the HOS is a survey administered to Medicare beneficiaries enrolled in M+C managed care plans. It is used to determine the change in health status over a 2-year time period. Two years after administration of the baseline survey, a follow-up survey using a similar instrument is administered. Medicare beneficiaries who responded to the baseline survey are eligible for the follow-up survey, if they remained enrolled in the same health plan as at the time of the baseline survey. Beneficiaries who die between completion of the baseline and follow-up surveys are considered respondents for purposes of measuring change in physical health status but excluded from analysis of change in emotional health. The other exclusion criteria used for the baseline survey are also in effect for the follow-up survey.

All health plans with Medicare contracts in place that administered the Cohort 1 Baseline Survey in 1998 were required to administer the Cohort 1 Follow-up Survey. There were 225 plans that participated in the follow-up survey, and the sample consisted of 88,129 eligible individuals. The HOS follow-up survey is administered by mail with telephone follow-up of mail non-respondents during the same time frame as the baseline survey. As with the baseline survey and for purposes of this analysis, a respondent is a beneficiary for whom a PCS or MCS score could be calculated. The overall response rate was 85 percent for the Cohort 1 Follow-up Survey.

Across all analyses, the patterns observed in the Cohort 1 Follow-up Survey are very similar to those reported for the Cohort 3 Baseline Survey. Response rates vary by all enrollment and demographic characteristics and health status measures evaluated in this survey and in the same general patterns observed for the baseline survey. As previously observed, mean PCS scores differ substantially across respondents based on sociodemographic characteristics and levels of health status. Further, the direction and magnitude of the odds ratios from the multivariate analysis of response propensity among the entire sample of eligible beneficiaries for the included beneficiary-level variables are consistent with the descriptive comparisons between respondents and non-respondents and with the direction of effect observed in the multivariate

8

modeling of response for the Cohort 3 Baseline Survey. However, the magnitude of effect of the predictor variables collectively is lower than observed in the Cohort 3 Baseline Survey.

A comparison of PIP-DCG risk scores for respondents and the entire sample of eligible beneficiaries showed only a 1 percent difference between them, suggesting that respondents, on average, have a modestly higher level of health status than the entire surveyed population. This was supported by a similar analysis by level of beneficiary characteristics that found differences in health status risk scores between respondents (healthier) and the entire sample for persons 75 to 84 years of age, Whites, females, community dwellers, and enrollees in Medicare only. And, as with the analysis of the Cohort 3 Baseline HOS, a comparison of the differences between eligibles and respondents by plan response rate deciles does not suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles.

E.5.3 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey

The CAHPS® Medicare+Choice (M+C) Enrollee Survey is an annual survey conducted by CMS to assess the experience of Medicare beneficiaries enrolled in Medicare+Choice organizations (MCOs). It was developed to capture critical information about the enrollees' perception of quality of care in MCOs. The sample for the survey was designed to allow CAHPS® outcomes to be compared between plans, as well as with Original Fee-for-Service (FFS) Medicare. Each Medicare managed care plan comprises a reporting unit. Within each reporting unit, a random sample of plan enrollees was selected. Eligible plans for the 2000 survey administration included all MCOs and continuing cost contracts with contracts in effect as of July 1, 1999. The survey included 292 plans.

To be eligible for sample selection, beneficiaries had to have been enrolled in the selected MCO continuously for at least 6 months and at the time of selection could not have been institutionalized. Approximately 600 beneficiaries were sampled from each MCO in the survey. Beneficiaries declared ineligible ex post were those reported to be deceased or institutionalized, those with bad addresses and telephone numbers, and beneficiaries who had switched MCOs. The 2000 sample consisted of 216,919 Medicare eligible beneficiaries. The survey was conducted by mail with a telephone follow-up of mail non-respondents. A Spanish version of the survey was available. A questionnaire was considered complete if 10 or more key questions were answered. The survey response rate for the 2000 CAHPS® M+C Enrollee Survey was 83 percent.

Response rates vary by all enrollment and demographic characteristics and health status measures evaluated in this survey and in the same general patterns observed for the two Medicare Health Outcomes Surveys. Comparison of response rates for categories of demographic, enrollment, health risk status, and utilization characteristics shows that they are significantly lower for beneficiaries who are Black rather than White, under age 65 and above 74 rather than age 65 to 74, also enrolled in Medicaid rather than enrolled in Medicare alone, hospitalized in the year prior to survey rather than those not hospitalized in the prior year, with PIP-DCG risk scores indicating poorer health status, and disabled and aged beneficiaries with ESRD as compared to Medicare beneficiaries enrolled in Medicare due to age alone.

For the CAHPS® M+C Enrollee Survey, we display estimates of the average rating of respondents to two measures of satisfaction—satisfaction with the health plan and satisfaction

with getting care when needed—as the outcome measures. If outcome measures, such as satisfaction with care, vary by demographic characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias existing increases. There is considerable variation in self-reported satisfaction with care received from the beneficiaries' health plans. Beneficiaries under age 65, or those entitled to Medicare because of disability, reported less satisfaction with their health plan than beneficiaries age 65 and older. A similar pattern is observed when evaluating reason for Medicare entitlement. Beneficiaries dually enrolled in Medicare and Medicaid are modestly less satisfied with their health plans than beneficiaries not dually enrolled. Beneficiaries in the risk score quintiles indicating the best health status (lowest PIP-DCG scores) have lower rates of satisfaction than beneficiaries in average health status. There is virtually no variation in self-reported satisfaction with getting care when needed across any of the demographic or health status categories.

The direction and magnitude of the odds ratios from our multivariate logistic regression analysis of response are consistent with the descriptive comparisons between respondents and non-respondents. Beneficiaries under the age of 65 and age 75 and older are roughly 20 percent to 45 percent less likely to respond to the CAHPS® M+C Enrollee Survey than beneficiaries age 65 to 74. Beneficiaries of American Indian descent and Hispanic and Asian beneficiaries are more likely to respond than Whites. Blacks are far less likely than White beneficiaries to respond to the survey. Men are less likely to respond than women. Beneficiaries dually enrolled in Medicare and Medicaid are almost 50 percent less likely to respond than beneficiaries not also enrolled in Medicaid. Compared to beneficiaries with an average health status score, those with a higher level of health status are generally more likely to respond to the survey. Beneficiaries with poorer health status or with ESRD are less likely to respond than those with average health status or without ESRD, respectively.

A comparison of PIP-DCG risk scores for respondents and the entire sample of eligible beneficiaries shows a small (2 percent) difference between them, with respondents having lower scores, indicating better health status. Mean number of hospitalizations and mean number of inpatient days are modestly lower for respondents than for survey eligibles. Differences in the mean health status between survey eligibles and respondents display a general trend in which health status estimates derived using the PIP-DCG risk score are often lower (better health) than those derived for survey eligibles across some of the major subpopulations of Medicare beneficiaries. This suggests that health status estimates derived from respondents only tend to modestly overestimate the health of Medicare M+C enrollees. And, this overestimate tends to be for several of the healthier subpopulations (e.g, aged without ESRD and not dually enrolled in Medicare and Medicaid).

A comparison of the differences between eligibles and respondents by plan response rate deciles does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles. In fact, there are limited observed differences between eligibles and respondents for the health plans with the lowest level of response. Of more interest is the difference in the characteristics of eligible beneficiaries in the health plans with the lowest response rates. These health plans tend to have considerably larger proportions of non-White beneficiaries as well as beneficiaries dually enrolled in Medicare and Medicaid and beneficiaries with ESRD.

E.5.4 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey

All Medicare managed care plans that have contracts with physicians or physician groups are required to conduct annual enrollment and disenrollment surveys and report the results to CMS. Legislation requires that CMS make consumer assessment information on the plans available to Medicare beneficiaries to assist them in making plan choice decisions regarding participation in the program. The enrollment survey requirement is satisfied by the annual nationwide administration of the Medicare CAHPS® M+C Enrollee Survey. However, the Enrollee Survey includes only those who have been continuously enrolled in a plan for 6 months or more at the time of the survey and excludes beneficiaries who voluntarily disenrolled from the plan during the previous year. Hence, there was a need to separately survey plan disenrollees and add their responses to those of enrollees.

The CAHPS® M+C Disenrollment Survey has been conducted annually since 2000. It consists of two different component surveys. The Assessment Survey component is intended to collect beneficiaries' assessment of their experiences while they were in the managed care plan. Since they were in essence to be added together, the Assessment Survey component of the CAHPS® M+C Disenrollment Survey was created to be virtually identical in content to the CAHPS® M+C Enrollee Survey. The survey is administered by mail with telephone follow-up of non-respondents. There is a Spanish language version.

The sample for the Assessment Survey was selected at about the same time and in the same proportions in each health plan as that used in the Enrollee Survey to minimize design effects in the combined survey estimates. Data collection activities for the 2000 Assessment Survey were conducted between October 2000 and February 2001. Beneficiaries who had voluntarily left their plan between May and July 2000 were eligible to be included in the Assessment Survey sample if they had 6 months of continuous enrollment in the plan. The 2000 CAHPS® M+C Disenrollment Assessment Survey sample consisted of 22,272 eligible Medicare beneficiaries from 281 managed care plans. Deceased disenrollees were removed from the sampling frame before the sample was selected. Returns from the survey process resulted in further exclusions of persons considered ineligible because of death, institutionalization, or being erroneously categorized as disenrolled. About 55 percent of eligibles completed a questionnaire. A questionnaire was considered complete if the respondent answered at least one question other than screening questions.

Response rates vary by all enrollment and demographic characteristics and health status measures evaluated in this survey and in the same general patterns observed for the CAHPS® M+C Enrollee Survey. The response rates of beneficiaries under age 65 and above 74 years are significantly lower than those for beneficiaries 65 to 74 years of age. This finding is particularly true for the youngest and oldest age groups. The response rates for all of the racial/ethnic groups except Asians are significantly lower than those for Whites. Beneficiaries not also enrolled in Medicaid have a significantly higher response rate than beneficiaries dually enrolled in Medicare and Medicaid. Beneficiaries entitled to Medicare because they are disabled (with or without ESRD) or aged with ESRD responded at a significantly lower rate than aged beneficiaries without ESRD. Beneficiaries with a PIP-DCG score lower than the category containing 1.00 (in better health because they are below the mean) have significantly higher response rates than those in the category containing 1.00. And, beneficiaries with two or more hospital discharges

have significantly lower response rates than those who had not been hospitalized at all in the prior year.

As observed in the CAHPS® M+C Enrollee Survey, there is considerable variation in self-reported satisfaction with care received from the beneficiaries' health plans. Beneficiaries under age 65 report less satisfaction with their health plan than beneficiaries age 65 to 74, while beneficiaries age 75 to 84 are more satisfied. Female beneficiaries are more satisfied with their plan than males. Beneficiaries dually enrolled in Medicare and Medicaid are modestly less satisfied with their health plan than beneficiaries not dually enrolled. Disabled beneficiaries (without ESRD) are less satisfied with their plan than the aged (without ESRD). Beneficiaries in the risk score quintiles indicating the best health status and the worst health status have lower rates of satisfaction than beneficiaries in average health status. And, beneficiaries with three or more hospital stays are less satisfied with their plan than persons who had none. There is less variation in self-reported satisfaction with getting care when needed across most of the demographic, enrollment, health status, and utilization categories.

The direction and magnitude of the odds ratios from our multivariate logistic regression analysis of response are consistent with the descriptive comparisons between respondents and non-respondents. Beneficiaries under the age of 65 and age 85 and older are less likely to respond than beneficiaries age 65 to 74. American Indian and Hispanic beneficiaries are significantly more likely to respond than White beneficiaries. In contrast, Blacks and Asians are less likely to respond than White beneficiaries. Beneficiaries dually enrolled in Medicare and Medicaid are less likely to respond than beneficiaries not also enrolled in Medicaid. Beneficiaries with ESRD are significantly less likely to respond than beneficiaries without ESRD. Compared to beneficiaries with an average health status score, those in both better and poorer health status are more likely to respond. Beneficiaries who were hospitalized two or three times during the year prior to the survey are less likely to respond than persons with no hospital stays.

A comparison of average PIP-DCG risk scores for respondents and the entire sample of eligible beneficiaries shows a 4 percent overall difference between them, with respondents having lower scores, indicating better health status. This was supported by a similar analysis by level of beneficiary characteristics that found statistically significant differences in risk score between respondents (healthier) and the entire sample for persons 65 years of age and older; Whites, Blacks, and Hispanics; females and males; aged; and those not enrolled in Medicaid.

We examine the differences between eligibles and respondents by plan response rate deciles to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles. There are limited observed differences between eligibles and respondents at any level of plan response; however, the characteristics of survey eligibles (as well as respondents) do appear to differ across the deciles of plan response. There appears to be a slight trend for average age to increase as the plan response rate increases. Ignoring the two sets of deciles with very few plans and beneficiaries, the same pattern exists for the percent White and the percent of beneficiaries who are enrolled in Medicare because they are elderly without ESRD. The proportion of beneficiaries dually enrolled in Medicare and Medicaid seems to decrease as the plan response rate increases.

E.5.5 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Reasons Survey

The 2000 CAHPS® M+C Disenrollment Reasons Survey is the second component of the CAHPS® M+C Disenrollment Survey. The Reasons Survey was conducted for the first time in the summer of 2000 with a sample of Medicare beneficiaries who voluntarily left their managed care plan during 2000. The Reasons Survey is conducted quarterly, as opposed to once a year like the other Medicare CAHPS® Surveys. Although data collection and processing are implemented on a quarterly basis, the survey results are reported annually. Our analysis is limited to the 3rd quarter (July to September 2000) sample of eligible beneficiaries.

The survey is conducted by mail with telephone follow-up of mail non-respondents. The questionnaire for this survey was designed to collect information on the reasons that members left their former Medicare managed care plan. A Spanish version of the questionnaire was available. There were 12,659 beneficiaries eligible for this CAHPS® M+C Disenrollment Reasons Survey. To be eligible, beneficiaries had to have been contacted and indicated that they voluntarily disenrolled from their health plan. Beneficiaries unable to be contacted or those without good addresses or phone numbers were considered ineligible. In the 3rd quarter 2000 CAHPS® M+C Disenrollment Reasons Survey, 7,395 beneficiaries were deemed respondents, for a response rate of 58 percent.

In comparison to all the other surveys included in this report, fewer subpopulations exhibited differential rates of response. The response rates of beneficiaries above 74 years of age are significantly lower than those for beneficiaries age 65 to 74. The response rates for Blacks and Hispanics are significantly lower than those for Whites. Beneficiaries dually enrolled in Medicare and Medicaid have significantly lower response rates than Medicare beneficiaries not dually enrolled in Medicaid. Beneficiaries with a PIP-DCG score in the lowest three categories (best health status) have significantly higher response rates than those in the category reflecting average health status. And, beneficiaries who had two hospital discharges have a significantly lower response rate than those who had not been hospitalized at all in the prior year.

The direction and magnitude of the odds ratios from our multivariate logistic regression analysis of response are consistent with the descriptive comparisons between respondents and non-respondents. Beneficiaries age 75 and older are less likely to respond than beneficiaries age 65 to 74. Beneficiaries who are Black, Hispanic, and Asian are less likely to respond than White beneficiaries. Men are more likely to respond than women. Beneficiaries dually enrolled in Medicare and Medicaid are 35 percent less likely to respond than beneficiaries not enrolled in Medicaid. Compared to beneficiaries with an average health status score, those in both better and poorer health are less likely to respond. Interestingly, beneficiaries who were hospitalized one or three or more times during the year prior to survey have a higher likelihood of responding than beneficiaries who did not have had any hospitalizations. In contrast, beneficiaries who were hospitalized two times in the year prior to survey are less likely to respond.

A comparison of average PIP-DCG risk scores for respondents and the entire sample of eligible beneficiaries shows a 3 percent difference, with respondents having lower scores (healthier). This was supported by a similar analysis by level of beneficiary characteristics that found differences in average risk scores between respondents (healthier) and the entire sample for beneficiaries 85 years of age and older, Whites, females, and not dually enrolled in Medicare and Medicaid. Of all the categories of reasons that persons have for receiving Medicare, only

respondents who are aged without ESRD are in significantly better health status as measured by the PIP-DCG score.

A comparison of the differences between eligibles and respondents by plan response rate deciles once again does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles. In fact, there are very few observable differences between eligibles and respondents in health plans at any level of plan response. Of more interest is the difference in the characteristics of eligible beneficiaries in the health plans with the lowest response rates. These health plans tend to have considerably larger proportions of non-White beneficiaries, as well as beneficiaries dually enrolled in Medicare and Medicaid, and more and longer hospital stays.

E.5.6 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey

The annual Medicare CAHPS® Fee-for-Service (FFS) Survey was first fielded in October 2000 and completed in February 2001. The primary mode of data collection was a mail survey with telephone follow-up for mail non-respondents. A Spanish version of the questionnaire was available.

The sample of 162,130 eligible beneficiaries selected for the 2000 Medicare CAHPS® FFS Survey was drawn from a frame created from the August 2000 version of the Medicare Enrollment Data Base (EDB). The frame comprised 30.1 million persons who were enrolled in Medicare FFS for at least the prior 6 months and resided in the 50 states, the District of Columbia (DC), or Puerto Rico. Beneficiaries identified in the survey as being under 18 years of age, not identifying themselves as enrolled in the Original FFS Medicare plan, or deceased before or during the data collection period were treated as ineligible.

The sample was drawn from 280 distinct geographic areas in the United States and Puerto Rico, with approximately 600 sample members selected from each geographic area. The geographic areas combine to represent the 50 states, DC, and Puerto Rico. A total of 103,551 surveys were completed, resulting in a 64 percent response rate.

With very few exceptions, the distribution of response rates differs significantly by category of enrollment, demographic, health status, and use measures. The response rates of beneficiaries under age 65 and above age 74 are significantly lower than those for beneficiaries 65 to 74 years of age. This finding is particularly true for the youngest and oldest age groups. The response rates for all of the racial/ethnic groups are significantly lower than those for Whites. Beneficiaries not dually enrolled in Medicare and Medicaid have a significantly higher response rate than dually enrolled beneficiaries. Beneficiaries entitled to Medicare because they are disabled or because of ESRD only responded at a significantly lower rate than aged beneficiaries without ESRD. Beneficiaries in better health have a significantly higher response rate than those in an average health state, and those in poor health have significantly lower response rates than those in an average health state. Only beneficiaries who had three or more hospital discharges have a significantly lower response rate than those who had not been hospitalized in the prior year.

There are statistically significant differences in the mean PCS and MCS scores according to categories of enrollment, demographic, health status, and medical use variables. Medicare beneficiaries younger and older than 65 to 74 years of age have much lower mean PCS scores than beneficiaries age 65 to 74. Hispanic, American Indian, and Black Medicare beneficiaries

have much lower mean PCS scores than White beneficiaries. Female beneficiaries have a slightly lower mean PCS score than males. Beneficiaries dually enrolled in Medicare and Medicaid have a much lower mean PCS score than beneficiaries in Medicare alone. Beneficiaries entitled to Medicare because of disability or ESRD have considerably lower mean PCS scores than those whose only entitlement to Medicare is because of age. Compared to beneficiaries with average health status, beneficiaries with better health have progressively higher mean PCS scores, while beneficiaries with poor health have progressively lower mean PCS scores. Beneficiaries with hospital stays during the prior year have progressively lower mean PSC scores, as the number of stays increase when compared to Medicare beneficiaries without a prior hospital stay. Mean MCS scores also differ substantially across respondents based on sociodemographic characteristics and levels of health status. A similar pattern as that observed for PCS scores is observed for MCS scores for the demographic characteristics of age, race, gender, dual Medicare and Medicaid enrollment, disability and ESRD, health status, and number of hospitalizations.

With respect to beneficiaries' rating of their satisfaction with Original FFS Medicare, persons under 65 (the disabled) rate satisfaction with Medicare lower than persons in the 65 to 74 age category, while those over 74 self-report higher rates of satisfaction than beneficiaries age 65 to 74. Women rate Medicare higher than men. Beneficiaries who are entitled to Medicare because of their disability (without ESRD) or only because of ESRD rate Medicare lower than those entitled because they are aged (without ESRD). Beneficiaries with a PIP-DCG risk score in the better health status categories rate Medicare lower than those in the average health status category. And, beneficiaries with one or two hospital stays in the year prior to survey have higher levels of satisfaction than persons with no prior hospitalizations.

With respect to beneficiaries' reported level of satisfaction with getting needed care, beneficiaries under 65 years of age report slightly lower satisfaction with getting needed care than those 65 to 74 years of age, while persons 75 to 84 years of age report slightly higher satisfaction. Beneficiaries dually enrolled in Medicare and Medicaid have a lower level of satisfaction with getting needed care than those not also enrolled in Medicaid. And, disabled beneficiaries (without ESRD) have a lower level of satisfaction with getting needed care than aged beneficiaries (without ESRD).

A comparison of PIP-DCG risk scores for respondents and the entire sample of eligible beneficiaries shows a 2 percent difference between them, with respondents having lower scores, indicating better health. Differences in the mean health status (as represented by the PIP-DCG risk score) between survey eligibles and respondents display a general trend in which health status estimates for respondents derived using the PIP-DCG risk score are modestly lower (better health) than those derived for survey eligibles across most major subpopulations of Medicare beneficiaries. There is one noted exception. Respondents who are entitled to Medicare because of a disability produce an average health status estimate that is 2 percent worse than an estimate derived for all survey eligibles.

A comparison of the differences between eligibles and respondents by state-level response rate deciles does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles. Although we observe statistically significant differences between eligibles and respondents for some subpopulations,

15

the level of difference is small. As the state-level response rate increases, the proportion of eligibles who are White increases as well, and the difference in the proportion of Whites between respondents and eligibles declines. The same is true for the proportion of beneficiaries who are eligible for Medicare because they are aged without ESRD. The percent of eligibles dually enrolled in Medicare and Medicaid declines as state-level response rates increase. As noted with race, the difference in the proportion of dual enrollees between respondents and eligibles declines as response rates increase. A similar pattern emerges from the analysis of mean PIP-DCG risk scores: as the response rate at the state level increases, the mean risk score for eligibles and respondents both decline, but the respondents are significantly less healthy across the board. Mean number of hospitalizations and hospital inpatient days follow a similar pattern, but the difference between respondents and eligibles is only significant for one of the levels of state response rate.

CHAPTER 1
INTRODUCTION

1.1   Purpose and Background

The Centers for Medicare & Medicaid Services (CMS) annually conducts several large surveys of Medicare beneficiaries to assess, among other things, their self-reported health status, their recollection of health services used, and their reported satisfaction with their health plans and care. The goal of this project is to examine the potential degree of non-response bias in two major survey efforts that collect information from five different Medicare beneficiary populations: the 2000 Medicare Health Outcomes Survey (HOS) and the 2000 Consumer Assessment of Health Plans Survey (CAHPS®).[6]

Both surveys are important instruments that were designed and administered as a part of a larger CMS initiative to monitor and improve the quality of care provided to Medicare beneficiaries. The HOS is a HEDIS®[7] effectiveness-of-care measure that monitors the quality of care provided to Medicare beneficiaries by measuring changes in health status between two years. The CAHPS® Survey is actually a family of surveys designed to collect information that may help beneficiaries make informed Medicare enrollment choices. The CAHPS® Survey provides a set of meaningful and reliable consumer-oriented measures on beneficiaries' experiences and satisfaction with health care. A variation of the CAHPS® Survey has been developed to ascertain reasons why Medicare beneficiaries voluntarily disenroll from a Medicare+Choice (M+C)[8] managed care plan.

Non-response may be a major threat to the validity of survey sample estimates obtained from these two important surveys of Medicare beneficiaries. There are two possible types of non-response in surveys. One type occurs when a selected sample member does not respond at all to the survey. The second occurs when a selected sample member responds to some items but fails to answer all of them. Typically, the first type is referred to as survey non-response and the second as missing data or item non-response. Non-response bias is the systematic difference between the outcome scores for survey respondents and the (unknown) scores that would have been obtained if all subjects had completed the entire survey. The degree of bias is determined by two factors: (1) the difference in characteristics of interest (e.g., health status) between respondents and non-respondents, and (2) the non-response rate.

In this study, we examine the degree of potential non-response bias to six surveys of Medicare beneficiaries:

- 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey

---

[6] CAHPS® is a registered trademark of the Agency for Healthcare Research and Quality.

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

[8] The Medicare Modernization Act renamed Medicare+Choice (M+C) managed care plans as Medicare Advantage health plans. We retain the reference to M+C health plans for editorial convenience.

- 2000 Cohort 3 Baseline Medicare Health Outcomes Survey

- 2000 CAHPS® M+C Enrollee Survey

- 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey

- 2000 CAHPS® M+C Disenrollment Assessment Survey

- 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only)

Each survey is conducted on a large, national probability sample of Medicare beneficiaries. However, the surveys differ in a number of important ways. One of the surveys is conducted on a different segment of the Medicare population—persons in Medicare fee-for-service—than the other five, all of which are conducted with beneficiaries either enrolled in a Medicare managed care plan or recently withdrawn from one. Another is a survey employing a longitudinal design, the Medicare HOS, whereas the others employ more typical cross-sectional designs.

In addition, the sampling frames have different eligibility criteria and differing levels of information available to eliminate Medicare beneficiaries ineligible for the survey. Beneficiaries determined to be deceased during the time of survey administration are removed from the HOS Baseline Survey and all four CAHPS® Surveys as ineligibles; however, they are retained for analytic purposes in the HOS Follow-up Survey. Beneficiaries with end-stage renal disease (ESRD) may or may not be included in the surveys. Institutionalized beneficiaries are removed from some but not all of the sampling frames before sample selection. However, for several of the surveys, institutionalized beneficiaries are only removed from the sample when someone reports in a returned survey or on a follow-up phone contact that the person to whom the survey is addressed is in a nursing home. Clearly, some institutionalized beneficiaries still could be included among the non-respondents to these surveys. In stark contrast, the Disenrollment Reasons Survey considers only those beneficiaries who are contacted and confirmed to have voluntarily disenrolled from a particular health plan to be eligible for the survey. All non-respondents are considered ineligible.

Further, the definition of a respondent is considerably different across the six surveys. For example, the CAHPS® FFS Survey considers a beneficiary who answers a single question to be a respondent. The Medicare HOS Survey considers a beneficiary who provides answers to a sufficient number of questions to allow for the calculation of a physical component or mental component summary score to be a respondent. Once again, the Disenrollment Reasons Survey differs substantially from the other surveys. A respondent is defined as a beneficiary that provided an affirmative response to one of the preprinted reasons for disenrollment from the plan.

These and other factors may independently or collectively affect the number and types of beneficiaries who respond to the six surveys as well as the analytic estimates of interest derived from the surveys. Thus, it is important to understand the sampling frame, eligibility criteria, and survey response definition of each survey to allow for an informed interpretation of the

non-response bias analyses. More detailed information on these issues is provided for each of the surveys in the chapters that follow.

Current weighting of the survey responses to account for design and non-response effects differs across the surveys as well. The HOS does not employ design or non-response weights. In contrast, the M+C CAHPS® Disenrollment Assessment Survey uses response propensities from logistic regression models to adjust the initial design-based weights upward for respondents so that they represent both respondents and non-respondents. Further, the research to understand the degree of potential non-response bias within each of the survey efforts has been limited to the use of only demographic and entitlement data available from CMS' enrollment files.

This project extends prior non-response bias research by including in our analyses a claims-based measure of health status, the Principal Inpatient Diagnostic Cost Group (PIP-DCG) risk score, available for both survey respondents and non-respondents. To the extent that a claims-based measure of health status is a reasonable proxy for self-reported health status obtained in the HOS and the FFS CAHPS® Surveys, we will be able to directly assess the probable degree of non-response bias for estimates of health status derived from survey respondents only. Similarly, if health status is correlated with measures of satisfaction and experiences with care that are ascertained in the CAHPS® Surveys, we will also be able to assess the probable degree of non-response bias for satisfaction estimates from the CAHPS® Surveys.

We conduct our assessment of probable degree of non-response bias in two steps. First, we evaluate how respondents differ from non-respondents by demographic, entitlement, and health status characteristics. If non-respondents are drastically different from respondents but represent only a negligibly small fraction of survey eligibles, then overall bias might not be significant.

Second, to understand how biased overall survey estimates become by using data from respondents only, we compare the PIP-DCG health status scores between eligibles and respondents, rather than between respondents and non-respondents. In doing so, we are assuming that the average health status estimate for the eligibles closely approximates the average health status estimate for the population from which the sample was drawn. Because the number of eligible beneficiaries is very large in most of the surveys (e.g., almost 300,000 for the M+C Cohort 1 Baseline HOS), we consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the probable degree of bias that is present.

1.2   Methods

1.2.1 Data Sources and Linkage

For each survey, we calculate response rates in total, by plan or state, and by demographic and enrollment characteristics of the beneficiaries. We estimate the scope of potential non-response bias by examining the differences in demographic characteristics, enrollment, health status, and service utilization between respondents and non-respondents. Demographic and enrollment information is obtained from CMS' Denominator file and Group Health Plan file. To evaluate the differences in health status and service utilization, we use M+C

19

inpatient encounter data and Medicare FFS claims data that are available for both respondents and non-respondents. The claims and M+C inpatient encounter data provide information on hospitalization rates, inpatient days, and diagnoses, which allows for the calculation of the PIP-DCG risk adjustment score. The PIP-DCG risk score is used as a measure of predicted Medicare expenditures and health status. More detailed information on several of the data sources and methods of obtaining and linking data from the various files is provided below.

Claims and Encounter Data—CMS collects encounter data from M+C plans for use in "claims-based" diagnostic risk adjustment. The Balanced Budget Act of 1997 mandated Medicare to implement risk adjusted payment for M+C plans in the year 2000. It also required plans to supply encounter data to CMS to support risk adjusted payments. Hospitals submit data to plans for plan enrollees who have a hospital discharge using the CMS 1450 (UB-92) Uniform Institutional Provider Claim Form or the Medicare Part A ANSI ASC X12 837 record. Plans may either submit a complete UB-92/ANSI 837 or an abbreviated UB-92 record. M+C organizations have been submitting inpatient encounter data from the start date of July 1997. Based on our previous experience in working with encounter data, even the first year of collected data proved sufficient for conducting plan-level analysis (McCall, Harlow, and Dayhoff, 2001).

Inpatient encounter data are available for all managed care enrollees, including both respondents and non-respondents to the Medicare HOS and CAHPS® Surveys. These data are used to compare the health status of survey respondents and non-respondents. For example, the presence and number of hospitalizations are used in these comparisons as markers of poor health. Since PIP-DCG risk scores predict future Medicare expenditures and compare these expenditures to the general Medicare population, mean PIP-DCG risk scores may be used as an overall measure of beneficiary group health status.

Inpatient hospital claims data for Medicare beneficiaries in FFS are obtained from CMS' standard analytic inpatient file derived from claims submitted from hospitals treating Medicare FFS beneficiaries using the CMS 1450 (UB-92) Uniform Institutional Provider Claim Form. Thus, it contains diagnostic information similar to that found for M+C inpatient encounters.

Risk Score—We use PIP-DCG risk adjustment scores as a measure of future expenditures and health status. The PIP-DCG model was implemented in 2000 by CMS to adjust a portion of capitation payments to M+C organizations. As its name suggests, the PIP-DCG model combines principal inpatient diagnoses with demographic information to develop an index of predicted (future) health care expenditures. A risk adjustment score of 1.0 indicates an average level of predicted future expenditures. The PIP-DCG model includes 16 diagnostic categories with numerical labels ranging from 4 to 29. Each numerical label is intended to roughly indicate the predicted expenditure level, in thousands of 1996 dollars, for people classified in this group. People assigned to PIP-DCG category 4, which includes beneficiaries with no hospital admissions in the previous year, receive a risk score based on demographic factors only. For other model categories, a person's principal inpatient diagnosis contributes to the risk score. For full information on this model, consult RTI's report *Principal Inpatient Diagnostic Cost Group Models for Medicare Risk Adjustment* (Pope et al., 1999).

Matching of Claims and Encounter Data with Survey Data—In this project, we use M+C encounter data and FFS inpatient claims to measure utilization, such as the number of

hospitalizations and inpatient days, as well as to calculate PIP-DCG risk scores for HOS and CAHPS® respondents and non-respondents within each plan. The source of these data is FU and Associates, CMS' contractor responsible for maintenance of the M+C encounter data files and the annual construction of PIP-DCG scores for all Medicare beneficiaries.

Annually, FU and Associates constructs PIP-DCG scores for all Medicare beneficiaries to be used for payment calculations for M+C plans for a future time period. To be included in the file that develops payment rates using the full PIP-DCG model, a beneficiary must be enrolled in Medicare Part A for the full 12-month time period. Beneficiaries who are new enrollees receive a PIP-DCG score that is calculated only with demographic information. Since we are interested in obtaining estimates of health status during the survey year, we requested utilization data and PIP-DCG risk scores that were constructed with 1999 data to predict expenditure year 2000. Specifically, RTI submitted to FU and Associates a finder file of unique cross-referenced health insurance claim (HIC) numbers and requested the following information for all survey eligibles:

- final reconciliation Part A and Part B PIP-DCG risk scores that are based on July 1, 1998, through June 30, 1999, inpatient encounter data that are used to predict 2000 expenditures

- a (0,1) binary variable flag that indicates the following about the risk score: = 1 if reconciled risk score and = 0 if new enrollee risk score

- number of hospitalizations during July 1, 1998, through June 30, 1999

- number of inpatient days during July 1, 1998, through June 30, 1999

Survey and Claims Data Statistics—**Table 1-1** contains a summary of key elements of the six surveys including

- number of health plans (states) included in the survey;

- number of eligible beneficiaries;

- number of eligible beneficiaries with a calculated PIP-DCG score and the proportion that had scores calculated with demographic and claims data versus just demographic information;

- response rate using each survey's definition of response; and

- mean PIP-DCG risk score, mean number of hospitalizations, and mean number of inpatient days.

Table 1-1
Survey and Medicare+Choice (M+C) Hospital Encounter Data Match Rates, Mean Response Rates,
and Mean Health Status and Hospital Use Statistics among Eligibles for Six Surveys of Medicare Beneficiaries

Analytic Variable	Cohort 3 Baseline Medicare HOS	Cohort 1 Follow-up Medicare HOS	CAHPS® M+C Enrollee Survey	Medicare CAHPS® FFS Survey	CAHPS® M+C Disenrollment Assessment Survey	CAHPS® M+C Disenrollment Reasons Survey
Number of Health Plans or States (CAHPS® FFS)	306	225	292	52	279	239
Match between Eligibles and Encounter Data						
Number of Eligibles	291,221	88,129	216,919	162,130	22,272	12,658
Number of Eligibles with PIP-DCG Scores	291,205	88,129	216,915	162,126	22,272	12,658
Risk Score Calculation Indicator						
Calculated from Demographic Information Only	6.75	0.12	7.84	10.10	9.89	11.00
Calculated from Demographic and Hospitalization Data	93.25	99.88	92.16	89.90	90.11	89.00
Response Rates						
Unweighted Survey-Specific Response Rate[1] (%)	72	85	83	64	55	58
Mean of the Means of Survey-Specific Response Rate[2] (%)	73	86	83	65	55	60
Selection Probability Weighted Mean Survey-Specific Response Rate[3] (%)	71	85	82	63	55	57
Descriptive Health Status and Utilization Statistics						
Mean PIP-DCG Risk Score[1]	0.90	0.91	0.88	0.97	0.91	0.91
Mean Number of Hospitalizations[1]	0.20	0.20	0.19	0.25	0.22	0.21
Mean Number of Inpatient Days[1]	7.17	7.07	7.05	9.33	8.92	8.80

[1] An equal weight whereby all *sampled beneficiaries* are given a weight of 1.

[2] An equal weight whereby all *health plans or states* are given a weight of 1.

[3] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS), 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS), 2000 CAHPS® M+C Enrollee Survey, 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey, 2000 CAHPS® M+C Disenrollment Assessment Survey, and 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only).

We provide mean response rates for each survey calculated using three different methods: a mean calculated with each health plan given equal weight, a mean calculated as the mean of the means of individual health plans' response rate, and a mean calculated using the number of enrollees in the individual plans as the weight. We do so because researchers and CMS report information at the survey level using these three methods. For example, the HOS typically is reported using equal weighting of all plans, while the CAHPS® FFS Survey reports comparative information using the mean of the means approach to weighting.[9] The mean PIP-DCG risk scores and mean number of hospitalizations and number of inpatient days are enrollment weighted.

There is considerable variation across the six surveys in number of Medicare beneficiaries eligible for survey, ranging from 12,658 for the third quarter of the M+C Disenrollment Reasons Survey to almost 300,000 for the Cohort 3 HOS Baseline Survey. Virtually all eligible sample members received a PIP-DCG score; however, the proportion that received a score based upon the full year claims model, rather than just demographic characteristics, varied considerably across the surveys. Not surprisingly, virtually all beneficiaries who were eligible for the HOS Follow-up Survey were scored using the full year claims model, because they would have had to have been enrolled in Medicare 2 years earlier for survey at baseline. The CAHPS® FFS and both M+C Disenrollment Surveys have about 10 percent of their eligibles scored using demographics only. Thus, these three surveys appear to have the highest proportions of new Medicare enrollees.

Survey-specific response rates ranged from the mid-fifties for the two M+C Disenrollment Surveys to a high of 85 percent for the HOS Follow-up Survey. There is limited variation in response rates regardless of method used to weight plan- or state-specific estimates of response. Of note is the fact that despite making contact with beneficiaries to determine eligibility, each of the CAHPS® Disenrollment Surveys had less than a 60 percent response rate. This may reflect the dissatisfaction of the beneficiaries with the health plan spilling over to dissatisfaction with completing a survey about their experiences.

There is a six percentage point difference in mean PIP-DCG risk scores across the surveys, ranging from a low of 0.88 in the CAHPS® M+C Enrollee Survey to a high of 0.97 in the CAHPS® FFS Survey. This indicates that the CAHPS® M+C Enrollee Survey has the healthiest beneficiaries on average, while the CAHPS® FFS Survey has the sickest set of beneficiaries, as a higher score indicates worse health status. A similar pattern is observed for mean number of hospitalizations and mean number of inpatient days.

1.2.2 Analysis Approach

Across all plans and states for the HOS and CAHPS® Surveys, we analyze survey response rates and health status and medical care utilization differences between respondents and non-respondents and between respondents and survey eligibles and present this aggregate

---

[9] Note that because it does not include M+C plans, the CAHPS® FFS Survey reports estimates of experience and satisfaction with care by *state* rather than by plan. Thus, in CAHPS® FFS, state is analogous to health plan, which is used to report information for the other five surveys.

information in tabular format in the main body of this report. Plan-specific data are presented in an Appendix contained on a separate CD-ROM. This report consists of seven analysis chapters: six chapters contain analysis of differences between respondents and non-respondents to each survey using survey-specific response definitions. Each chapter contains a comparison of respondents and non-respondents by demographic characteristics, utilization measures, and PIP-DCG risk scores and is organized in a similar manner. The seventh analysis chapter provides across-survey comparisons of response rates and uses a uniform definition of a respondent, which is a beneficiary that provided a response to the "General self-rated health" question present on all surveys under study. Five key sets of analyses were performed for each survey and are described below.

Survey-Specific Response Rates—We begin our detailed examination of non-response bias by first exploring differences in response rates by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. **Table 1-2** displays the levels of stratification for these variables. Three sets of weights were constructed to allow for an evaluation of the influence of the size of the health plan (state) on calculation of response rates or measures of health status and satisfaction:

- An equal weight whereby all *sampled beneficiaries* are given a weight of 1. These weights are used to produce statistics that we refer to as unweighted.

- An equal weight whereby all *health plans or states* are given a weight of 1. These weights are used to produce statistics that we refer to as the mean of the means, with all health plans or states contributing equally to the calculation of the statistic.

- A selection probability weight whereby all beneficiaries are given a weight based on the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state. Thus, beneficiaries from very large health plans would contribute substantially more influence on the calculation of a statistic than beneficiaries from very small health plans. These weights are used to produce statistics that we refer to as selection weighted.

Pairwise comparisons of differences selection weighted response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment.

Differences in Characteristics of Respondents and Non-respondents—Second, we explore differences in the distribution of beneficiary characteristics between respondents and non-respondents using the stratifying variables displayed in **Table 1-2**. This review provides an overall sense of how different respondents are from non-respondents in terms of demographic and health status characteristics and is a critical factor in the determination of potential non-response bias. Statistical significance testing is performed using the chi-square test and $p<0.05$ level of significance.

Table 1-2
Levels of Stratification for Demographic, Enrollment, and Health Status and Medical Use

Age
  Under 65
  65-74
  75-84
  85 +

Race
  Unknown
  White
  Black
  Other
  Asian
  Hispanic
  American Indian

Gender
  Male
  Female

Medicaid Status
  Not Enrolled
  Enrolled

Institutionalized Status
  Community Dwelling
  Long-term Institutionalized
  Nursing Home Certifiable

Reason for Medicare Entitlement
  Aged without ESRD
  Aged with ESRD
  Disabled without ESRD
  Disabled with ESRD
  ESRD Only

Risk Score Decile
  0.36 - 0.45
  0.46 - 0.53
  0.54 - 0.57
  0.58 - 0.70
  0.71 - 0.73
  0.74 - 0.87
  0.88 - 0.91
  0.92 - 1.07
  1.08 - 1.26
  1.27 - 6.91

Number of Hospitalizations
  Zero
  One
  Two
  Three or More

Differences in Outcomes by Demographic and Health Status Characteristics—Third, we explore differences in survey-specific outcome scores by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. For the HOS and the CAHPS® FFS Surveys, we report mean physical component summary (PCS) and mental component summary (MCS) scores as the outcome measure. For the M+C CAHPS® Enrollee and Disenrollment Assessment Surveys and the CAHPS® FFS Survey, we report estimates of plan satisfaction and getting care as the outcome measures. If outcome measures, such as satisfaction with care and physical health, vary by demographic characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias increases. Pairwise comparisons of differences in response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment.

Factors that Predict Likelihood of Response—Univariate and bivariate analysis of response rates by sociodemographic and health status characteristics (e.g., race/ethnicity), while useful, may result in misleading conclusions, especially when many of the beneficiary factors are highly correlated. Our fourth analysis is a multivariate logistic regression analysis of likelihood of response that is estimated as a function of demographic and health status variables to provide estimates of the *independent* effect of beneficiary characteristics on response. We estimate separate models for each of the surveys.

The logistic regression model of response propensity takes the form

$$Log[P/(1-P)] = \beta_1 X_1 + \beta_2 X_2 + e$$

where
- P is the probability of the beneficiary responding to the survey;

- X_1 is a set of variables containing demographic and enrollment indicators for age, gender, race/ethnicity, dually enrolled Medicare/Medicaid status, institutionalized status (HOS only), and ESRD status;

- X_2 denotes a set of health status variables available from claims data and includes the PIP-DCG risk score and number of hospitalizations in 1999; and

- e is an error term.

The β coefficients are not directly interpretable; therefore, they are transformed into odds ratios that reflect the increased or decreased likelihood of responding when the independent variable is present. Odds ratio must be greater than zero; variables having a negative effect on the outcome variable will have an odds ratio between 0 and 1. An odds ratio of 1.35, for example, would indicate that a male beneficiary is 35 percent more likely to respond than a female beneficiary, while an odds ratio of 0.50 indicates the male beneficiary is only half as likely to respond as a female beneficiary.

26

We estimate the model unweighted and weighted by the probability of selection for survey.

Probable Degree of Non-response Bias—Fifth, we directly explore the degree of bias that may be present in estimates of health status and medical care usage by comparing means of these variables for respondents to those obtained for eligible beneficiaries, including non-respondents. We also report differences in mean PIP-DCG scores between respondents and survey eligibles, stratified by sociodemographic and medical care usage characteristics.

Last, we examine the differences between eligibles and respondents by plan response rate deciles to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles. Between eligibles and respondents, we compare average age; the proportion that are female, White, enrolled in Medicaid, and aged without ESRD; average PIP-DCG risk score; number of hospitalizations; and number of inpatient days. Pairwise comparisons of differences in mean estimates between eligibles and respondents are made within deciles of response using a two-sided z-test for differences in means or proportions at the significance level of $p<0.01$ to account for multiple comparisons. Because the number of eligible beneficiaries is very large for all but the CAHPS® Disenrollment Survey (e.g., almost 300,000 for the Cohort 3 Baseline HOS), we consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

1.3   Organization of Report

This report consists of eight chapters. The first chapter contains the introduction and background to the report. Chapters 2 through 7 contain analysis of differences between respondents and non-respondents to each survey using survey-specific response definitions. Each chapter contains a comparison of respondents and non-respondents by demographic characteristics, utilization measures, and PIP-DCG risk scores and is organized in a similar manner. Chapter 2 contains the non-response bias analysis for the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey. Chapter 3 contains the non-response bias analysis for the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey. This chapter also contains a unique analysis of response to the Follow-up Survey using Baseline Survey characteristics. Chapter 4 contains the non-response bias analysis for the 2000 CAHPS® M+C Enrollee Survey. Chapter 5 contains the non-response bias analysis for the 2000 CAHPS® M+C Disenrollment Assessment Survey. Chapter 6 contains the non-response bias analysis for the 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only). Chapter 7 contains the non-response bias analysis for the 2000 Medicare CAHPS® Fee-for-Service Survey. And, Chapter 8 provides across-survey comparisons using a uniform definition of a respondent, which is a beneficiary that provided a response to the "General self-rated health" question present on all surveys under study.

CHAPTER 2
ANALYSIS OF NON-RESPONSE BIAS IN THE 2000 COHORT 3 BASELINE MEDICARE HEALTH OUTCOMES SURVEY

2.1 Description of the Medicare Health Outcomes Survey

The Medicare Health Outcomes Survey (HOS), formerly known as the Health of Seniors Survey, is a survey of Medicare beneficiaries enrolled in Medicare+Choice (M+C) managed care organizations[10] (MCOs). The HOS instrument is based on the SF-36®[11] Health Survey, or SF-36, which asks the respondent to rate general health, ability to perform certain physical tasks, level of pain, and emotional state. Summary scales of physical and mental health, denoted as physical component summary (PCS) and mental component summary (MCS), respectively, are calculated using eight scales based on all 36 questions. Both components are normed such that the mean score is 50 with a standard deviation of 10 points in the general U.S. population. The HOS also includes additional questions on health status, activities of daily living, specific medical conditions, and demographics.

PCS scores are a reliable and valid measure of physical health. Very high PCS scores indicate no physical limitations, disabilities, or decline in well-being; high energy level; and a rating of health as excellent. Very low PCS scores indicate limitations in self-care and physical, social, and role activities; severe bodily pain; frequent tiredness; and a rating of health as poor.

MCS scores are a reliable and valid measure of mental health. Very high MCS scores indicate frequent positive affect, absence of psychological distress, and no limitation in usual social and role activities due to emotional problems. Low MCS scores indicate frequent psychological distress, and social and role disability due to emotional problems.

Since 1998, six baseline and four follow-up surveys have been successfully administered with the seventh HOS round fielded during spring 2004. The year 2000 was the first year both the follow-up and the baseline surveys were administered. In the first HOS cohort of 1998, enrollees in social health maintenance organizations (SHMOs) and the Medicare Choices demonstrations were sampled. The second cohort (1999) added specialty organizations such as PACE and EverCare, but EverCare plans were omitted starting with the 2000 Cohort 3 Baseline Survey. The 2000 survey included a wide range of health plan types: all M+C organizations, continuing cost contracts, PACE plans, SHMOs, Medicare Choices, and DoD Subvention Demonstration plans with a contract effective on or before January 1, 1999.

This chapter reports analyses of possible non-response bias for the 2000 Cohort 3 Baseline HOS Survey, which sampled a wide range of managed care organizations. We defined survey eligibles as Medicare beneficiaries who had to have been continuously enrolled in the same health plan for at least 6 months at the time the sample was drawn. Beneficiaries with end-stage renal disease (ESRD) were excluded from the sampling frame. One thousand eligible

---

[10] HEDIS® 2000 manual, volume 6 (NCQA, 2000), is used for the background information on the HOS Survey.

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

29

beneficiaries are sampled from each participating MCO. In health plans with 3,000 or more members, those who were sampled and participated in the Cohort 2 Baseline Survey were excluded. In plans with 1,000 or fewer enrollees, the entire eligible plan population is sampled.

In addition, some adjustments in calculating the number of eligible beneficiaries are made based on the survey disposition codes. The following individuals are declared ineligible ex post:

- those reported deceased

- those unable to complete the survey because of language barriers

- those with bad addresses and non-working telephone numbers

- those not enrolled in the appropriate MCO

- those with ESRD

The M+C HOS is a self-administered mail survey with telephone follow-up. The sample is drawn in March of each year, and surveying begins shortly thereafter and concludes in early summer. About 300 plans participated in the HOS 2000. The overall response rate was 72 percent for the 2000 Cohort 3 Baseline Survey. A Spanish version of the survey was completed by 337 beneficiaries.

The definition of a complete survey, as specified by the HOS protocol, was 80 percent or more of the total questions answered. Since the SF-36 instrument is the core component of the HOS, producing PCS and MCS scores is a central objective of the survey. As a result, researchers may define respondents as those for whom PCS and MCS scores can be calculated. In this project, we used the latter definition requiring the calculation of a PCS or MCS score to be considered a respondent.

The M+C HOS does not use design or non-response weights to adjust survey responses for differential rates of response among survey eligibles.

2.2   Survey-Specific Response Rates

We begin our detailed examination of non-response bias by first exploring differences in response rates by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. **Table 2-1** displays three sets of response rates. The first set of response rates is calculated using equal weighting, whereby all *sampled beneficiaries* are given a weight of 1. The second set of response rates is calculated as the mean of the mean, whereby all *health plans or states* are given a weight of 1. The third set of response rates is calculated using a selection probability weight, whereby all beneficiaries are given a weight based upon the likelihood of selection. Pairwise comparisons of differences in selection weighted response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p < 0.05$ with the Bonferroni multiple comparison adjustment.

Table 2-1
Survey-Specific Response Rates by Demographic and Health Status Characteristics for the Cohort 3 Baseline Medicare Health Outcomes Survey

Characteristic	Unweighted Response Rate[1] (%)	Mean of the Means Response Rate[2] (%)	Selection Probability Weighted Response Rate[3] (%)	
Age[4]				
Under 65	65	66	63	*
65-74	74	74	72	
75-84	73	73	72	*
85 +	63	64	62	*
Race				
Unknown	66	65	63	*
White	74	73	72	
Black	62	65	61	*
Other	67	67	67	*
Asian	75	74	74	
Hispanic	64	67	62	*
American Indian	71	72	62	
Gender				
Male	72	72	70	
Female	72	72	71	*
Medicaid Status				
Not Enrolled	73	73	71	
Enrolled	60	63	60	*
Institutionalized Status				
Community Dwelling	72	72	71	
Long-term Institutionalized	29	32	28	*
Nursing Home Certifiable	67	66	70	
Reason for Medicare Entitlement				
Aged without ESRD	73	73	71	
Aged with ESRD	71	69	83	*
Disabled without ESRD	65	66	64	*
Disabled with ESRD	60	63	35	
ESRD Only	85	85	91	
Risk Score Decile				
0.36 - 0.45	75	75	74	*
0.46 - 0.53	74	74	72	*
0.54 - 0.57	75	75	73	*
0.58 - 0.70	73	73	71	
0.71 - 0.73	74	74	71	*
0.74 - 0.87	74	73	73	*
0.88 - 0.91	73	73	72	*
0.92 - 1.07	71	72	70	
1.08 - 1.26	66	66	64	*
1.27 - 6.91	65	66	64	*
Number of Hospitalizations				
Zero	73	73	71	
One	70	70	69	*
Two	68	68	66	*
Three or More	63	63	62	*

[1] An equal weight whereby all *sampled beneficiaries* are given a weight of 1.

[2] An equal weight whereby all *health plans or states* are given a weight of 1.

[3] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[4] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS).

- Response rates vary by all enrollment and demographic characteristics and health status measures evaluated in this study.

  – Compared to beneficiaries age 65 to 74, beneficiaries under the age of 65, or those enrolled in Medicare because of a disability, and those 85 years of age or older have significantly lower rates of response.

  – Beneficiaries whose race/ethnicity is White or Asian are the most likely to respond to the HOS Baseline Survey, while Blacks and Hispanics are the least likely to respond.

  – Females are modestly more likely than males to respond.

  – Beneficiaries dually enrolled in Medicare and Medicaid are significantly less likely to respond. The response rate for dual enrollees is 11 percentage points lower than the response rate of beneficiaries not dually enrolled.

  – The response rate for beneficiaries residing in long-term institutionalized facilities is shockingly low—only 28 percent responded. This is in stark contrast to the response rate of 71 percent for beneficiaries residing in the community. Survey methodologists specializing in the Medicare population have typically found gaining the cooperation of gatekeepers at nursing homes to be very difficult. This response rate grimly reflects the reality of attempting to survey Medicare beneficiaries residing in nursing homes.

  – Interestingly, aged beneficiaries with ESRD are significantly more likely to respond to the HOS Survey than aged beneficiaries without ESRD.

  – The response rate for beneficiaries with an average health status score, or PIP-DCG score of 1.0, is 70 percent. In contrast, the response rate is 6 percentage points lower, or 64 percent, for beneficiaries with poorer health status or higher PIP-DCG scores. In contrast, response rates for beneficiaries in virtually every other PIP-DCG risk score category are higher than 70 percent.

  – The response rate for beneficiaries without any hospitalizations during the year prior to survey is 71 percent. Response rates decline as numbers of hospitalizations increase.

2.3 Differences in Characteristics of Respondents and Non-respondents

Second, we explore differences in the selection probability weighted distribution of beneficiary characteristics between eligibles, respondents, and non-respondents using the previously specified stratifying variables from Chapter 1. This review provides an overall sense of how different respondents are from non-respondents in terms of demographic and health status characteristics and is a critical factor in the determination of potential non-response bias. The distributions are weighted using selection probability. Statistical significance testing is performed using the chi-square test and $p<0.05$ level of significance. We summarize our findings below:

- The distribution of beneficiary characteristics and health status scores in **Table 2-2** differs systematically between respondents and non-respondents.

  – A greater proportion of non-respondents are under age 65 or age 85 and older.

  – The proportion of Whites is lower and Blacks is higher in the non-respondent population.

  – The proportion of Medicare and Medicaid dual enrollees is higher for non-respondents.

  – The proportions of Medicare beneficiaries residing in a long-term institution and the proportion of beneficiaries entitled to Medicare because of disability are higher among non-respondents, albeit the actual percentage point differences are very modest.

  – The distribution of health status risk scores is skewed more toward higher levels of disability among the non-respondents as compared to respondents. Twenty-five percent of the non-respondents have risk scores 8 percent or higher than the average score. This is in contrast to 18 percent for respondents.

  – There are also modestly more hospitalizations among non-respondents as compared to respondents.

2.4   Differences in Outcomes by Demographic and Health Status Characteristics

Third, we explore differences in survey-specific outcome scores by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. For the HOS, we report mean PCS and MCS scores as the outcome measure. If outcome measures, such as physical health, vary by demographic characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias increases. Pairwise comparisons of differences in selection probability weighted response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. Our findings are summarized as follows:

- Across all respondents in **Table 2-3**, the mean PCS score is 41.88. This is about 8 percentage points lower than the norm-based mean of 50 for the general population, indicating that respondents to the Cohort 3 Baseline HOS have, in general, a lower level of physical health than the general population.

- Mean PCS scores differ substantially across respondents based on sociodemographic characteristics and levels of health status.

Table 2-2
Distribution of Demographic and Health Status Characteristics among
Cohort 3 Baseline Medicare Health Outcomes Survey Eligibles, Respondents,
and Non-respondents, Selection Probability Weighted[1]

Characteristic	Eligibles (%)	Respondents (%)	Non-respondents (%)	
Age[2]				
Under 65	7	6	8	*
65-74	49	51	45	
75-84	34	34	33	
85 +	10	9	13	
Race				
Unknown	0.5	0.4	0.6	*
White	86	87	81	
Black	9	8	12	
Other	2	2	3	
Asian	1	1	1	
Hispanic	2	2	2	
American Indian	0.1	0.1	0.1	
Gender				
Male	43	43	43	
Female	57	57	57	
Medicaid Status				
Not Enrolled	95	96	93	*
Enrolled	5	4	7	
Institutionalized Status				
Community Dwelling	99	99	98	*
Long-term	0.6	0.2	1	
Nursing Home Certifiable	0.2	0.2	0.3	
Reason for Medicare				
Aged without ESRD	93	94	91	*
Aged with ESRD	0	0	0	
Disabled without ESRD	7	6	9	
Disabled with ESRD	0	0	0	
ESRD Only	0	0	0	
Risk Score Decile				
0.36 - 0.45	10	11	9	*
0.46 - 0.53	10	10	9	
0.54 - 0.57	14	14	12	
0.58 - 0.70	5	5	5	
0.71 - 0.73	9	9	8	
0.74 - 1.87	12	13	12	
0.88 - 0.91	12	12	11	
0.92 - 1.07	8	8	8	
1.08 - 1.26	10	9	12	
1.27 - 6.91	10	9	13	
Number of Hospitalizations				
Zero	87	87	85	*
One	9	9	10	
Two	3	2	3	
Three or More	1	1	2	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Statistical significance tested using chi-square distribution differences between respondents and non-respondents. An asterisk (*) denotes significance at <0.05 level.

Source: RTI analysis of the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS).

Table 2-3
Average Physical and Mental Health Component Scores by Demographic and Health Status Characteristics of Respondents to the Cohort 3 Baseline Medicare Health Outcomes Survey, Selection Probability Weighted[1]

Characteristic	Physical Health Component Score (PCS) (mean)	Mental Health Component Score (MCS) (mean)
Across all Respondents	41.88	50.84
Age[2]		
Under 65	32.09 *	40.91 *
65-74	44.65	52.47
75-84	40.85 *	50.69 *
85 +	36.55 *	48.34 *
Race		
Unknown	42.24	50.68
White	42.02	51.09
Black	39.78 *	49.04 *
Other	43.45 *	49.86 *
Asian	42.98 *	51.04 *
Hispanic	40.72 *	47.94 *
American Indian	36.87 *	46.19 *
Gender		
Male	42.81	51.17
Female	41.19 *	50.60 *
Medicaid Status		
Not Enrolled	42.12	51.08
Enrolled	35.82 *	44.80 *
Institutionalized Status		
Community Dwelling	41.93	50.88
Long-term Institutionalized	28.53 *	37.25 *
Nursing Home Certifiable	29.77 *	43.29 *
Reason for Medicare Entitlement		
Aged without ESRD	42.47	51.42
Aged with ESRD	30.74	45.98 *
Disabled without ESRD	32.18 *	41.28 *
Disabled with ESRD	52.98 *	57.28
ESRD Only	27.27 *	52.38
Risk Score Decile		
0.36 - 0.45	45.68 *	52.43 *
0.46 - 0.53	45.69 *	52.35 *
0.54 - 0.57	44.88 *	52.61
0.58 - 0.70	41.99 *	49.93 *
0.71 - 0.73	45.64 *	53.09 *
0.74 - 0.87	41.00 *	50.81 *
0.88 - 0.91	40.88 *	50.67
0.92 - 1.07	39.22	49.68
1.08 - 1.26	36.56 *	47.96
1.27 - 6.91	35.43 *	47.08
Number of Hospitalizations		
Zero	42.63	51.24
One	37.35 *	48.67 *
Two	35.69 *	47.15 *
Three or More	33.35 *	45.83 *

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS).

–   Compared to beneficiaries age 65 to 74, beneficiaries of all other age groups have lower physical functioning.

–   Compared to Whites, Blacks and Hispanics have lower mean PCS scores, while Asians and beneficiaries with race coded as "other" have higher mean PCS scores.

–   Men self-report a higher level of physical health than women.

–   Beneficiaries dually enrolled in Medicare and Medicaid self-report a significantly lower level of physical health (PCS of 35.8) than beneficiaries not enrolled in Medicaid (PCS of 42.1).

–   Beneficiaries residing in long-term care facilities and beneficiaries residing in the community who are deemed nursing home certifiable report significantly lower levels of physical functioning than community residents.

–   Compared to beneficiaries with average health status (PIP-DCG score range from 0.92 to 1.07), beneficiaries with better health status (lower scores) have higher average PCS scores and beneficiaries with worse health status (higher scores) have lower PCS scores. The negative relationship with PCS scores and PIP-DCG scores also appears linear in nature. A simple test of the correlation between the two health status scores revealed a negative but fairly weak correlation of -0.25. However, the Pearson product-moment correlation coefficient test assumes normality of distribution of both variables, and both exhibit some non-normality tendencies. Further research into the relationship between the two scores would appear to be warranted before a definitive statement regarding correlation and substitutability is made.

–   There is also an observed negative relationship with PCS scores and number of hospitalizations; as frequency of prior year hospitalizations increase, one observes declining average PIP-DCG scores.

•   Across all respondents, the mean MCS score is 50.84, indicating that respondents to the HOS have, in general, a similar level of self-reported mental health as the general population (mean of 50).

•   Mean MCS scores also differ substantially across respondents based on sociodemographic characteristics and levels of health status.

–   A similar pattern as that observed for PCS scores is observed for MCS scores for the demographic characteristics of age, race, gender, Medicaid enrollment, and institutionalized status, and with number of hospitalizations.

–   The aged with ESRD and the disabled without ESRD have lower self-reported mental health than beneficiaries who are aged without ESRD.

- The PIP-DCG score is not as strongly associated with the MCS score as it was for the PCS score. Beneficiaries with the highest health status (lowest PIP-DCG scores) have MCS scores higher than beneficiaries who have average health status (PIP-DCG score = 1.0). However, there are no statistically significant differences between mean MCS scores for beneficiaries with average health status and those with worse health status as measured by the PIP-DCG score (scores greater than 1.0).

2.5 Factors that Predict Likelihood of Response

Fourth, we predict the likelihood of response as a function of sociodemographic and health status characteristics of all eligible beneficiaries using a multivariate regression model. In **Table 2-4** we estimate the model unweighted and weighted by the inverse of the probability of selection for survey.

- The direction and magnitude of the odds ratios for the included beneficiary-level variables are consistent with the descriptive comparisons between respondents and non-respondents. The results from the weighted regression model follow:

    - Beneficiaries under the age of 65 and age 85 and older are about 25 percent less likely to respond than beneficiaries age 65 to 74.

    - Beneficiaries of Asian descent are about 10 percent more likely to respond than White beneficiaries. In contrast, all other minority races are far less likely than White beneficiaries to respond to the HOS.

    - Men are less likely to respond than women.

    - Beneficiaries dually enrolled in Medicare and Medicaid are 17 percent less likely to respond than beneficiaries not enrolled in Medicaid.

    - After controlling for health status, race, and age, beneficiaries with ESRD are significantly more likely than beneficiaries without ESRD to respond to the HOS.

    - The long-term institutionalized are 80 percent less likely to respond to the HOS as compared to community residing beneficiaries, while those that are nursing home certifiable are about 12 percent more likely to respond than community residents.

    - Compared to beneficiaries with average health status, those with lower PIP-DCG scores, which equates to a higher level of health status, are generally more likely to respond to the HOS. Beneficiaries in poorer health status, or higher PIP-DCG scores, are less likely to respond.

    - The likelihood of response to the HOS declines as the number of hospitalizations experienced during the year prior to survey increases.

Table 2-4
Logistic Regression of Likelihood of Response to the Cohort 3 Baseline
Medicare Health Outcomes Survey

Characteristic	Unweighted Regression Odds Ratio	Selection Probability Weighted Regression Odds Ratio[1]
Beneficiary Characteristics		
Under 65	0.703	0.728 [2]
75 to 84	0.946	0.978
85 +	0.711	0.716
Black	0.604	0.622
Unknown or Other Race	0.730	0.776
Asian	1.092	1.092
Hispanic	0.664	0.660
American Indian	0.972	0.665
Male	0.967	0.945
Medicaid	0.803	0.830
ESRD	1.328	1.756
Institutionalized Status		
Long-Term Institutionalized	0.219	0.212
Nursing Home Certifiable	0.897	1.117
Risk Score Decile		
0.36 - 0.45	1.100	1.055
0.46 - 0.53	1.070	1.046
0.54 - 0.57	1.072	1.023
0.58 - 0.70	1.071	1.009
0.71 - 0.73	1.050	0.988
0.74 - 0.87	1.070	1.053
0.88 - 0.91	1.044	1.003
1.08 - 1.26	0.919	0.876
1.27 - 6.91	0.831	0.867
Number of Hospitalizations		
One	1.041	1.013
Two	0.989	0.956
Three or More	0.859	0.835
No. of Observations	291221	291221
Overall Chi-Sq (p-value)	5732***	99215***

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Highlighted odds ratios are statistically significant at the $p<0.05$ level of significance. Asterisks (***) denote $p<0.001$ level of significance.

Source: RTI analysis of the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS).

2.6 Probable Degree of Non-response Bias

Fifth, we directly explore the degree of bias that may be present in estimates of health status and medical care usage by comparing means of these variables for respondents to those obtained for eligible beneficiaries, including non-respondents. We also report differences in mean PIP-DCG scores between respondents and survey eligibles stratified by sociodemographic and medical care usage characteristics.

And, last, we examine the differences between eligibles and respondents by plan response rate deciles to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles. Between eligibles and respondents, we compare average age; the proportion that are female, White, enrolled in Medicaid, and aged without ESRD; average PIP-DCG risk score; number of hospitalizations; and number of inpatient days.

Pairwise comparisons of differences in selection probability weighted mean estimates between eligibles and respondents are made using a two-sided z-test for differences in means or proportions at the significance level of $p<0.01$ to account for multiple comparisons. We consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

- Differences in the mean health status and medical use statistics between eligibles and respondents presented in **Table 2-5** reflect the differences previously observed in the underlying distribution of characteristics of respondents and non-respondents, suggesting that respondents, on average, have a modestly higher level of health status than the surveyed population. We draw this conclusion given that we have previously observed a negative correlation between PCS and PIP-DCG risk scores.

  – The mean PIP-DCG score is 2 percent lower for respondents than for survey eligibles, implying a modest degree of non-response bias, which overestimates the health status of the survey population.

  – Mean number of hospitalizations and inpatient days are both lower for respondents than for survey eligibles, also suggesting a small degree of non-response bias.

- Differences in the mean health status between survey eligibles and respondents display a general trend in which health status estimates derived using the PIP-DCG risk score are often lower (better health) than those derived for survey eligibles across most major subpopulations of Medicare beneficiaries (**Table 2-6**). This suggests that health status estimates derived from respondents only tend to modestly overestimate the health of M+C Medicare enrollees. There is a notable exception. Health status estimates derived from the survey respondents who are under age 65 or are entitled to Medicare because of ESRD or a disability are not statistically different from those derived for survey eligibles for similar populations, suggesting limited if any non-response bias in health status estimates for these populations.

39

- A comparison of the differences between eligibles and respondents by plan response rate deciles does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles (**Table 2-7**). In fact, there are no observed differences between eligibles and respondents for the health plans with the lowest level of response. Although we observe statistically significant differences between eligibles and respondents for some subpopulations (e.g., dually enrolled in Medicare and Medicaid), the level of difference is very small. The statistical difference is a function of the very large sample size for this survey.

Table 2-5
Average Health Status and Hospital Use among Cohort 3 Baseline
Medicare Health Outcomes Survey Eligibles, Respondents and Non-respondents,
Selection Probability Weighted[1]

Analytic Variable	Eligibles	Respondents	Non-respondents	Degree of Bias Difference in Means[2]	
Mean PIP-DCG Risk Score	0.90	0.88	0.95	-0.02	*
Mean Number of Hospitalizations	0.20	0.18	0.23	-0.02	*
Mean Number of Inpatient Days	7.17	6.81	7.92	-0.36	*

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of $p<0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS).

Table 2-6
Weighted[1] Average PIP-DCG Score for Eligibles and Respondents by Beneficiary and Enrollment Characteristics, Cohort 3 Baseline Medicare Health Outcomes Survey

Characteristic	Eligibles Mean	Respondents Mean	Degree of Bias Difference in Means[2]
Total	0.90	0.88	-0.02 *
Age			
Under 65	0.83	0.84	0.01
65-74	0.71	0.70	-0.01 *
75-84	1.03	1.01	-0.02 *
85 +	1.36	1.34	-0.03 *
Race			
Unknown	0.86	0.84	-0.02
White	0.89	0.88	-0.01 *
Black	0.94	0.91	-0.03 *
Other	0.79	0.78	-0.01
Asian	0.94	0.90	-0.04 *
Hispanic	0.97	0.95	-0.02
American Indian	0.96	1.01	0.05
Gender			
Male	0.95	0.94	-0.01 *
Female	0.86	0.83	-0.03 *
Medicaid Status			
Not Enrolled	0.88	0.86	-0.02 *
Enrolled	1.35	1.31	-0.04 *
Institutionalized Status			
Community Dwelling	0.89	0.88	-0.01 *
Long-term Institutionalized	1.64	1.56	-0.08
Nursing Home Certifiable	1.28	1.30	0.02
Reason for Medicare Entitlement			
Aged without ESRD	0.90	0.88	-0.02 *
Aged with ESRD	2.11	2.15	0.04
Disabled without ESRD	0.84	0.85	0.01
Disabled with ESRD	0.62	0.70	0.08
ESRD Only	0.80	0.83	0.03

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of $p < 0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS).

Table 2-7
Average Demographic and Health Status Characteristics of Eligibles and Respondents by Decile of Health Plan Response Level to the Cohort 3 Baseline Medicare Health Outcomes Survey, Selection Probability Weighted[1]

Analytic Variables	Level of Health Plan Response				
	31-50%	51-60%	61-70%	71-80%	81-90%
Number of Plans (Total = 306)	4	22	91	159	30
Number of Beneficiaries (Total = 291,221)	3,257	20,704	87,388	151,937	27,935
Demographics					
Average Age					
Eligibles[2]	74	74	74	74	74
Respondents	73	74	74	74	74
Percent Female					
Eligibles	57	56	57	58	58
Respondents	56	57	57	58	58
Percent White					
Eligibles	73	73*	83*	88*	96
Respondents	76	76	85	88	96
Percent Medicaid Enrolled					
Eligibles	7	8*	4*	4*	3*
Respondents	7	7	4	3	2
Percent Aged without ESRD					
Eligibles	93	92	93*	94*	96
Respondents	93	92	94	95	96
Health Status and Use					
Average PIP-DCG Risk Score					
Eligibles	0.94	0.94*	0.90*	0.89*	0.89*
Respondents	0.92	0.92	0.88	0.87	0.87
Average Number of Hospitalizations					
Eligibles	0.28	0.22	0.20*	0.18*	0.20
Respondents	0.26	0.21	0.19	0.18	0.18
Average Number of Inpatient Days					
Eligibles	9.9	7.6	7.2	6.9	7.2
Respondents	9.8	6.9	6.9	6.7	6.7

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents made within the decile of response rate category using a two-sided z-test of differences at the significance level of $p < 0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS).

CHAPTER 3
ANALYSIS OF NON-RESPONSE BIAS IN THE 2000 COHORT 1 FOLLOW-UP MEDICARE HEALTH OUTCOMES SURVEY

3.1   Description of the Medicare Health Outcomes Survey

As described in Chapter 2, the Medicare Health Outcomes Survey (HOS) is a survey administered to Medicare beneficiaries enrolled in Medicare+Choice (M+C) managed care organizations[12] (MCOs) and is used to determine the change in health status over a 2-year time period. The HOS instrument is based on the SF-36®[13] Health Survey, or SF-36, which asks the respondent to rate general health, ability to perform certain physical tasks, level of pain, and emotional state. Summary scales of physical and mental health, denoted as PCS and MCS, are calculated using the SF-36 questions. Two years after administration of the baseline survey, a follow-up survey using a similar instrument is administered. Change in health status at the health plan level is calculated and is provided to health plans for their use as a quality improvement and monitoring tool.

Medicare beneficiaries who are respondents to the Baseline HOS are eligible for re-survey, if they remain enrolled in the same health plan as they were at the time of baseline survey administration. Beneficiaries who die between completion of the baseline survey and follow-up survey administration are considered respondents for purposes of calculating the change in the PCS score; however, they are excluded from analysis of change in the MCS score. As with the baseline survey, some adjustments in calculating numbers of beneficiaries eligible for re-survey are made based on the survey disposition codes accredited during survey administration. The following individuals are declared ineligible ex post:

- those reported deceased (excluded from MCS analysis)

- those unable to complete the survey because of language barriers

- those with bad addresses and non-working telephone numbers

- those not enrolled in the baseline MCO

- those with ESRD

All participating health plans with Medicare contracts in place on or before January 1, 1997, that administered a Cohort 1 Baseline Survey in 1998 were required to administer the Cohort 1 Follow-up Survey. The follow-up survey is administered as a mail survey with telephone follow-up during the same general timeframe as the baseline survey. The definition of a complete survey, as specified for the HOS protocol, for the follow-up survey is 80 percent or more of the total questions answered. As with the baseline survey and for purposes of this

---

[12] HEDIS® 2000 manual, volume 6 (NCQA, 2000), is used for the background information on the HOS Survey.

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

analysis, a respondent is a beneficiary for whom a PCS or MCS score could be calculated. The overall response rate was 85 percent for the Cohort 1 Follow-up Survey. The HOS Follow-up Survey does not use design or non-response weights to adjust survey responses for differential rates of response among survey eligibles.

3.2 Survey-Specific Response Rates

Differences in response rates by beneficiary demographic and enrollment characteristics, health status, and medical care use rates are displayed in **Table 3-1**. As for the Cohort 3 Baseline Survey, three sets of weights are used to present response rates. Pairwise comparisons of differences in selection probability weighted response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. We summarize our findings below, focusing on the response rates in the enrollment weighted column:

- The overall response rate to the Cohort 1 Follow-up HOS was 85 percent. This is a 15 percentage point higher response rate than observed for the 2000 Cohort 3 Baseline HOS administered at the same time as the follow-up survey.

- Response rates vary by all enrollment and demographic characteristics and health status measures evaluated in this survey and in the same general patterns observed for the baseline survey.

  – Compared to beneficiaries age 65 to 74, beneficiaries under the age of 65, or those entitled to Medicare because of disability, and those 85 years of age or older have significantly lower rates of response.

  – Beneficiaries whose race/ethnicity is White or Asian are the most likely to respond to the survey, while Blacks, Hispanics, and American Indians are the least likely to respond.

  – Beneficiaries dually enrolled in Medicare and Medicaid are significantly less likely to respond than beneficiaries not enrolled in Medicaid. The response rate for dual enrollees is 8 percentage points lower than the response rate of beneficiaries not dually enrolled.

  – The response rate for beneficiaries residing in long-term institutionalized facilities is low—only 46 percent responded. This is in contrast to the response rate of 86 percent for beneficiaries residing in the community.

  – Beneficiaries enrolled in Medicare due to disability and without ESRD are significantly less likely to respond to the HOS Survey than aged beneficiaries without ESRD.

Table 3-1
Survey-Specific Response Rates by Demographic and Health Status Characteristics
for the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey

Characteristic	Unweighted Response Rate (%)	Mean of the Means Response Rate (%)	Selection Probability Weighted[1] Response Rate (%)	
Age[2]				
Under 65	80	81	80	*
65-74	87	86	88	
75-84	86	85	86	*
85 +	79	78	79	*
Race				
Unknown	81	79	81	
White	86	85	86	
Black	76	77	76	*
Other	84	84	84	
Asian	88	84	88	
Hispanic	82	81	82	*
American Indian	69	68	69	*
Gender				
Male	85	84	85	
Female	86	85	86	
Medicaid Status				
Not Enrolled	86	85	86	
Enrolled	76	78	78	*
Institutionalized Status				
Community Dwelling	86	85	86	
Long-term Institutionalized	46	41	46	*
Nursing Home Certifiable	75	75	75	*
Reason for Medicare Entitlement				
Aged without ESRD	86	85	86	
Aged with ESRD	83	83	83	
Disabled without ESRD	80	81	80	*
Disabled with ESRD	N/A	N/A	N/A	
ESRD Only	100	100	100	
Risk Score Decile				
0.36 - 0.45	87	86	87	*
0.46 - 0.53	88	86	88	*
0.54 - 0.57	88	87	88	
0.58 - 0.70	87	86	87	*
0.71 - 0.73	87	86	87	*
0.74 - 0.87	85	84	85	
0.88 - 0.91	86	86	86	*
0.92 - 1.07	85	84	85	
1.08 - 1.26	81	80	81	*
1.27 - 6.91	80	80	80	*
Number of Hospitalizations				
Zero	86	85	86	
One	83	83	83	*
Two	81	80	81	*
Three or More	78	78	78	*

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS).

– The response rate for beneficiaries with an average health status score, or PIP-DCG score of 1.0, is 85 percent. The response rate is lower for beneficiaries with poorer health status or higher PIP-DCG scores. In contrast, response rates for beneficiaries with better self-reported health status are higher.

– The response rate for beneficiaries without any hospitalizations during the year prior to re-survey is 86 percent. As exhibited in the 2000 Baseline Survey, response rates decline as numbers of hospitalizations increase.

3.3 Differences in Characteristics of Respondents and Non-respondents

Differences in the selection probability weighted distribution of beneficiary characteristics between respondents and non-respondents using the same stratifying variables displayed in **Table 3-1** are presented in **Table 3-2**. The distributions are weighted using selection probability weights. Statistical significance testing is performed using the chi-square test and p<0.05 level of significance. We summarize our findings below:

- The distribution of beneficiary characteristics and health status scores differs systematically between respondents and non-respondents.

  – A greater proportion of non-respondents are under age 65 or age 85 and older.

  – The proportion of Whites is lower and Blacks is higher in the non-respondent population.

  – The proportion of Medicare and Medicaid dual enrollees is modestly higher for non-respondents.

  – The proportion of Medicare beneficiaries residing in a long-term institution and the proportion of beneficiaries entitled to Medicare because of disability are higher among the non-respondents, albeit the actual percentage point differences are very modest.

  – The distribution of health status risk scores is skewed more toward higher levels of disability among the non-respondents as compared to respondents. Twenty-seven percent of the non-respondents have risk scores 8 percent or higher than the average score. This is in contrast to 20 percent for respondents.

  – There are also modestly more hospitalizations among non-respondents as compared to respondents.

3.4 Characteristics of Respondents and Non-respondents at Follow-up using Baseline Characteristics

An alternative way of evaluating non-response is to look at response rates based on baseline characteristics. We evaluate whether the distribution of beneficiary characteristics differs systematically between follow-up respondents and follow-up non-respondents, both alive and deceased, at time of follow-up. This latter comparison is important as decedents are handled

Table 3-2
Distribution of Demographic and Health Status Characteristics among Cohort 1
Follow-up Medicare Health Outcomes Survey Eligibles, Respondents, and
Non-respondents, Selection Probability Weighted[1]

Characteristic	Eligibles (%)	Respondents (%)	Non-respondents (%)	
Age[2]				
Under 65	5	4	6	*
65-74	46	47	40	
75-84	39	39	38	
85 +	11	10	15	
Race				
Unknown	0.3	0.3	0.4	*
White	88	89	83	
Black	7	6	12	
Other	2	2	2	
Asian	1	1	1	
Hispanic	2	2	2	
American Indian	0.1	0.1	0.1	
Gender				
Male	42	42	43	
Female	58	58	57	
Medicaid Status				
Not Enrolled	96	96	93	*
Enrolled	4	4	7	
Institutionalized Status				
Community Dwelling	99	99	98	*
Long-term Institutionalized	0.4	0.2	2	
Nursing Home Certifiable	0.4	0.3	1	
Reason for Medicare Entitlement				
Aged without ESRD	95	95	93	*
Aged with ESRD	0	0	0	
Disabled without ESRD	5	5	7	
Disabled with ESRD	0	0	0	
ESRD Only				
Risk Score Decile				
0.36 - 0.45	10	10	9	*
0.46 - 0.53	7	8	6	
0.54 - 0.57	12	13	10	
0.58 - 0.70	13	13	11	
0.71 - 0.73	11	11	10	
0.74 - 1.87	5	5	5	
0.88 - 0.91	13	13	12	
0.92 - 1.07	9	9	9	
1.08 - 1.26	10	10	13	
1.27 - 6.91	10	10	14	
Number of Hospitalizations				
Zero	86	87	84	*
One	10	9	11	
Two	3	2	3	
Three or More	1	1	2	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Statistical significance tested using chi-square distribution differences between respondents and non-respondents. An asterisk (*) denotes significance at <0.05 level.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS).

differently in the calculation of the PCS and MCS. A clearer understanding of the types of beneficiaries who are non-respondents at follow-up because of death may help to guide interpretation and use of the change in MCS and PCS scores for quality improvement. Differences in rates and proportions across strata are evaluated for statistical significance by use of the chi-square test of differences for categorical data at the $p<0.05$ significance level. Three sets of statistical comparisons are made between (A) respondents and live non-respondents, (B) respondents and deceased non-respondents, and (C) live and deceased non-respondents.

- There are significant differences in the types of beneficiaries who respond at follow-up versus those that did not respond, and there are significant differences in the types of beneficiaries who did not respond due to death prior to follow-up versus those that were alive at time of follow-up (**Table 3-3**).

    – 71,094 beneficiaries who completed a baseline survey responded at follow-up—a response rate of 85 percent.

    – Of the 19,578 non-respondents, roughly 60 percent were alive at the time of re-survey and 40 percent were deceased.

    – Older beneficiaries, beneficiaries from racial minorities or dually enrolled in Medicare and Medicaid at baseline, and beneficiaries residing in institutions or who were nursing home certifiable at baseline are more likely to be non-respondents at follow-up.

    – Among non-respondents, the mean age at baseline is roughly 3 years younger for live non-respondents as compared to those that were deceased at the time of follow-up. Among non-respondents, males are more likely to be non-respondents due to death as compared to females, and White beneficiaries are more likely to be non-respondents due to death than racial minorities. Beneficiaries who at baseline were enrolled in Medicare because of disability or were aged with ESRD are more likely to be non-respondents due to death as compared to aged beneficiaries without ESRD.

- Respondents at follow-up had better health status at baseline than non-respondents as measured by PCS and MCS scores as well as a general health status question at time of baseline survey and as compared with a prior year (**Table 3-4**). Not surprisingly, respondents alive at follow-up were also healthier at baseline than non-respondents who died prior to follow-up.

    – Mean baseline PCS is 44 for respondents as compared to 43 for live non-respondents and 35 for deceased non-respondents.

    – Mean baseline MCS is 53 for respondents as compared to 51 for live non-respondents and 47 for deceased non-respondents.

    – Roughly 25 percent of respondents at follow-up reported being in fair or poor health at baseline. This is in contrast to 31 percent of live non-respondents and 58 percent of deceased non-respondents.

50

Table 3-3
2000 Cohort 1 Follow-up Medicare Health Outcomes Survey Eligibles:
Baseline Survey Demographic Characteristics by Follow-up Response Status

Characteristic	Follow-up Respondents	Follow-up Alive Non-respondents	Follow-up Decedents	Statistical Testing[1,2]		
	N=71,094	N=11,531	N=8,047			
Age (mean)	73.9	74.9	78.4	A	B	C
Race	%	%	%			
White	89.3	83.3	88.9	A	B	C
Black	5.9	10.6	7.3			
Other	1.5	1.8	1.3			
Asian	1.3	1.3	0.8			
Hispanic	1.6	2.6	1.4			
North American Native	0.1	0.1	0.1			
Unknown	0.3	0.4	0.4			
Gender						
Male	41.8	41.9	50.9		B	C
Female	58.2	58.1	49.1			
Institutionalized Status						
Community Dwelling	99.7	99.1	97.3	A	B	
Long-term Institutionalized	0.1	0.5	1.9			
Nursing Home Certifiable	0.2	0.4	0.8			
Medicaid Status						
Enrolled	2.23	4.1	6.1	A	B	C
Not Enrolled	97.8	95.9	93.9			
Reason for Entitlement						
Aged without ESRD	99.9	99.9	99.7		B	C
Aged with ESRD	0	0	0.1			
Disabled without ESRD	0.1	0.1	0.2			

[1] Differences in rates and proportions across strata are evaluated for statistical significance by use of chi-square test for categorical data and a two-sided z-test for continuous data at p<0.05 level of significance.
A: statistically significant difference between respondents and alive non-respondents
B: statistically significant difference between respondents and deceased non-respondents
C: statistically significant difference between alive and deceased non-respondents
[2] Data are unweighted.

Source: RTI analysis of 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS) and 1998 Cohort 1 Baseline HOS.

Table 3-4
2000 Cohort 1 Follow-up Medicare Health Outcomes Survey Eligibles:
Baseline Health Status Characteristics by Follow-up Response Status

Characteristic	Follow-up Respondents	Follow-up Alive Non-respondents	Follow-up Decedents	Statistical Testing[1,2]		
	N=71,094	N<11,531	N=8,047			
Baseline PCS Score (mean)	43.96	42.69	34.93	A	B	C
Baseline MCS Score (mean)	52.65	50.83	46.57	A	B	C
General Health	%	%	%	A	B	C
Excellent	6.6	6.4	2.0			
Very good	27.4	22.2	10.9			
Good	42.1	40.4	28.7			
Fair	20.7	25.1	36.0			
Poor	3.2	5.9	22.3			
General Health Compared to Year Ago				A	B	C
Much better	4.8	6.1	3.9			
Somewhat better	10.7	10.1	9.4			
About the same	68.9	65.1	47.4			
Somewhat worse	13.9	15.6	26.8			
Much worse	1.7	3.1	12.5			

[1] Differences in rates and proportions across strata are evaluated for statistical significance by use of chi-square test for categorical data and a two-sided z-test for continuous data at $p<0.05$ level of significance.
A: statistically significant difference between respondents and alive non-respondents
B: statistically significant difference between respondents and deceased non-respondents
C: statistically significant difference between alive and deceased non-respondents
[2] Data are unweighted.

Source: RTI analysis of 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS) and 1998 Cohort 1 Baseline HOS.

- Deceased non-respondents were far more likely to have reported at baseline that they were in worse health than they had been during the prior year.

- Follow-up respondents were less likely to report functional status limitations across a broad array of activities at baseline than non-respondents alive at time of follow-up and considerably less likely than non-respondents not alive at time of follow-up (**Table 3-5**).

  - Over 70 percent of decedents at time of follow-up reported being limited a lot in doing vigorous activities at time of baseline. This is in contrast to 49 percent of non-respondents alive at follow-up and 47 percent of follow-up respondents.

  - Over one-half of decedents at time of follow-up reported being limited a lot in climbing several stairs at time of baseline. This is in contrast to 31 percent of non-respondents alive at follow-up and 25 percent of follow-up respondents.

  - And, 35 percent of decedents at time of follow-up reported being limited a lot in walking one block at time of baseline. This is in contrast to 14 percent of non-respondents alive at follow-up and 9 percent of follow-up respondents.

- Similarly, follow-up respondents were less likely to report limitations across all types of activities of daily living (ADL) at baseline than non-respondents alive at time of follow-up and considerably less likely than non-respondents not alive at time of follow-up (**Table 3-6**).

  - The mean number of ADL limitations was 0.8 at baseline for follow-up respondents; the mean number of ADL limitations was 1.1 and 2.2 at baseline for living and deceased non-respondents, respectively.

  - Thirteen percent of deceased respondents at baseline reported being unable to bathe. Less than 2 percent of respondents reported being unable to bathe at baseline.

  - Almost 10 percent of deceased respondents at baseline reported being unable to walk. Once again, less than 2 percent of respondents reported being unable to walk at baseline.

- Baseline respondents who were deceased at the time of follow-up were far more likely to self-report the presence of at least 1 of 17 chronic conditions at baseline as compared to respondents and live non-respondents (**Table 3-7**). There are only limited differences in the proportion of respondents and living non-respondents reporting the presence of specific chronic conditions.

  - Twenty-six percent of deceased non-respondents reported angina or coronary artery disease at baseline. About 15 percent of respondents and living non-respondents reported the same condition at baseline.

53

Table 3-5
2000 Cohort 1 Follow-up Medicare Health Outcomes Survey Eligibles:
Baseline Survey Functional Status Limitations by Follow-up Response Status

Characteristic	Follow-up Respondents	Follow-up Alive Non-respondents	Follow-up Decedents	Statistical Testing[1,2]		
	N=71,094	N=11,531	N=8,047			
Vigorous Activities	%	%	%	A	B	C
Limited a lot	46.8	49.0	71.9			
Limited a little	38.7	34.6	18.5			
Not limited at all	14.4	16.4	9.6			
Moderate Activities				A	B	C
Limited a lot	16.2	23.0	48.5			
Limited a little	34.8	34.1	29.4			
Not limited at all	49.1	43.0	22.1			
Lifting Groceries				A	B	C
Limited a lot	11.0	17.3	39.2			
Limited a little	29.7	31.3	32.6			
Not limited at all	59.3	51.4	28.2			
Climbing Several Stairs				A	B	C
Limited a lot	25.2	30.8	57.0			
Limited a little	37.0	35.2	25.9			
Not limited at all	37.8	33.9	17.1			
Climbing 1 Flight of Stairs				A	B	C
Limited a lot	11.0	16.7	37.8			
Limited a little	27.1	28.9	32.1			
Not limited at all	61.9	54.4	30.1			
Bending and Kneeling				A	B	C
Limited a lot	19.0	23.3	41.9			
Limited a little	42.3	41.5	36.8			
Not limited at all	38.7	35.1	21.2			
Walking More than 1 Mile				A	B	C
Limited a lot	30.5	37.0	64.8			
Limited a little	29.2	27.6	18.4			
Not limited at all	40.3	35.4	16.8			
Walking Several Blocks				A	B	C
Limited a lot	20.8	26.9	54.4			
Limited a little	24.0	25.4	21.9			
Not limited at all	55.2	47.7	23.8			
Walking One Block						
Limited a lot	9.3	14.1	35.6	A	B	C
Limited a little	19.2	22.1	27.9			
Not limited at all	71.4	63.9	36.6			

[1] Differences in rates and proportions across strata are evaluated for statistical significance by use of chi-square test for categorical data and a two-sided z-test for continuous data at p<0.05 level of significance.
A: statistically significant difference between respondents and alive non-respondents
B: statistically significant difference between respondents and deceased non-respondents
C: statistically significant difference between alive and deceased non-respondents
[2] Data are unweighted.

Source: RTI analysis of 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS) and 1998 Cohort 1 Baseline HOS.

Table 3-6
2000 Cohort 1 Follow-up Medicare Health Outcomes Survey Eligibles:
Baseline ADL Limitations by Follow-up Response Status

Characteristic	Follow-up Respondents	Follow-up Alive Non-respondents	Follow-up Decedents	Statistical Testing[1,2]		
	N=71,094	N=11,531	N=8,047			
Mean # of ADL Limitations	0.79	1.06	2.23	A	B	C
Bathing	%	%	%	A	B	C
Unable to do	1.5	3.5	12.8			
Have difficulty	8.6	12.7	26.5			
No difficulty	89.9	83.9	60.7			
Dressing				A	B	C
Unable to do	1.1	2.4	9.1			
Have difficulty	6.9	10.8	24.0			
No difficulty	92.0	86.8	66.9			
Eating				A	B	C
Unable to do	0.8	1.2	3.3			
Have difficulty	3.2	5.5	15.3			
No difficulty	96.0	93.4	81.5			
Transferring				A	B	C
Unable to do	0.9	1.9	6.4			
Have difficulty	22.0	25.3	40.6			
No difficulty	77.2	72.9	53.0			
Walking				A	B	C
Unable to do	1.6	3.1	9.5			
Have difficulty	28.2	32.3	52.8			
No difficulty	70.2	64.6	37.7			
Using the Toilet				A	B	C
Unable to do	0.8	1.5	5.5			
Have difficulty	5.1	8.1	18.1			
No difficulty	94.1	90.4	76.4			
Difficulty controlling urination	24.8	26.2	36.0	A	B	C

[1] Differences in rates and proportions across strata are evaluated for statistical significance by use of chi-square test for categorical data and a two-sided z-test for continuous data at p<0.05 level of significance.
A: statistically significant difference between respondents and alive non-respondents
B: statistically significant difference between respondents and deceased non-respondents
C: statistically significant difference between alive and deceased non-respondents
[2] Data are unweighted.

Source: RTI analysis of 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS) and 1998 Cohort 1 Baseline HOS.

Table 3-7
2000 Cohort 1 Follow-up Medicare Health Outcomes Survey Eligibles:
Baseline Survey Self-Reported Chronic Conditions by Follow-up Response Status

Characteristic	Follow-up Respondents	Follow-up Alive Non-respondents	Follow-up Decedents	Statistical Testing[1,2]		
	N=71,094 %	N=11,531 %	N=8,047 %			
Hypertension	52.0	53.1	57.9	A	B	C
Angina or CAD	14.8	15.2	25.6		B	C
CHF	5.7	7.5	20.8	A	B	C
AMI or heart attack	9.4	10.4	19.6	A	B	C
Other heart conditions	20.0	19.4	31.4		B	C
Stroke	6.7	9.6	18.1	A	B	C
Emphysema, asthma, or COPD	11.4	11.7	22.3		B	C
Cron's disease, colitis, or other GI	5.1	5.2	7.4		B	C
Arthritis (hip or knee)	37.4	37.9	40.3		B	C
Arthritis (hand or wrist)	33.6	34.3	36		B	C
Sciatica	21.5	21.5	22.7		B	
Diabetes	15.1	17.7	23.8	A	B	C
Cancer other than skin	12.5	12	24		B	C
Colon cancer	2.8	3.8	6.9	A	B	C
Lung cancer	1.0	1.5	8.5	A	B	C
Breast cancer	5.5	5.1	6.4		B	C
Prostate cancer	7.9	7.9	11.8		B	C

[1] Differences in rates and proportions across strata are evaluated for statistical significance by use of chi-square test for categorical data and a two-sided z-test for continuous data at p<0.05 level of significance.
A: statistically significant difference between respondents and alive non-respondents
B: statistically significant difference between respondents and deceased non-respondents
C: statistically significant difference between alive and deceased non-respondents
[2] Data are unweighted.

Source: RTI analysis of 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS) and 1998 Cohort 1 Baseline HOS.

- Twenty-one percent of deceased non-respondents reported congestive heart failure at baseline, while 6 and 8 percent of respondents and living non-respondents, respectively, reported the same condition at baseline.

- There is also a stark difference in the proportion of respondents, living non-respondents, and deceased non-respondents that reported lung cancer at baseline, reflecting the high rate of mortality associated with the disease.

3.5 Differences in Outcomes by Demographic and Health Status Characteristics

Next, we explore differences in survey-specific outcome scores by beneficiary demographic and enrollment characteristics, health status and medical care use rates. For the HOS, we report mean PCS and MCS scores at the time of follow-up as the outcome measure. Pairwise comparisons of differences in selection probability weighted response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. Our findings are summarized as follows:

- Across all respondents, the mean PCS score is 41.32 (**Table 3-8**). This is about 9 percentage points lower than the norm-based mean of the 50 for the general population indicating that respondents to the Cohort 1 Follow-up HOS have, in general, a lower level of physical health than the general population.

- Mean PCS scores differed substantially across respondents based on sociodemographic characteristics and levels of health status; however, the patterns are remarkably similar to those observed for the Cohort 3 Baseline HOS.

  - Compared to beneficiaries age 65 to 74, beneficiaries of all other age groups have lower physical functioning.

  - Compared to Whites, Blacks and Hispanics have lower mean PCS scores, while Asians and beneficiaries with race coded as "other" have higher mean PCS scores.

  - Men self-report a higher level of physical health than women. However, men were modestly more likely to die than women between baseline and follow-up. Thus, in estimation of PCS change scores between the two survey waves, one would probably not see men with better physical health status because death is considered a worse outcome.

  - Beneficiaries dually enrolled in Medicare and Medicaid self-report a significantly lower level of physical health (PCS of 34.27) than beneficiaries not also enrolled in Medicaid (PCS of 41.58).

  - Beneficiaries residing in long-term facilities and beneficiaries residing in the community who were deemed nursing home certifiable report significantly lower levels of physical functioning than community residents.

57

Table 3-8
Average Physical and Mental Health Component Scores by Demographic and Health Status Characteristics of Respondents to the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey, Selection Probability Weighted[1]

Characteristic	Physical Health Component Score (PCS) (mean)		Mental Health Component Score (MCS) (mean)	
Across all Respondents[2]	41.32		50.97	
Age				
Under 65	31.76	*	41.25	
65-74	43.97		52.49	
75-84	40.57	*	50.86	*
85 +	35.97	*	48.56	*
Race				
Unknown	42.64		50.72	
White	41.43		51.19	
Black	39.20	*	48.99	*
Other	43.01	*	50.13	*
Asian	43.59	*	52.03	
Hispanic	39.80	*	47.08	
American Indian	38.46		48.14	*
Gender				
Male	42.25		51.30	
Female	40.63	*	50.73	*
Medicaid Status				
Not Enrolled	41.58		51.21	
Enrolled	34.27	*	44.51	*
Institutionalized Status				
Community Dwelling	41.38		51.02	
Long-term Institutionalized	29.02	*	39.45	*
Nursing Home Certifiable	30.28	*	43.76	*
Reason for Medicare Entitlement				
Aged without ESRD	41.79		51.44	
Aged with ESRD	29.47		56.56	
Disabled without ESRD	31.79	*	41.62	*
Disabled with ESRD	N/A		N/A	
ESRD Only	57.70		60.31	
Risk Score Decile				
0.36 - 0.45	44.28	*	51.57	*
0.46 - 0.53	46.88	*	54.02	*
0.54 - 0.57	43.85	*	52.59	*
0.58 - 0.70	44.29	*	52.53	*
0.71 - 0.73	41.41	*	51.32	*
0.74 - 0.87	38.63		48.81	*
0.88 - 0.91	41.19	*	51.36	*
0.92 - 1.07	38.95		49.80	
1.08 - 1.26	36.20	*	48.64	*
1.27 - 6.91	34.86	*	47.14	*
Number of Hospitalizations				
Zero	42.07		51.37	
One	37.19	*	48.96	*
Two	34.92	*	47.27	*
Three or More	32.95	*	46.16	*

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS).

- Compared to beneficiaries with average health status (PIP-DCG score range from 0.92 to 1.07), beneficiaries with better health status (lower scores) generally have higher average PCS scores and beneficiaries with worse health status (higher scores) have lower PCS scores.

- There is also an observed negative relationship with PCS scores and number of hospitalizations; as frequency of hospitalization increases, one observes a declining average PIP-DCG score.

• Across all respondents, the mean MCS score is 50.97, indicating that respondents to the Cohort 1 Follow-up HOS have, in general, a similar level of self-reported mental health as the general population (mean of 50).

• Mean MCS scores also differ substantially across respondents based on sociodemographic characteristics and levels of health status.

- A set of patterns similar to those observed for PCS scores is generally seen with the MCS scores for the demographic characteristics of age, race, gender, Medicaid enrollment, institutionalized status, health status, and number of hospitalizations.

- The disabled without ESRD have lower self-reported mental health than beneficiaries aged without ESRD.

- Beneficiaries with the highest health status (lowest PIP-DCG scores) have MCS scores higher than beneficiaries with average health status. Beneficiaries with poor health status as measured by the PIP-DCG score have the lowest MCS scores.

3.6   Factors that Predict Likelihood of Response

We predict the likelihood of response as a function of sociodemographic and health status characteristics of all sampled beneficiaries using a multivariate regression model. We estimate the model unweighted and weighted by the inverse likelihood of the beneficiary being selected for survey.

• The direction and magnitude of the odds ratios for the included beneficiary-level variables are consistent with the descriptive comparisons between respondents and non-respondents (**Table 3-9**). However, the magnitude of effect of the predictor variables collectively is lower than observed in the Cohort 3 Baseline HOS as measured by the chi-square statistic.

- Beneficiaries under the age of 65 and age 85 and older are more than 25 percent less likely to respond than beneficiaries age 65 to 74.

- All minority races other than Asians are far less likely than White beneficiaries to respond.

Table 3-9
Logistic Regression of Likelihood of Response to the 2000 Cohort 1
Follow-up Medicare Health Outcomes Survey

Characteristic	Unweighted Regression Odds Ratio[1]	Selection Probability Regression Odds Ratio[2]
Beneficiary Characteristics		
Under 65	0.661	0.661
75 to 84	0.961	0.961
85 +	0.720	0.720
Black	0.508	0.508
Unknown or Other Race	0.824	0.824
Asian	1.151	1.151
Hispanic	0.731	0.731
American Indian	0.410	0.410
Male	0.993	0.993
Medicaid	0.816	0.816
ESRD	1.179	1.179
Institutionalized Status		
Long-term Institutionalized	0.197	0.197
Nursing Home Certifiable	0.616	0.616
Risk Score Decile		
0.36 - 0.45	1.206	1.206
0.46 - 0.53	1.159	1.159
0.54 - 0.57	1.240	1.240
0.58 - 0.70	1.123	1.123
0.71 - 0.73	1.134	1.134
0.74 - 0.87	1.144	1.144
0.88 - 0.91	1.067	1.067
1.08 - 1.26	0.921	0.921
1.27 - 6.91	0.840	0.840
Number of Hospitalizations		
One	0.977	0.977
Two	0.931	0.931
Three or More	0.805	0.805
No. of Observations	88129	88129
Overall Chi-Sq (p-value)	1720***	1720***

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Highlighted odds ratios are statistically significant at the p<0.05 level of significance. Asterisks (***) denote p<0.001 level of significance.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS).

- Beneficiaries dually enrolled in Medicare and Medicaid are about 20 percent less likely to respond than beneficiaries not also enrolled in Medicaid.

- The long-term institutionalized are 80 percent less likely to respond to the follow-up HOS as compared to community residing beneficiaries. Beneficiaries who are nursing home certifiable are about 40 percent less likely to respond than community residents. At baseline, nursing home certifiable beneficiaries were 12 percent more likely to respond than community residents.

- Compared to beneficiaries with average health status, those with lower PIP-DCG scores, which equates to a higher level of health status, are generally more likely to respond to the Follow-up HOS. Beneficiaries in the poorest health, or with the highest PIP-DCG scores, are less likely to respond than beneficiaries with average health status.

- The likelihood of response to the Follow-up HOS declines only for those beneficiaries who were hospitalized three or more times during the year prior to re-survey.

3.7   Probable Degree of Non-response Bias

We directly explore the degree of bias that may be present in estimates of health status and medical care usage by comparing means of these variables for respondents to those obtained for eligible beneficiaries, including non-respondents. We also report differences in mean PIP-DCG scores between respondents and survey eligibles stratified by sociodemographic and medical care usage characteristics. We also examine the differences between eligibles and respondents by plan response rate deciles to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles.

Pairwise comparisons of differences in selection probability weighted mean estimates between eligibles and respondents are made using a two-sided z-test for differences in means or proportions at the significance level of $p<0.01$ to account for multiple comparisons. We consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

- Differences in the mean health status and mean number of hospitalizations between eligibles and respondents reflect the differences previously observed in the underlying distribution of characteristics of respondents and non-respondents (**Tables 3-1 and 3-2**), suggesting that respondents, on average, have a modestly higher level of health status than the surveyed population (**Table 3-10**).

  - The mean PIP-DCG score is only 1 percent lower for respondents than for survey eligibles, implying modest non-response bias, which overstates average health status of the survey population.

  - Mean number of hospitalizations is also modestly lower for respondents than for survey eligibles.

61

Table 3-10
**Average Health Status and Hospital Use among 2000 Cohort 1 Follow-up Medicare
Health Outcomes Survey Eligibles, Respondents, and Non-respondents,
Selection Probability Weighted[1]**

Analytic Variable	Eligibles	Respondents	Non-respondents	Degree of Bias Difference in Means[2]	
Mean PIP-DCG Risk Score	0.91	0.90	1.00	-0.01	*
Mean Number of Hospitalizations	0.20	0.19	0.26	-0.01	*
Mean Number of Inpatient Days	7.07	6.78	8.39	-0.29	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS).

– We observe no statistical difference in mean number of inpatient days.

- Differences in the mean health status between survey eligibles and respondents display a more modest general trend than observed for the Cohort 3 Baseline HOS; however, health status estimates derived using the PIP-DCG risk score continue to be lower (better health) than those derived for survey eligibles but for fewer subpopulations of Medicare beneficiaries (**Table 3-11**). As with the baseline survey, health status estimates derived from respondents at follow-up tend to modestly overestimate the health of M+C Medicare enrollees. And, this overestimation tends to be for several of the healthier subpopulations (e.g, aged without ESRD and beneficiaries not dually enrolled in Medicare and Medicaid).

- As with the analysis of the Cohort 3 Baseline HOS, a comparison of the differences between eligibles and respondents by plan response rate deciles does not suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles (**Table 3-12**). One interesting note is that health plans with low response rates tend to have very high rates of Medicare and Medicaid dual enrollees at follow-up when compared to plans with higher response rates. This could reflect movement into nursing homes and spending down of assets. Once again, we observe statistically significant differences between eligibles and respondents for some subpopulations, but the actual difference in proportions or means is very small. The statistical difference is a function of the large sample size for this survey.

Table 3-11

Average PIP-DCG Score for Eligibles and Respondents by Beneficiary and Enrollment Characteristics, 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey, Selection Probability Weighted[1]

Characteristic	Eligibles Mean	Respondents Mean	Degree of Bias Difference in Means[2]	
Total	0.91	0.90	-0.01	*
Age				
Under 65	0.87	0.86	-0.01	
65-74	0.73	0.72	-0.01	
75-84	1.02	1.01	-0.01	*
85 +	1.35	1.33	-0.02	
Race				
Unknown	0.88	0.85	-0.03	
White	0.91	0.90	-0.01	*
Black	0.96	0.94	-0.02	
Other	0.82	0.81	-0.01	
Asian	0.92	0.92	0.00	
Hispanic	0.94	0.93	-0.01	
American Indian	1.00	0.97	-0.03	
Gender				
Male	0.98	0.97	-0.01	
Female	0.87	0.85	-0.02	*
Medicaid Status				
Not Enrolled	0.89	0.88	-0.01	*
Enrolled	1.40	1.38	-0.02	
Institutionalized Status				
Community Dwelling	0.91	0.90	-0.01	*
Long-term Institutionalized	1.62	1.60	-0.02	
Nursing Home Certifiable	1.37	1.36	-0.01	
Reason for Medicare Entitlement				
Aged without ESRD	0.91	0.90	-0.01	*
Aged with ESRD	1.75	2.01	0.26	
Disabled without ESRD	0.89	0.87	-0.02	
Disabled with ESRD	NA	NA	NA	
ESRD Only	0.80	0.45	-0.35	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS).

Table 3-12
Average Demographic and Health Status Characteristics of Eligibles and Respondents by Decile of Health Plan Response Level to the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey, Selection Probability Weighted[1]

	Level of Health Plan Response			
Analytic Variables	41-70%	71-80%	81-90%	91-100%
Number of Plans (Total = 225)	5	45	158	17
Number of Beneficiaries (Total=88,129)	883	13,823	65,341	8,082
Demographics				
Average Age[2]				
Eligibles	73	75	75	75
Respondents	73	75	75	75
Percent Female				
Eligibles	59	58	57	58
Respondents	58	58	58	58
Percent White				
Eligibles	68	79*	89*	96
Respondents	70	81	90	96
Percent Medicaid Enrolled				
Eligibles	21*	6	4*	2
Respondents	16	6	3	2
Percent Aged without ESRD				
Eligibles	88	93	95	97
Respondents	88	94	95	97
Health Status and Use				
Average PIP-DCG Risk Score				
Eligibles	1.00	.94*	.91*	0.90
Respondents	0.94	0.92	0.90	0.89
Average Number of Hospitalizations				
Eligibles	0.22	0.22	0.20	0.20
Respondents	0.20	0.20	0.19	0.19
Average Number of Inpatient Days				
Eligibles	7.2	7.9	6.9	7.1
Respondents	6.5	7.5	6.7	6.6

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents made within the decile of response rate category using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS).

CHAPTER 4
ANALYSIS OF NON-RESPONSE BIAS IN THE 2000 CAHPS® MEDICARE+CHOICE (M+C) ENROLLEE SURVEY

4.1   Description of the CAHPS® M+C Enrollee Survey

The Consumer Assessment of Health Plans Study (CAHPS®) is a set of survey and reporting formats developed by a consortium of researchers at Harvard Medical School, RAND, and RTI, and is sponsored by AHRQ and CMS. This effort produced a family of surveys that target specific populations, such as Medicaid and Medicare beneficiaries, consumers of health care in managed care and fee-for-service settings, as well as adults and children. This chapter reports on our examination of potential non-response bias to the CAHPS® Medicare+Choice (M+C) Enrollee Survey.

The CAHPS® M+C Enrollee Survey is an annual survey conducted by CMS to assess the experience of Medicare beneficiaries enrolled in Medicare+Choice organizations (MCOs). It was developed to capture critical information about the quality of care in MCOs, such as overall ratings for each health plan, ease of getting needed care and specialist referrals, and ratings on how well doctors communicate.

The sample design for the survey was developed to allow CAHPS® outcomes to be compared between plans, as well as between Medicare managed care and Original Fee-for-Service (FFS) Medicare. Each Medicare managed care plan comprised a reporting unit. In cases where a contract covered a wide geographic area, some plans (reporting units) were further defined by geographic location. Thus, a single plan with wide geographic coverage in a large state might have multiple reporting units. Within each reporting unit, a simple random sample was drawn of plan enrollees who had been enrolled in the plan for 6 months or longer. Eligible plans for the 2000 survey administration included all M+C organizations and continuing cost contracts with contracts in effect as of July 1, 1999.

To be eligible for sample selection, beneficiaries had to have been enrolled in a selected MCO, had to have had at least 6 months of continuous coverage, and could not have been institutionalized at the time of sample selection. Approximately 600 beneficiaries were sampled from each organization selected to participate in the survey. In MCOs with fewer than 600 Medicare beneficiaries, all beneficiaries were selected for survey. Beneficiaries who switched from their MCO between sampling and survey administration were later excluded from the denominator in calculating response rates. In addition, some adjustment in calculating the number of eligibles was made by the CAHPS® survey team based on the survey disposition codes. The following individuals were declared non-eligible ex post:

- those reported deceased

- those institutionalized

- those who switched managed care plans

- those with bad addresses and non-working telephone numbers

The survey was a telephone survey with mail follow-up. Numerous attempts were made to reach the beneficiaries. Multiple mailing attempts included Federal Express and Priority Mail service. The survey instrument was available in both English and Spanish. The CAHPS® survey team employed the following definition of a complete survey: if 10 or more key questions were answered, then the questionnaire was counted as complete.

The final adjusted response rate is the number of complete questionnaires divided by the number of eligibles minus deceased and those later found to be ineligible. The 2000 sample frame consisted of 225,171 Medicare beneficiaries of whom 216,919 beneficiaries were deemed eligible. The survey response rate for the 2000 CAHPS® M+C Enrollee Survey was 83 percent.

4.2   Survey-Specific Response Rates

We begin our detailed examination of possible non-response bias in the CAHPS® M+C Enrollee Survey by first exploring differences in response rates by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. **Table 4-1** displays three sets of response rates, with alternative weights applied. Pairwise comparisons of differences in selection probability weighted response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. We summarize our findings below, focusing on the response rates in the enrollment weighted column:

- Looking across the three sets of response rates for each category of the characteristics shows the rates to be fairly similar to one another, usually within 1 or 2 percentage points.

- With very few exceptions, the selection probability weighted distribution of response rates differs significantly by category of enrollment, demographic, health status, and use measures.

  – The response rates of beneficiaries under the age of 65 and above 74 years are significantly lower than those for beneficiaries 65 to 74 years of age. This finding is particularly true for the youngest and oldest age groups.

  – The response rates for Blacks and those with a race/ethnicity code of "other" are significantly lower than for Whites. Beneficiaries of Hispanic and American Indian race/ethnicity have response rates that are significantly higher than Whites.

  – The response rate for males is significantly lower than for females, although the difference is small, only 1 percentage point.

  – Beneficiaries dually enrolled in Medicare and Medicaid have a significantly lower response rate than those not dually enrolled.

  – Beneficiaries entitled to Medicare because they are disabled (without ESRD) or aged with ESRD only responded at a significantly lower rate than aged beneficiaries without ESRD.

68

Table 4-1

Survey-Specific Response Rates by Demographic and Health Status Characteristics for the CAHPS® M+C Enrollee Survey

Characteristic	Unweighted Response Rate[1] (%)	Mean of the Means Response Rate[2] (%)	Selection Probability Weighted Response Rate[3] (%)	
Age[4]				
Under 65	77	78	77	*
65-74	86	85	84	
75-84	83	82	82	*
85 +	72	72	70	*
Race				
Unknown	14	13	15	*
White	84	82	82	
Black	76	77	74	*
Other	11	13	11	*
Asian	81	82	82	
Hispanic	90	93	89	*
American Indian	96	95	92	*
Gender				
Male	82	82	81	
Female	83	83	82	*
Medicaid Status				
Not Enrolled	84	83	82	
Enrolled	68	70	68	*
Reason for Medicare Entitlement				
Aged without ESRD	83	83	82	
Aged with ESRD	70	73	69	*
Disabled without ESRD	77	78	77	*
Disabled with ESRD	79	80	82	
ESRD Only	81	84	83	
Risk Score Decile				
0.36 - 0.45	87	86	86	*
0.46 - 0.53	85	84	84	*
0.54 - 0.57	87	86	85	*
0.58 - 0.70	85	85	84	*
0.71 - 0.73	85	84	84	*
0.74 - 0.87	85	85	84	*
0.88 - 0.91	83	83	82	*
0.92 - 1.07	81	81	80	
1.08 - 1.26	75	75	73	*
1.27 - 6.91	75	75	73	*
Number of Hospitalizations				
Zero	83	83	82	
One	81	80	80	*
Two	80	80	79	*
Three or More	74	74	73	*

[1] An equal weight whereby all *sampled beneficiaries* are given a weight of 1.

[2] An equal weight whereby all *health plans or states* are given a weight of 1.

[3] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[4] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes these comparisons that exceed the significance level. The reference stratum with each set of characteristics is in bold.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey.

– Beneficiaries with a PIP-DCG risk score above the category containing average health status (PIP-DCG score of 1.0), or in poorer health status, have significantly lower response rates than those with average health status. Response rates increase as health status improves (i.e., declining risk score).

– Response rates decline as number of hospitalizations in the year prior to survey increases. We use hospitalizations as a proxy for health status.

4.3   Differences in Characteristics of Respondents and Non-respondents

We explore differences in the selection probability weighted distribution of beneficiary characteristics between respondents and non-respondents. This review provides an overall sense of how different respondents are from non-respondents in terms of demographic and health status characteristics and is a critical factor in the determination of potential non-response bias. Statistical significance testing is performed using the chi-square test and $p<0.05$ level of significance. We summarize our findings below:

- The selection probability weighted distribution of response rates differs significantly by category of enrollment, demographic, health status, and use measures (**Table 4-2**).

  – Respondents differ by age from non-respondents, with disproportionately more beneficiaries age 65 to 74 responding, and fewer responding in the disabled under 65 group and in the 85 and older age group.

  – Respondents also differ by race/ethnicity from non-respondents, with disproportionately more White beneficiaries responding and fewer Black beneficiaries responding.

  – With respect to gender, respondents are modestly more likely to be female.

  – Beneficiaries dually enrolled in Medicare and Medicaid are less likely to respond than beneficiaries not also enrolled in Medicaid.

  – Beneficiaries entitled to Medicare exclusively because they are age 65 or older are more likely to respond than those entitled to Medicare for other reasons.

  – When rates of response are arrayed according to health status as measured by PIP-DCG score deciles, respondents are seen to be healthier (PIP-DCG scores below 1.0).

  – There is a lower percentage of respondents who were hospitalized in the year prior to survey than non-respondents.

4.4   Differences in Outcomes by Demographic and Health Status Characteristics

We explore differences in survey-specific outcome scores by beneficiary demographic and enrollment characteristics, health status, and medical care use rates (**Table 4-3**). For the CAHPS® M+C Enrollee Survey, we display estimates of the average rating of

Table 4-2
Distribution of Demographic and Health Status Characteristics among
CAHPS® M+C Enrollee Survey Eligibles, Respondents, and
Non-respondents, Selection Probability Weighted[1]

Characteristic	Eligibles (%)	Respondents (%)	Non-respondents (%)	
Age[2]				
Under 65	7	6	9	*
65-74	50	51	41	
75-84	34	34	34	
85 +	10	8	16	
Race				
Unknown	0.1	0	0.3	*
White	83	83	79	
Black	9	8	12	
Other	1	0.1	3	
Asian	2	2	2	
Hispanic	5	6	3	
American Indian	0.3	0.3	0.1	
Gender				
Male	43	43	44	*
Female	57	57	56	
Medicaid Status				
Not Enrolled	94	95	90	*
Enrolled	6	5	11	
Reason for Medicare Entitlement				
Aged without ESRD	93	94	91	*
Aged with ESRD	0.2	0.2	0.4	
Disabled without ESRD	7	6	9	
Disabled with ESRD	0.1	0.1	0.1	
ESRD Only	0	0	0	
Risk Score Decile				
0.36 - 0.45	10	10	8	*
0.46 - 0.53	11	11	9	
0.54 - 0.57	14	14	11	
0.58 - 0.70	5	5	4	
0.71 - 0.73	9	9	8	
0.74 - 1.87	12	12	11	
0.88 - 0.91	13	13	12	
0.92 - 1.07	8	8	9	
1.08 - 1.26	10	9	14	
1.27 - 6.91	10	9	15	
Number of Hospitalizations				
Zero	87	88	85	*
One	9	9	10	
Two	3	2	3	
Three or More	1	1	2	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Statistical significance tested using chi-square distribution differences between respondents and non-respondents. An asterisk (*) denotes significance at <0.05 level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey.

Table 4-3

Mean CAHPS® Plan Satisfaction Rating and Mean CAHPS® Composite for Getting Care When Needed for CAHPS® M+C Enrollee Survey Respondents for Levels of Demographic, Health Status, and Utilization Measures, Selection Probability Weighted[1]

	Self Reported Satisfaction with Plan				Self Reported Satisfaction with Getting Care when Needed		
Characteristic	Mean	SE	Number of Cases		Mean	SE	Number of Cases
Overall	8.5		169,171		2.8		155,958
Age[2]							
Under 65	7.7	0.024	10,816		2.7	0.005	10,446
65-74	8.4	0.007	87,519	*	2.8	0.001	79,958
75-84	8.6	0.008	57,311	*	2.8	0.002	53,022
85 +	8.6	0.017	13,525	*	2.8	0.004	12,532
Race							
White	8.5	0.005	142,365		2.8	0.001	131,557
Black	8.4	0.020	13,012		2.8	0.004	12,049
Other	8.5	0.018	13,794		2.8	0.004	12,352
Gender							
Male	8.4	0.008	72,533		2.8	0.002	66,498
Female	8.5	0.006	96,638	*	2.8	0.001	89,460
Medicaid Status							
Not Enrolled	8.5	0.005	161,372		2.8	0.001	148,661
Enrolled	8.3	0.027	7,799	*	2.7	0.006	7,297
Reasons for Medicare Entitlement							
Aged without ESRD	8.5	0.005	158,023		2.8	0.001	145,184
Aged with ESRD	8.7	0.118	321		2.8	0.026	317
Disabled without ESRD	7.7	0.024	10,692	*	2.7	0.005	10,326
Disabled with ESRD	7.7	0.254	87	*	2.8	0.046	86
ESRD Only	7.6	0.365	35		2.6	0.105	33
Risk Score Quintile							
0.36 - 0.53	8.3	0.011	36,611		2.8	0.002	33,321
0.54 - 0.70	8.4	0.011	32,711	*	2.8	0.002	29,938
0.71 - 0.87	8.5	0.010	36,431	*	2.8	0.002	33,442
0.88 - 1.07	8.6	0.011	34,273		2.8	0.002	31,702
1.08 -6.91	8.5	0.012	29,145	*	2.8	0.003	27,555
Number of Hospitalizations							
Zero	8.5	0.005	148,198		2.8	0.001	135,815
One	8.5	0.017	14,660		2.8	0.004	14,034
Two	8.5	0.032	4,238		2.8	0.007	4,101
Three or More	8.4	0.047	2,075		2.8	0.011	2,008

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of p<0.05 with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes these comparisons that exceed the significance level. The reference stratum with each set of characteristics is in bold.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey.

respondents to two measures of satisfaction—satisfaction with the health plan and satisfaction with getting care when needed—as the outcome measures. If outcome measures, such as satisfaction with care, vary by demographic characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias existing increases. Pairwise comparisons of differences in selection probability weighted response rates between the various levels of stratification and a reference group are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. Because of the large sample sizes in the CAHPS® M+C Enrollee Survey, many statistical comparisons are statistically significant. We reserve our comments to substantive differences. Our findings are summarized as follows:

- There is considerable variation in self-reported satisfaction with care received from the beneficiaries' health plans.

  – Beneficiaries under age 65, or those entitled to Medicare because of disability, reported less satisfaction with their health plan than beneficiaries age 65 and older. A similar pattern is observed when evaluating reason for Medicare entitlement.

  – Beneficiaries dually enrolled in Medicare and Medicaid were modestly less satisfied with their health plans than beneficiaries not dually enrolled.

  – Beneficiaries in the risk score quintiles indicating the best health status (lowest PIP-DCG scores) had lower rates of satisfaction than did beneficiaries in average health status.

- There is virtually no variation in self-reported satisfaction with getting care when needed across any of the demographic or health status categories.

4.5   Factors that Predict Likelihood of Response

We predict the likelihood of response as a function of sociodemographic and health status characteristics of all sampled beneficiaries using a multivariate regression model. We estimate the model unweighted and weighted by the inverse likelihood of the beneficiary being selected for survey.

- **Table 4-4** contrasts the statistically significant odds ratios resulting from a logistic regression model intended to predict response that is not weighted and one that is weighted. The odds ratios barely differ between the models. There were three fewer statistically significant categories in the model without weights.

- The direction and magnitude of the odds ratios for the included beneficiary-level variables are consistent with the descriptive comparisons between respondents and non-respondents. Following are the results from the weighted regression model:

  – Beneficiaries under the age of 65 and age 75 and older are roughly 20 percent to 45 percent less likely to respond to the CAHPS® M+C Enrollee Survey than beneficiaries age 65 to 74.

73

Table 4-4
Logistic Regression of Likelihood of Response to the CAHPS® M+C Enrollee Survey

Characteristic	Unweighted Regression Odds Ratio[1]	Selected Probability Weighted Regression Odds Ratio[2]
Beneficiary Characteristics		
Under 65	0.655	0.678
75 to 84	0.800	0.811
85 +	0.528	0.554
Black	0.658	0.666
Unknown or Other Race	0.024	0.027
Asian	0.822	1.015
Hispanic	1.796	1.812
American Indian	4.990	2.904
Male	0.934	0.960
Medicaid	0.521	0.563
ESRD	0.731	0.749
Risk Score Decile		
0.36 - 0.45	1.155	1.218
0.46 - 0.53	1.097	1.107
0.54 - 0.57	1.123	1.152
0.58 - 0.70	1.230	1.222
0.71 - 0.73	1.048	1.039
0.74 - 0.87	1.234	1.259
0.88 - 0.91	1.100	1.107
1.08 - 1.26	0.889	0.870
1.27 - 6.91	0.808	0.765
Number of Hospitalizations		
One	1.036	1.102
Two	1.054	1.131
Three or More	0.820	0.862
No. of Observations	216919	216919
Overall Chi-Sq (p-value)	10494***	336647***

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Highlighted odds ratios are statistically significant at the <0.05 level of significance. Asterisks (***) denote p<0.001 level of significance.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey.

– Beneficiaries of American Indian descent are almost three times more likely to respond than White beneficiaries. Hispanic and Asian beneficiaries are also more likely to respond than Whites. In contrast, Blacks are far less likely than White beneficiaries to respond to the survey.

– Men are less likely to respond than women.

– Beneficiaries dually enrolled in Medicare and Medicaid are almost 50 percent less likely to respond than beneficiaries not also enrolled in Medicaid.

– After controlling for health status, race, and age, beneficiaries with ESRD are significantly less likely than beneficiaries without ESRD to respond to the survey.

– Compared to beneficiaries with an average health status score, those with lower PIP-DCG scores, which equates to a higher level of health status, are generally more likely to respond to the survey. Beneficiaries with poorer health status, or higher PIP-DCG scores, are less likely to respond than those with average health status.

– Interestingly, beneficiaries who were hospitalized one or two times during the year prior to survey are more likely to respond than beneficiaries who have not had any hospitalizations. In contrast, beneficiaries who were hospitalized three or more times in the year prior to survey were about 14 percent less likely to respond.

4.6   Probable Degree of Non-response Bias

Earlier descriptive statistics showed that satisfaction with the beneficiary's health plan increased as health status, measured by the PIP-DCG score, declined. In contrast, beneficiaries who are entitled to Medicare because of the presence of ESRD or a disability expressed lower levels of satisfaction with their health plan than beneficiaries without ESRD or those entitled to Medicare because of age. Thus, there appears to be a relationship between satisfaction with a health plan and health status, although there is not necessarily a clear pattern across subpopulations.

We indirectly explore the degree of bias that may be present in estimates of satisfaction by using health status and medical care usage as proxies. We compare means of these variables for respondents to those obtained for eligible beneficiaries, including non-respondents. We also report differences in mean PIP-DCG scores between respondents and survey eligibles stratified by sociodemographic and medical care usage characteristics.

Last, we examine the differences between eligibles and respondents by plan response rate deciles to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles based on health status. Between eligibles and respondents, we compare average age; the proportion that are female, White, enrolled in Medicaid, and aged without ESRD; and average PIP-DCG risk score, number of hospitalizations, and number of inpatient days.

Pairwise comparisons of differences in selection probability weighted mean estimates between eligibles and respondents are made using a two-sided z-test for differences in means or proportions at the significance level of p<0.01 to account for multiple comparisons. We consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

- The means of selected continuous variables tell a less specific story of differences in characteristics between eligibles and respondents, despite achieving statistical significance (**Table 4-5**).

  – The mean PIP-DCG score is only 2 percent lower for respondents than for survey eligibles, implying modestly better health status.

  – Mean number of hospitalizations is modestly lower for respondents than for survey eligibles.

  – Mean number of inpatient days is also modestly lower for respondents than for survey eligibles.

- Differences in the mean health status between survey eligibles and respondents display a general trend in which health status estimates derived using the PIP-DCG risk score are often lower (better health) than those derived for survey eligibles across some of the major subpopulations of Medicare beneficiaries (**Table 4-6**). This suggests that health status estimates derived from respondents only tend to modestly overestimate the health of Medicare M+C enrollees. And, this overestimate tends to be for several of the healthier subpopulations (e.g, aged without ESRD and not dually enrolled in Medicare and Medicaid).

- A comparison of the differences between eligibles and respondents by plan response rate deciles does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles (**Table 4-7**). In fact, there are limited observed differences between eligibles and respondents for the health plans with the lowest level of response. Although we observe statistically significant differences between eligibles and respondents for some subpopulations (e.g., dually enrolled in Medicare and Medicaid), the level of difference is relatively small. The statistical difference is a function of the very large sample size for this survey.

Of more interest is the difference in the characteristics of eligible beneficiaries in the health plans with the lowest response rates. These health plans tend to have considerably larger proportions of non-White beneficiaries as well as beneficiaries dually enrolled in Medicare and Medicaid and beneficiaries with ESRD.

Table 4-5
Average Health Status and Hospital Use among CAHPS® M+C Enrollee Survey Eligibles, Respondents, and Non-respondents, Selection Probability Weighted[1]

Analytic Variable	Eligibles	Respondents	Non-respondents	Degree of Bias Difference in Means[2]	
Mean PIP-DCG Risk Score	0.88	0.86	0.98	-0.02	*
Mean Number of Hospitalizations	0.19	0.18	0.23	-0.01	*
Mean Number of Inpatient Days	7.05	6.59	8.76	-0.46	*

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey.

Table 4-6

Average PIP-DCG Score for Eligibles and Respondents by Beneficiary and Enrollment Characteristics, CAHPS® M+C Enrollee Survey, Selection Probability Weighted[1]

Characteristic	Eligibles Mean	Respondents Mean	Degree of Bias Difference in Means[2]	
Total	0.88	0.86	-0.02	*
Age				
Under 65	0.84	0.84	0.00	
65-74	0.70	0.69	-0.01	*
75-84	1.01	1.00	-0.01	*
85 +	1.35	1.33	-0.02	*
Race				
Unknown	1.17	1.01	-0.16	
White	0.88	0.86	-0.02	*
Black	0.92	0.89	-0.03	*
Other	0.87	0.80	-0.07	
Asian	0.84	0.82	-0.02	
Hispanic	0.86	0.85	-0.01	
American Indian	0.91	0.89	-0.02	
Gender				
Male	0.94	0.92	-0.02	*
Female	0.84	0.82	-0.02	*
Medicaid Status				
Not Enrolled	0.86	0.84	-0.02	*
Enrolled	1.27	1.25	-0.02	
Institutionalized Status				
Community Dwelling	NA	NA	NA	
Long-term Institutionalized	NA	NA	NA	
Nursing Home Certifiable	NA	NA	NA	
Reason for Medicare Entitlement				
Aged without ESRD	0.88	0.86	-0.02	*
Aged with ESRD	1.92	1.80	-0.12	
Disabled without ESRD	0.83	0.83	0.00	
Disabled with ESRD	1.72	1.79	0.07	
ESRD Only	0.78	0.81	0.03	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of $p<0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey.

Table 4-7
Average Demographic and Health Status Characteristics of Eligibles and Respondents by Decile of Health Plan Response Level to the CAHPS® M+C Enrollee Survey, Selection Probability Weighted[1]

	Level of Health Plan Response				
Analytic Variables	0-60%	61-70%	71-80%	81-90%	91-100%
Number of Plans (Total =409)	5	17	115	257	15
Numberof Beneficiaries (Total=216,919)	1300	8605	59617	139800	7597
Demographics					
Average Age[2]					
Eligibles	69	75	74	74*	73
Respondents	68	74	74	73	73
Percent Female					
Eligibles	56	56	58	57	56
Respondents	57	57	58	57	56
Percent White					
Eligibles	24	58*	77	86	96
Respondents	27	55	77	86	96
Percent Medicaid Enrolled					
Eligibles	36*	10*	6*	5*	3
Respondents	42	8	5	4	3
Percent Aged without ESRD					
Eligibles	80*	93	93*	94	94
Respondents	74	93	94	94	95
Health Status and Use					
Average PIP-DCG Risk Score					
Eligibles	0.95	.93*	.90*	0.87*	0.86
Respondents	0.98	0.90	0.88	0.85	0.85
Average Number of Hospitalizations					
Eligibles	0.19	0.20	0.19*	0.18	0.18
Respondents	0.20	0.19	0.18	0.18	0.17
Average Number of Inpatient Days					
Eligibles	7.5	7.5	7.5*	6.7	7.4
Respondents	6.7	7.0	6.7	6.5	6.8

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents made within the decile of response rate category using a two-side z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey.

CHAPTER 5
ANALYSIS OF NON-RESPONSE BIAS IN THE 2000 CAHPS® MEDICARE+CHOICE (M+C) DISENROLLMENT ASSESSMENT SURVEY

5.1   Description of the CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey

All Medicare plans that have contracts with physicians or physician groups are required to conduct both annual enrollment and disenrollment surveys and report the results to CMS. Legislation requires that CMS make consumer assessment information on the plans available to Medicare beneficiaries to assist them in making plan choice decisions regarding participation in the program.

The enrollment survey requirement is satisfied by the annual nationwide administration of the Medicare CAHPS® M+C Enrollee Survey, which CMS has sponsored since 1998. However, because the Enrollee Survey includes only those who have been enrolled in a plan for 6 months or more at the time of survey administration, it excludes beneficiaries who voluntarily disenrolled during the previous year. The General Accounting Office (GAO) and others have pointed out that the results of this satisfaction survey may be biased in favor of the plans, given that disenrollees, who may be among the most dissatisfied plan members, have voluntarily withdrawn from it and are excluded from the Enrollee Survey sample. Hence, there was a need to separately survey plan disenrollees and add their responses to those of enrollees.

The CAHPS® M+C Disenrollment Survey actually consists of two different but related surveys. One is intended to collect beneficiaries' assessment of their experiences while they were in the managed care plan (the Assessment Survey), while the other queries beneficiaries about their reasons for disenrolling (the Reasons Survey). Since they were in essence to be added together, the Assessment Survey component of the CAHPS® M+C Disenrollment Survey was created to be virtually identical in content to the CAHPS® M+C Enrollee Survey.

The first national implementation of the CAHPS® M+C Disenrollment Assessment Survey was conducted in 2000, and it has been repeated annually. The Assessment Survey is conducted in close coordination with the Enrollee Survey. The sample for the Assessment Survey is drawn at about the same time and in the same proportion in each health plan as that used in the Enrollee Survey to minimize design effects in the combined survey estimates. Both surveys require that beneficiaries must have had 6 months of continuous enrollment in the health plan in order to be eligible for sample selection, and they both employ the same 6-month reference period.

The 2000 CAHPS® M+C Disenrollment Assessment Survey sample consisted of 31,041 Medicare beneficiaries from a total of 281 managed care health plans. All living non-institutionalized Medicare beneficiaries who had voluntarily left their managed care plan between May and July 2000, after having been continuously enrolled in the plan for at least 6 months, were eligible to be included in the Assessment Survey sample. Persons whose plan left their area or who left the plan's area were considered non-voluntary disenrollees and not eligible for the sample. Deceased disenrollees were removed from the sampling frame before the sample was selected. Returns from the survey process also resulted in further exclusion of persons

considered ineligibles because of death or institutionalization or an administrative error related to enrollment or disenrollment.

Data collection activities for the 2000 Assessment Survey were conducted between October 6, 2000, and February 21, 2001. The multi-wave survey process involved numerous attempts to reach respondents in English and/or Spanish by regular mail, telephone, and express mail. Approximately 28.2 percent (8,769) of the initial sample of 31,041 was considered ineligible to participate in the survey; that is, the sample members had died or become institutionalized after the sample was selected, or they were considered involuntarily or mistakenly disenrolled from their health plan. About 54.8 percent (12,208) of the 22,272 eligibles completed a questionnaire. A questionnaire was considered complete if the respondent answered at least one question other than screening questions designed to identify involuntary disenrollees. Nearly 45.2 percent of eligibles (10,064) did not respond, primarily because they could not be contacted. Only 17.3 percent of eligible non-respondents actually refused, and an additional 4 percent were unable to respond due to a language barrier or disability.

As part of survey administration, non-response analysis was conducted on the 2000 CAHPS® M+C Disenrollment Assessment Survey data. For that analysis, sample members were classified as respondents or non-respondents; response propensities were then modeled using logistic regression in SUDAAN. The predicted response propensities were used to adjust the initial design-based weights (the inverse of the selection probability) upward for respondents so that they represented both respondents and non-respondents, while weights for non-respondents were set to zero. This general approach used to adjust weights for non-response has been described by Folsom (1991) and Iannacchione, Milne, and Folsom (1991).

5.2   Survey-Specific Response Rates

As with the CAHPS® M+C Enrollee Survey, we first examine differences in response by beneficiary and enrollment characteristics, health status, and health care use rates. **Table 5-1** compares the response rates for each category of enrollment, demographic, health status, and use measures presented in three ways: unweighted, as a mean of the means across plans, and weighted using the selection probability weight. Looking across the three estimates of response for each category of the characteristics shows them to be fairly similar to one another, usually within 1 or 2 percentage points. The weighting does not seem to make a great deal of difference. If anything, the mean of means approach more often results in the highest response rate calculation, while the unweighted and selection probability weighted response rates are most similar to one another. Pair-wise comparisons of differences in weighted response rates between the variables' categories and the reference level are made using a two-sided z-test for proportions at the significance level of p<0.05 with the Bonferroni multiple comparison adjustment. Findings are summarized below, focusing on the response rates in the weighted column of the table:

- With few exceptions, the weighted distribution of response rates differs significantly by category of demographic, enrollment, health status, and use measures.

Table 5-1

Survey-Specific Response Rates by Demographic and Health Status Characteristics for the 2000 CAHPS® M+C Disenrollment Assessment Survey

Characteristic	Unweighted Response Rate[1] (%)	Mean of the Means Response Rate[2] (%)	Selection Probability Weighted Response Rate[3] (%)	
Age[4]				
Under 65	48	51	48	*
65-74	59	60	58	
75-84	55	57	56	*
85 +	42	45	42	*
Race				
Unknown	29	29	35	
White	56	56	56	
Black	45	47	45	*
Other	7	7	5	*
Asian	59	64	54	
Hispanic	68	73	68	*
American Indian	84	85	87	*
Gender				
Male	56	58	56	
Female	54	55	54	*
Medicaid Status				
Not Enrolled	57	57	57	
Enrolled	42	46	41	*
Reason for Medicare Entitlement				
Aged without ESRD	56	57	56	
Aged with ESRD	38	39	40	
Disabled without ESRD	48	51	48	*
Disabled with ESRD	20	20	7	*
ESRD Only	0	0	0	
Risk Score Decile				
0.36 - 0.45	56	58	58	*
0.46 - 0.53	59	61	59	*
0.54 - 0.57	58	60	59	*
0.58 - 0.70	58	59	57	*
0.71 - 0.73	60	60	59	*
0.74 - 0.87	57	57	57	*
0.88 - 0.91	56	56	55	*
0.92 - 1.07	50	51	48	
1.08 - 1.26	49	52	49	
1.27 - 6.91	46	48	47	
Number of Hospitalizations				
Zero	56	57	55	
One	51	53	53	
Two	47	50	48	*
Three or More	44	47	44	*

[1] An equal weight whereby all *sampled beneficiaries* are given a weight of 1.

[2] An equal weight whereby all *health plans or states* are given a weight of 1.

[3] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[4] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey.

- The response rates of beneficiaries under age 65 and above 74 years are significantly lower than those for beneficiaries 65 to 74 years of age. This finding is particularly true for the youngest and oldest age groups.

- The response rates for all of the racial/ethnic groups except Asians are significantly lower than for Whites.

- The response rate for males is significantly higher than for females, although the difference is small.

- Beneficiaries not also enrolled in Medicaid have a significantly higher response rate than beneficiaries dually enrolled in Medicare and Medicaid.

- Beneficiaries entitled to Medicare because they are disabled (with or without ESRD) or aged with ESRD responded at a significantly lower rate than aged beneficiaries without ESRD.

- Beneficiaries with a PIP-DCG score lower than the category containing 1.00 (in better health because they are below the mean) have significantly higher response rates than those in the category containing 1.00.

- Beneficiaries with two or more hospital discharges have significantly lower response rates than those who had not been hospitalized at all in the prior year.

5.3   Differences in Characteristics of Respondents and Non-respondents

Next we examine differences in the weighted distribution of beneficiary characteristics between respondents and non-respondents. This review provides an overall sense of how different respondents are from non-respondents in terms of their demographic, enrollment, health status, and utilization characteristics and is a critical factor in the determination of potential non-response bias. The distributions are weighted using selection probability weights. Statistical significance testing is performed using the chi-square test and $p<0.05$ level of significance. We summarize our findings from **Table 5-2** below:

- The distribution of respondents is significantly different from non-respondents. There are more young elderly in the respondent group, and more old elderly and young disabled persons among the non-respondents.

- Respondents are significantly more likely to be White and Hispanic, and non-respondents are more likely to be Black.

- Respondents are significantly more likely to be male and non-respondents female.

- Respondents are significantly more likely to not be enrolled in Medicaid along with Medicare, while non-respondents are more likely to be dually enrolled.

- The distribution of reasons for Medicare entitlement is statistically significantly different between respondents and non-respondents. Generally speaking, those

Table 5-2

**Distribution of Demographic and Health Status Characteristics among 2000 CAHPS®
M+C Disenrollment Assessment Survey Eligibles, Respondents, and Non-respondents,
Selection Probability Weighted[1]**

Characteristic	Eligibles (%)	Respondents (%)	Non-respondents (%)	
Age[2]				
Under 65	12	11	14	*
65-74	49	52	45	
75-84	30	30	30	
85 +	9	7	11	
Race				
Unknown	0.1	0.1	0.2	*
White	75	77	73	
Black	17	14	20	
Other	1	0.1	2	
Asian	1	1	1	
Hispanic	6	7	4	
American Indian	0.3	0.4	0.1	
Gender				
Male	43	44	41	*
Female	57	56	59	
Medicaid Status				
Not Enrolled	87	90	83	*
Enrolled	13	10	17	
Reason for Medicare				
Aged without ESRD	87	89	85	*
Aged with ESRD	0.2	0.1	0.3	
Disabled without ESRD	12	11	14	
Disabled with ESRD	0.1	0	0.1	
ESRD Only	0.04	0	0.04	
Risk Score Decile				
0.36 - 0.45	10	10	10	*
0.46 - 0.53	13	14	11	
0.54 - 0.57	3	4	3	
0.58 - 0.70	14	14	13	
0.71 - 0.73	8	8	7	
0.74 - 0.87	13	13	12	
0.88 - 0.91	11	11	10	
0.92 - 1.07	10	9	11	
1.08 - 1.26	10	9	11	
1.27 - 6.91	10	8	12	
Number of Hospitalizations				
Zero	86	87	84	*
One	9	9	10	
Two	3	3	4	
Three or More	2	1	2	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Statistical significance tested using chi-square distribution differences between respondents and non-respondents. An asterisk (*) denotes significance at <0.05 level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey.

whose Medicare entitlement results from reaching age 65 are more likely to be respondents, while those who are entitled to Medicare as a result of disability are more likely to be non-respondents.

- The distribution of respondents and non-respondents differs significantly according to their PIP-DCG score deciles. Slightly larger proportions of respondents are in the deciles with lower PIP-DCG risk scores (0.91 or less), representing better health, and larger proportions of non-respondents have scores in the higher deciles (scores above 0.92), representing less healthy states.

- Respondents also differ significantly from non-respondents in the number of hospitalizations they had in the prior year. Respondents are slightly more likely to have had no hospital stays, while non-respondents are more likely to have had one or more hospitalizations.

5.4 Differences in Outcomes by Demographic and Health Status Characteristics

Next, we examine differences in survey-specific outcome measures by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. For the CAHPS® M+C Disenrollment Assessment Survey, we display estimates of the average ratings of respondents to two measures of satisfaction—satisfaction with the health plan and satisfaction with getting care when needed—as the outcome measure. If outcome measures, such as satisfaction with care, vary by demographic characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias existing increases. Pair-wise comparisons of differences in weighted response rates between the various categories of the characteristics and a reference level is made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. Because of the large sample sizes in the CAHPS® M+C Disenrollment Assessment Survey, many statistical comparisons are statistically significant. We reserve our comments to the largest differences. Our findings from **Table 5-3** are summarized below:

- There is considerable variation in self-reported satisfaction with care received from the beneficiaries' health plans.

    – Beneficiaries under age 65 reported less satisfaction with their health plan than beneficiaries age 65 to 74, while beneficiaries age 75 to 84 were more satisfied.

    – Female beneficiaries are more satisfied with their plan than males.

    – Beneficiaries dually enrolled in Medicare and Medicaid were modestly less satisfied with their health plan than beneficiaries not dually enrolled.

    – Disabled beneficiaries (without ESRD) were less satisfied with their plan than the aged (without ESRD).

86

Table 5-3
Mean CAHPS® Plan Satisfaction Rating and Mean CAHPS® Composite for Getting Care When Needed for 2000 CAHPS® M+C Disenrollment Assessment Survey for Levels of Demographic, Health Status, and Utilization Measures, Selection Probability Weighted[1]

Characteristic	Self Reported Satisfaction with Plan				Self Reported Satisfaction with Getting Care When Needed			
	Mean	SE	Number of Cases		Mean	SE	Number of Cases	
Age[2]								
Under 65	5.5	0.093	1,219	*	2.3	0.022	1,161	*
65-74	6.5	0.040	5,719		2.5	0.009	5,218	
75-84	6.8	0.053	3,247	*	2.6	0.012	2,990	
85 +	6.5	0.118	712		2.5	0.027	645	
Race								
White	6.5	0.033	8,518		2.5	0.008	7,867	
Black	6.5	0.089	1,420		2.5	0.019	1,238	
Other	6.7	0.104	959	*	2.5	0.024	909	
Gender								
Male	6.3	0.045	4,804		2.5	0.010	4,398	
Female	6.6	0.040	6,093	*	2.5	0.009	5,616	
Medicaid Status								
Not Enrolled	6.5	0.031	9,820		2.5	0.007	9,012	
Enrolled	6.2	0.105	1,077	*	2.4	0.023	1,002	*
Reasons for Medicare Entitlement								
Aged without ESRD	6.6	0.031	9,662		2.5	0.007	8,837	
Aged with ESRD	7.3	0.682	15		2.8	0.117	15	
Disabled without ESRD	5.5	0.093	1,217	*	2.3	0.022	1,159	*
Disabled with ESRD	8.2	1.500	2		1.8	0.125	2	*
ESRD Only								
Risk Score Quintile								
0.36 - 0.53	6.4	0.059	2,642	*	2.5	0.014	2,419	
0.54 - 0.70	6.4	0.067	2,001	*	2.5	0.015	1,822	
0.71 - 0.87	6.6	0.064	2,321		2.5	0.014	2,118	
0.88 - 1.07	6.7	0.068	2,090		2.5	0.016	1,930	
1.08 -6.91	6.4	0.075	1,843	*	2.5	0.017	1,725	
Number of Hospitalizations								
Zero	6.5	0.032	9,475		2.5	0.007	8,663	
One	6.4	0.103	981		2.5	0.023	919	
Two	6.2	0.187	291		2.5	0.043	282	
Three or More	5.6	0.256	150	*	2.5	0.063	150	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of p<0.05 with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey.

- Beneficiaries in the risk score quintiles indicating the best health status (PIP-DCG scores below 0.70) and the worst (1.08 and higher) had lower rates of satisfaction than did beneficiaries in average health status (category containing 1.00).

- Beneficiaries with three or more hospital stays were less satisfied with their plan than persons who had none.

- There is less variation in self-reported satisfaction with getting care when needed across most of the demographic, enrollment, health status, and utilization categories.

  - Beneficiaries under 65 years of age were significantly less satisfied with getting needed care in their plans than were persons 65 to 74 years of age.

  - Persons who were dually enrolled in Medicare and Medicaid were less satisfied with getting needed care from their plan than those who were enrolled in Medicare alone.

  - Disabled beneficiaries were less satisfied with being able to get needed care from their plan than persons who received their Medicare because they were age 65 or older.

5.5   Factors that Predict Probability of Response

In **Table 5-4** we present the results of predicting the probability of response as a function of demographic and health status characteristics of all sampled beneficiaries using a multivariate logistic regression model. We estimate the model unweighted and weighted.

- There are eight more significant predictors of response in the weighted model than the unweighted.

- The direction and magnitude of the odds ratios for the included beneficiary-level variables are consistent with the descriptive comparisons between respondents and non-respondents. The results from the weighted regression model are as follows:

  - Beneficiaries under the age of 65 and age 85 and older are roughly 20 percent to 40 percent less likely to respond than beneficiaries age 65 to 74.

  - The odds of beneficiaries of American Indian descent responding are five times higher than White beneficiaries. Hispanic beneficiaries have about 80 percent higher odds of responding than Whites. In contrast, Blacks and Asians have about 30 percent and 10 percent lower odds of responding, respectively, than White beneficiaries.

  - Beneficiaries dually enrolled in Medicare and Medicaid have about 40 percent lower odds of responding than beneficiaries not also enrolled in Medicaid.

  - Beneficiaries with ESRD have 60 percent lower odds of responding than beneficiaries without ESRD.

Table 5-4
Logistic Regression of Probability of Response to the 2000 CAHPS® M+C Disenrollment Assessment Survey

Characteristic	Unweighted Regression Odds Ratio[1]	Selection Probability Weighted Regression Odds Ratio[2]
Beneficiary Characteristics		
Under 65	0.749	0.768
75 to 84	0.901	0.982
85 +	0.576	0.623
Black	0.688	0.691
Unknown or Other Race	0.079	0.059
Asian	1.105	0.928
Hispanic	1.772	1.833
American Indian	4.251	5.311
Male	1.063	1.021
Medicaid	0.617	0.588
ESRD	0.495	0.401
Risk Score Decile		
0.36 - 0.45	1.046	1.223
0.46 - 0.53	1.139	1.250
0.54 - 0.57	1.031	1.156
0.58 - 0.70	1.092	1.147
0.71 - 0.73	1.108	1.200
0.74 - 0.87	1.162	1.188
0.88 - 0.91	1.032	0.991
1.08 - 1.26	1.045	1.088
1.27 - 6.91	0.943	1.010
Number of Hospitalizations		
One	0.950	1.025
Two	0.868	0.909
Three or More	0.802	0.852
No. of Observations	22272	22272
Overall Chi-Sq (p-value)	857***	11013***

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Highlighted odds ratios are statistically significant at the <0.05 level of significance. Asterisks (***) denote p<0.001 level of significance.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey.

- Compared to beneficiaries with an average health status score, those with lower PIP-DCG scores, which equates to a higher level of health status, have odds of responding that are 15 percent to 25 percent higher than beneficiaries in the category that contains 1.00.

- Beneficiaries who were hospitalized two or three times during the year prior to the survey have slightly lower odds of responding (10 percent to 15 percent) than persons with none.

5.6 Probable Degree of Non-response Bias

We indirectly explore the degree of bias that may be present in estimates of satisfaction by using health status (PIP-DCG) and medical care usage (hospital use) as proxies. In **Table 5-5,** we compare the means of these variables for respondents to those obtained for eligible beneficiaries, including non-respondents.

- The means of selected selection probability weighted continuous variables tell a less specific story of differences in characteristics between eligibles and respondents despite achieving statistical significance (**Table 5-5**).

  - The mean PIP-DCG score is 4 percent lower for respondents than for survey eligibles, implying modestly better health status.

  - Mean number of hospitalizations is also modestly lower for respondents than for survey eligibles.

  - Mean number of inpatient days is slightly higher for respondents than for survey eligibles, but it is not statistically significant.

In **Table 5-6** we report differences in mean PIP-DCG scores between respondents and survey eligibles according to the categories of demographic and enrollment characteristics. Differences in the mean health status score (from PIP-DCG as a proxy) between survey eligibles and respondents display a general trend in which health status estimates of respondents are often moderately better (lower mean PIP-DCG scores) than those for survey eligibles across many subpopulations of Medicare beneficiaries.

- Respondent beneficiaries between the ages of 65 and 84 have significantly better health status (lower PIP-DCG scores) than the sample of eligibles from which they came.

- White, Black, and Hispanic respondents have significantly lower PIP-DCG scores (better health status) than the pool of eligibles forming the sample.

- Both men and women respondents have significantly better health status than the eligible sample from which they came.

- Respondents were in a significantly better health state than the sample eligibles, regardless of whether they are also enrolled in Medicaid.

90

Table 5-5
Average Health Status and Hospital Use among 2000 CAHPS® M+C
Disenrollment Assessment Survey Eligibles, Respondents, and Non-respondents,
Selection Probability Weighted[1]

Analytic Variable	Eligibles	Respondents	Non-respondents	Degree of Bias Difference in Means[2]	
Mean PIP-DCG Risk Score	0.91	0.87	0.95	-0.04	*
Mean Number of Hospitalizations	0.22	0.19	0.25	-0.03	*
Mean Number of Inpatient Days	8.92	9.68	8.14	0.76	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey.

Table 5-6
Average PIP-DCG Score for Eligibles and Respondents by Beneficiary and Enrollment Characteristics, 2000 CAHPS® M+C Disenrollment Assessment Survey, Selection Probability Weighted[1]

Characteristic	Eligibles Mean	Respondents Mean	Degree of Bias Difference in Means[2]	
Total	0.91	0.87	-0.04	*
Age				
Under 65	0.80	0.82	0.02	
65-74	0.75	0.73	-0.02	*
75-84	1.07	1.03	-0.04	*
85 +	1.36	1.33	-0.03	
Race				
Unknown	1.34	0.77	-0.57	
White	0.89	0.87	-0.02	*
Black	0.99	0.92	-0.07	*
Other	0.87	0.66	-0.21	
Asian	0.80	0.75	-0.05	
Hispanic	0.91	0.87	-0.04	*
American Indian	0.95	0.92	-0.03	
Gender				
Male	0.94	0.92	-0.02	*
Female	0.88	0.84	-0.04	*
Medicaid Status				
Not Enrolled	0.86	0.83	-0.03	*
Enrolled	1.25	1.22	-0.03	
Reason for Medicare Entitlement				
Aged without ESRD	0.92	0.88	-0.04	*
Aged with ESRD	2.04	1.06	-0.98	*
Disabled without ESRD	0.80	0.82	0.02	
Disabled with ESRD	0.83	1.96	1.13	
ESRD Only	0.44	NA	NA	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of $p<0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey.

- Respondents who obtain their Medicare because they have reached 65 years of age, with or without ESRD, have significantly better health status (lower PIP-DCG scores) than the pool of eligibles from which they originated.

And, last, in **Table 5-7**, we examine the differences between eligibles and respondents by plan response rate deciles to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles. We compare eligibles and respondents with respect to their average age; the proportion that are female, White, enrolled in Medicaid, and aged without ESRD; and average PIP-DCG risk score, number of hospitalizations, and number of inpatient days. Pair-wise comparisons of differences in weighted mean estimates between eligibles and respondents are made using a two-sided z-test for differences in means or proportions at the significance level of $p<0.01$ to account for multiple comparisons. We consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

- There are no significant differences in mean age between the pool of eligibles and respondents, although there appears to be a slight trend for the means to increase as the plan response rate increases.

- Ignoring the two sets of deciles with very few plans and beneficiaries (0 to 10 percent and 91 percent to 100 percent), the same pattern exists for the percent female, the percent White, and the percent who are enrolled in Medicare because they are elderly without ESRD.

- A tendency in the opposite direction is suggested by the percent of beneficiaries who are dually enrolled in Medicare and Medicaid; dual enrollment seems to decrease as the plan response rate increases. Plans with response rates between 31 percent and 70 percent have a significantly higher proportion of dually enrolled in Medicare and Medicaid than they have respondents. This is consistent with the logistic regression analysis, indicating that dual enrollees are less likely to respond than beneficiaries not also enrolled in Medicaid.

Table 5-7

Average Demographic and Health Status Characteristics of Eligibles and Respondents by Decile of Health Plan Response Level to the 2000 CAHPS® M+C Disenrollment Assessment Survey, Selection Probability Weighted[1]

	Level of Health Plan Response							
Analytic Variable	0-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	100%
Number of Plans (Total =393)	20	35	77	114	93	40	8	6
Number of Beneficiaries (Total=22,272)	204	2,799	4,245	8,834	4,722	1,306	124	38
Demographics								
Average Age								
Eligibles	72	70	72	72	72	71	73	73
Respondents	73	70	72	72	72	71	74	73
Percent Female								
Eligibles	53	54	57	59	58	55	56	52
Respondents	51	55	55	57	58	55	57	53
Percent White								
Eligibles	61	60	69	78	81	84	89	71
Respondents	62	58	69	79	82	84	90	71
Percent Medicaid Enrolled								
Eligibles	16	24*	15	12*	11*	9	9	8
Respondents	16	20	13	9	8	7	6	10
Percent Aged w/o ESRD								
Eligibles	87	81	87	89*	89	88	91	93
Respondents	92	84	88	90	89	89	97	93
Health Status and Use								
Average PIP-DCG Risk Score								
Eligibles	0.89	0.98*	0.91*	0.92	0.89*	0.81	0.90	0.75
Respondents	0.93	0.92	0.85	0.90	0.85	0.80	0.88	0.71
Average Number of Hospitalizations								
Eligibles	0.14	0.26	0.22	0.23	0.20	0.17	0.24	0.02
Respondents	0.30	0.22	0.18	0.21	0.17	0.17	0.14	0.00
Average Number of Inpatient Days								
Eligibles	5.3	8.3	7.6	7.7	14.5	6.0	13.0	3.0
Respondents	5.6	7.9	7.5	7.4	18.0	6.0	4.8	0

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents made within the decile of response rate category using a two-sided z-test of differences at the significance level of $p < 0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Assessment Survey.

CHAPTER 6
ANALYSIS OF NON-RESPONSE BIAS IN THE 2000 CAHPS® MEDICARE+CHOICE (M+C) DISENROLLMENT REASONS SURVEY

6.1 Description of the CAHPS® Medicare+Choice (M+C) Disenrollment Reasons Survey

The 2000 Medicare CAHPS® M+C Disenrollment Reasons Survey is the second of the *two* surveys that form the Medicare CAHPS® M+C Disenrollment Survey—the other being the Medicare CAHPS® M+C Disenrollment Assessment Survey. The Reasons Survey was conducted for the first time in the summer of 2000 with a sample of Medicare beneficiaries who voluntarily left their managed care plan or continuing cost contracts during 2000. The Reasons Survey is conducted quarterly, as opposed to once a year like the other Medicare CAHPS® Surveys, to have it occur as close to the disenrollment date as possible to minimize the number of sample members who cannot be located, have recall problems, or are confused by having disenrolled from more than one plan. Although data collection and processing are implemented on a quarterly basis, the survey results are reported annually.

The survey data are collected via a mail survey with telephone follow-up of mail survey non-respondents. The Reasons Survey questionnaire was especially designed to collect information about the reasons that sample members left their former Medicare managed health care plan. To assist with respondent recall, virtually every item in the questionnaire is customized with the name of the plan from which the sample member disenrolled. The questionnaire contained 78 questions: 7 screening questions to verify that respondents were truly voluntary disenrollees and 35 questions about reasons for leaving the health plan. The questions that asked about reasons for leaving were grouped into seven domains: (1) plan availability; (2) doctors and other health providers; (3) access to care; (4) information about the plan; (5) pharmacy benefits; (6) costs and benefits; and (7) access to hospitals, medical equipment, and home health care. In addition there were 23 questions about health status and demographic characteristics and 5 questions asking the respondent to rate the sample health plan and the care received from that plan and questions about the experience with the plan, plus 8 questions about the appeals and grievance process. The questions were translated into Spanish and available by request as either a questionnaire or as a Spanish-language telephone interview.

Our non-response bias analysis is restricted to the 3rd quarter only eligible beneficiaries. The number determined to be eligible for the July to September 2000 CAHPS® M+C Disenrollment Survey was 12,659. The eligible beneficiaries represent approximately 55 percent of the initial sample of 23,219; eligible beneficiaries are those who were able to be contacted and determined to have voluntarily disenrolled from the assigned health plan. Beneficiaries unable to be contacted are deemed as ineligible. The response rate is calculated as a function of the number of respondents with any information divided by sample members for whom eligibility has been determined. In the 3rd quarter 2000 CAHPS® M+C Disenrollment Survey, 7,395 beneficiaries were deemed respondents, for a response rate of 58 percent.

As part of survey administration, non-response analysis was conducted on the 2000 Reasons Survey. For that analysis, sample members were classified as respondents or non-respondents and response propensities were modeled using logistic regression in SUDAAN

CAHPS®. The predicted response propensities were used to adjust the initial design-based weights upward for respondents so that they represented both respondents and non-respondents; weights for non-respondents were set to zero. The general approach used to adjust weights for non-response is described by Folsom (1991) or Iannacchione, Milne, and Folsom (1991).

6.2 Survey-Specific Response Rates

As with the CAHPS® M+C Disenrollment Assessment Survey, we first examine differences in response by beneficiary and enrollment characteristics, health status, and health care use rates. **Table 6-1** compares the response rates for each category of enrollment, demographic, health status, and use measures presented in three ways: unweighted, as a mean of the means across plans, and selection probability weighted. Looking across the three estimates of response for each category of the characteristics shows them to be fairly similar to one another, usually within 1 or 2 percentage points. The weighting does not seem to make a great deal of difference. Pair-wise comparisons of differences in enrollment weighted response rates between the variables' categories and the reference level are made using a two-sided z-test for proportions at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. Findings are summarized below, focusing on the response rates in the column.

- Only a very few of the weighted distributions of response rates differ significantly by category of demographic, enrollment, health status, and use measures.

    - The response rates of beneficiaries above 74 years of age are significantly lower than those for beneficiaries age 65 to 74.

    - The response rates for Blacks and Hispanics are significantly lower than for Whites.

    - The response rate for females is significantly lower than for males, although the difference is small.

    - Beneficiaries dually enrolled in Medicare and Medicaid have significantly lower response rates than Medicare beneficiaries not dually enrolled in Medicaid.

    - Beneficiaries with a PIP-DCG score in the lowest three categories (best health status) have significantly higher response rates than those in the category containing 1.00, reflecting average health status.

    - Beneficiaries who had two hospital discharges have a significantly lower response rate than those who had not been hospitalized at all in the prior year.

6.3 Differences in Characteristics of Respondents and Non-respondents

Next we examine differences in the weighted distribution of beneficiary characteristics between respondents and non-respondents. This review provides an overall sense of how

Table 6-1

Survey-Specific Response Rates by Demographic and Health Status Characteristics for the 2000 CAHPS® M+C Disenrollment Reasons Survey

Characteristic	Unweighted Response Rate[1] (%)	Mean of the Means Response Rate[2] (%)	Selection Probability Weighted Response Rate[3] (%)	
Age[4]				
Under 65	57	56	59	
65-74	63	62	61	
75-84	57	57	54	*
85 +	43	44	43	*
Race				
Unknown	33	33	44	
White	60	59	59	
Black	56	58	52	*
Other	48	54	49	*
Asian	53	54	47	
Hispanic	40	40	40	*
American Indian	55	55	59	
Gender				
Male	60	60	58	
Female	57	57	56	*
Medicaid Status				
Not Enrolled	59	60	58	
Enrolled	49	48	45	*
Reason for Medicare Entitlement				
Aged without ESRD	59	58	57	
Aged with ESRD	42	42	48	
Disabled without ESRD	57	56	58	
Disabled with ESRD	50	50	50	
ESRD Only	50	50	75	
Risk Score Decile				
0.36 - 0.45	62	61	62	*
0.46 - 0.53	64	64	62	*
0.54 - 0.57	63	62	62	*
0.58 - 0.70	61	63	58	
0.71 - 0.73	60	61	56	
0.74 - 0.87	58	58	52	
0.88 - 0.91	58	56	57	
0.92 - 1.07	55	57	55	
1.08 - 1.26	51	51	52	
1.27 - 6.91	49	49	47	
Number of Hospitalizations				
Zero	59	59	57	
One	56	55	56	
Two	52	51	48	*
Three or More	48	49	59	

[1] An equal weight whereby all *sampled beneficiaries* are given a weight of 1.

[2] An equal weight whereby all *health plans or states* are given a weight of 1.

[3] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[4] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of p<0.05 with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 CAHPS Medicare+Choice (M+C) Disenrollment Reasons Survey (3rd quarter only).

different respondents are from non-respondents in terms of their demographic, enrollment, health status, and utilization characteristics and is a critical factor in the determination of potential non-response bias. The distributions are weighted. Statistical significance testing is performed using the chi-square test and p<0.05 level of significance. We summarize our findings from **Table 6-2** below:

- Respondents differ by age from non-respondents, with disproportionately more beneficiaries 65 to 74 years old responding, and fewer responding in the disabled under 65 group and in the 85 and older age group.

- Respondents also differ by race/ethnicity from non-respondents, with disproportionately more White beneficiaries responding and slightly fewer Black and Hispanic beneficiaries responding.

- With respect to gender, male beneficiaries are slightly more likely to respond than females.

- When classified by their Medicaid status, beneficiaries dually enrolled in Medicare and Medicaid are less likely to respond than Medicare beneficiaries not dually enrolled.

- When rates of response are arrayed according to health status as measured by PIP-DCG score deciles, beneficiaries who are healthier (the categories including PIP-DCG scores below 0.74) respond more often, while categories including scores of 0.74 and higher (less healthy) have lower response rates.

- Beneficiaries who were not hospitalized in the prior year are slightly more likely to respond than beneficiaries who had been hospitalized.

6.4   Factors that Predict Probability of Response

In **Table 6-3**, we predict the probability of response as a function of demographic, enrollment, health status, and utilization characteristics of all sampled beneficiaries using a multivariate regression model. We estimate the model unweighted and weighted by the inverse of the probability of the beneficiary being selected for survey.

- There are 12 fewer statistically significant variables in the model without weights.

- The direction and magnitude of the odds ratios for the included beneficiary-level variables are consistent with the descriptive comparisons between respondents and non-respondents. The results from the weighted regression model are as follows:

  – Beneficiaries age 75 and older have odds of responding that are roughly 20 percent to 50 percent less than beneficiaries age 65 to 74.

  – Beneficiaries who are Black, Hispanic, and Asian have odds of responding that are from 20 percent to 50 percent less than White beneficiaries.

98

Table 6-2
Distribution of Demographic and Health Status Characteristics among 2000 CAHPS®
M+C Disenrollment Reasons Survey Eligibles, Respondents, and Non-respondents,
Selection Probability Weighted[1]

Characteristic	Eligibles (%)	Respondents (%)	Non-respondents (%)	
Age[2]				
Under 65	11	11	12	*
65-74	49	52	44	
75-84	31	30	32	
85 +	9	7	13	
Race				
Unknown	0.2	0.1	0.3	*
White	80	82	77	
Black	13	12	13	
Other	3	3	4	
Asian	1	1	1	
Hispanic	3	2	4	
American Indian	0.1	0.1	0.1	
Gender				
Male	41	42	39	*
Female	59	58	61	
Medicaid Status				
Not Enrolled	87	90	84	*
Enrolled	13	11	16	
Reason for Medicare Entitlement				
Aged without ESRD	89	89	88	
Aged with ESRD	0.1	0.1	0.1	
Disabled without ESRD	11	11	12	
Disabled with ESRD	0	0	0	
ESRD Only	0	0	0	
Risk Score Decile				
0.36 - 0.45	10	11	9	*
0.46 - 0.53	12	13	10	
0.54 - 0.57	12	13	10	
0.58 - 0.70	5	5	5	
0.71 - 0.73	15	15	14	
0.74 - 0.87	6	6	6	
0.88 - 0.91	11	11	12	
0.92 - 1.07	10	9	11	
1.08 - 1.26	10	8	11	
1.27 - 6.91	10	8	12	
Number of Hospitalizations				
Zero	86	87	85	*
One	9	9	10	
Two	3	3	4	
Three or More	2	1	2	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Statistical significance tested using chi-square distribution differences between respondents and non-respondents. An asterisk (*) denotes significance at <0.05 level.

Source: RTI analysis of the 2000 CAHPS Medicare+Choice (M+C) Disenrollment Reasons Survey (3rd quarter only).

Table 6-3
Logistic Regression of Probability of Response to the 2000 CAHPS® M+C Disenrollment Reasons Survey

Characteristic	Unweighted Regression Odds Ratio[1]	Selection Probability Weighted Regression Odds Ratio[2]
Beneficiary Characteristics		
Under 65	0.905	1.029
75 to 84	0.897	0.790
85 +	0.532	0.507
Black	0.872	0.807
Unknown or Other Race	0.585	0.652
Asian	0.801	0.687
Hispanic	0.460	0.502
American Indian	0.802	0.985
Male	1.073	1.067
Medicaid	0.734	0.646
ESRD	0.700	0.927
Risk Score Decile		
0.36 - 0.45	0.972	0.913
0.46 - 0.53	1.017	0.858
0.54 - 0.57	1.017	0.942
0.58 - 0.70	0.971	0.802
0.71 - 0.73	0.925	0.805
0.74 - 0.87	0.938	0.741
0.88 - 0.91	0.886	0.916
1.08 - 1.26	0.859	0.926
1.27 - 6.91	0.781	0.694
Number of Hospitalizations		
One	1.036	1.097
Two	0.903	0.867
Three or More	0.816	1.263
No. of Observations	12,658	12,658
Overall Chi-Sq (p-value)	315***	2691***

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Highlighted odds ratios are statistically significant at the <0.05 level of significance. Asterisks (***) denote p<0.001 level of significance.

Source: RTI analysis of the 2000 CAHPS Medicare+Choice (M+C) Disenrollment Reasons Survey (3rd quarter only).

–   The odds of men responding are about 7 percent higher than women.

–   Beneficiaries dually enrolled in Medicare and Medicaid have almost 35 percent lower odds of responding than beneficiaries not enrolled in Medicaid.

–   Compared to beneficiaries with an average health status score (PIP-DCG = 1.00), those in both better and poorer health are less likely to respond.

–   Interestingly, beneficiaries who were hospitalized one or three or more times during the year prior to survey have higher odds by 10 percent to 25 percent of responding than beneficiaries who did not have had any hospitalizations. In contrast, beneficiaries who were hospitalized two times in the year prior to survey are about 13 percent less likely to respond.

6.5   Probable Degree of Non-response Bias

We indirectly explore the degree of bias that may be present in estimates of reasons for disenrolling by using health status and medical care usage as proxies. In **Table 6-4**, we compare means of these variables for respondents to those obtained for eligible beneficiaries.

- The means of weighted continuous variables tell a less specific story of differences in characteristics between eligibles and respondents.

    –   The mean PIP-DCG score is only 3 percent lower for respondents than for survey eligibles, implying modestly better health status among respondents. It is statistically significant, however.

We also report differences in mean PIP-DCG scores between respondents and survey eligibles stratified by demographic and enrollment characteristics in **Table 6-5**. Pairwise comparisons of differences in weighted mean estimates between the pool of sample eligibles and respondents are made using a two-sided z-test for differences in means or proportions at the significance level of p<0.01 to account for multiple comparisons. We consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

- Differences in the mean health status between survey eligibles and respondents display a general trend in which health status estimates derived using the PIP-DCG risk score are often lower (better health) than those derived for survey eligibles across some of the major subpopulations of Medicare beneficiaries. This suggests that health status estimates derived from respondents alone tend to modestly overestimate the health of Medicare M+C disenrollees.

    –   While respondent beneficiaries across all age groups have slightly lower mean PIP-DCG score and are therefore in better health, the only age group for which the difference is statistically significant is that of beneficiaries 85 years of age and older.

101

Table 6-4
Average Health Status and Hospital Use among 2000 CAHPS® M+C
Disenrollment Reasons Survey Eligibles, Respondents, and Non-respondents,
Selection Probability Weighted[1]

Analytic Variable	Eligibles	Respondents	Non-respondents	Degree of Bias Differences in Means[2]	
Mean PIP-DCG Risk Score	0.91	0.88	0.96	-0.03	*
Mean Number of Hospitalizations	0.21	0.19	0.22	-0.02	
Mean Number of Inpatient Days	8.80	7.76	10.05	-1.14	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Reasons Survey (3rd quarter only).

Table 6-5
Average PIP-DCG Score for Eligibles and Respondents by Beneficiary and
Enrollment Characteristics, 2000 CAHPS® M+C Disenrollment Reasons Survey,
Selection Probability Weighted[1]

Characteristic	Eligibles Mean	Respondents Mean	Degree of Bias Difference in Means[2]	
Total	0.92	0.88	-0.04	*
Age				
Under 65	0.84	0.82	-0.02	
65-74	0.75	0.73	-0.02	
75-84	1.04	1.04	0.00	
85 +	1.38	1.28	-0.10	*
Race				
Unknown	1.32	0.87	-0.45	
White	0.90	0.87	-0.03	*
Black	0.99	0.93	-0.06	
Other	0.82	0.85	0.03	
Asian	1.03	0.89	-0.14	
Hispanic	0.95	0.93	-0.02	
American Indian	1.14	1.14	0.00	
Gender				
Male	0.96	0.93	-0.03	
Female	0.89	0.84	-0.05	*
Medicaid Status				
Not Enrolled	0.87	0.84	-0.03	*
Enrolled	1.23	1.19	-0.04	
Reason for Medicare Entitlement				
Aged without ESRD	0.92	0.88	-0.04	*
Aged with ESRD	2.06	2.19	0.13	
Disabled without ESRD	0.83	0.81	-0.02	
Disabled with ESRD	1.52	0.49	-1.03	
ESRD Only	0.83	0.88	0.05	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS® Medicare+Choice (M+C) Disenrollment Reasons Survey (3rd quarter only).

- Again, the direction and magnitude of difference between respondents and eligibles by race/ethnicity suggests that respondents are consistently healthier than eligibles, but the only statistically significant difference is for Whites.

- Male and female respondents both have lower PIP-DCG scores and are in better health than the pool of sample eligibles, but only the difference for females is statistically significant.

- The respondents not dually enrolled in Medicare and Medicaid have a significantly lower mean PIP-DCG score (better health) than the eligibles.

- Of all the categories of reasons that persons have for receiving Medicare, only respondents who are aged without ESRD are in significantly better health status as measured by the PIP-DCG score.

Last, in **Table 6-6**, we examine the differences between eligibles and respondents by plan response rate deciles. Our intent with this table is to investigate whether there is a response rate below which respondents are an unrepresentative sample of survey eligibles. Between eligibles and respondents, we compare average age; the proportion that are female, White, enrolled in Medicaid, and aged without ESRD; and average PIP-DCG risk score, number of hospitalizations, and number of inpatient days.

- A comparison of the differences between eligibles and respondents by plan response rate deciles does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles. In fact, there are very few observable and fewer statistically significant differences between eligibles and respondents in health plans at any level of plan response.

- Of more interest is the difference in the characteristics of eligible beneficiaries in the health plans with the lowest response rates. These health plans tend to have considerably larger proportions of non-White beneficiaries, as well as beneficiaries dually enrolled in Medicare and Medicaid, and more and longer hospital stays.

Table 6-6

Average Demographic and Health Status Characteristics of Eligibles and Respondents by Decile of Health Plan Response Level to the 2000 CAHPS® M+C Disenrollment Reasons Survey, Selection Probability Weighted[1]

Analytic Variables	Level of Health Plan Response							
	0-30%	31-40%	41-50%	51-60%	61-70%	71-80%	81-90%	91-100%
Number of Plans (Total = 271)	9	15	61	74	67	32	9	4
Number of Beneficiaries (Total = 12,658)	132	687	2,851	3,780	3,354	1,441	398	15
Demographics								
Average Age[2]								
Eligibles	73	73	74	73	73	72	71	75
Respondents	70	73	74	72	72	72	72	75
Percent Female								
Eligibles	62	58	59*	61	58	60	66	60
Respondents	62	56	55	61	58	60	64	60
Percent White								
Eligibles	52	72	83*	76	84	87	77	80
Respondents	39	77	87	77	86	87	75	80
Percent Medicaid Enrolled								
Eligibles	20	15	12*	13	11	14	14	0
Respondents	18	11	7	12	9	12	12	0
Percent Aged without ESRD								
Eligibles	90	92	93	89	87	89	83	93
Respondents	86	94	93	89	87	90	85	93
Health Status and Use								
Average PIP-DCG Risk Score								
Eligibles	1.02	0.92	0.96*	0.91	0.89	0.9	0.84	0.75
Respondents	0.9	0.87	0.9	0.89	0.87	0.85	0.84	0.75
Average Number of Hospitalizations								
Eligibles	0.31	0.17	0.25	0.21	0.19	0.22	0.17	0.07
Respondents	0.06	0.2	0.23	0.19	0.19	0.17	0.17	0.07
Average Number of Inpatient Days								
Eligibles	13.1	7.3	10.4	9.4	7.5	7.7	7.6	1.0
Respondents	4.7	5.3	8	8.9	7.5	5.8	7.4	1.0

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents are made within decile of response rate category using a two-sided z-test of differences at the significance level of $p<0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 CAHPS Medicare+Choice (M+C) Disenrollment Reasons Survey (3rd quarter only).

CHAPTER 7
ANALYSIS OF NON-RESPONSE BIAS IN THE 2000 MEDICARE CAHPS® FEE-FOR-SERVICE (FFS) SURVEY

7.1   Description of the CAHPS® Fee-for-Service Survey

The CAHPS® family of surveys is conducted on an annual basis to fulfill a requirement of Congress (under the Balanced Budget Act of 1997 and the Medicare Prescription Drug, Improvement, and Modernization act of 2003) to provide information to Medicare beneficiaries on the quality of health care services provided through the Medicare program. The information collected in the surveys of beneficiaries is intended to allow beneficiaries to compare the information on experiences in Original Fee-for-Service (FFS) Medicare to similar information collected from beneficiaries enrolled in Medicare+Choice (managed care health) plans. In fall 2000, the CAHPS® Medicare Disenrollment and Medicare Fee-for-Service Surveys were being conducted for the first time, while the Medicare Managed Care CAHPS® Survey was in its fourth implementation.

The primary mode of data collection for the CAHPS® FFS Survey was a self-administered mail survey. The option to complete the survey by telephone was offered to provide a way to include sample members for whom completing a written survey might not be possible—for example, sample members with vision, reading, or other impairments that might otherwise preclude their participation. A Spanish version of the questionnaire was available on request, and there were bilingual interviewers able to complete Spanish language interviews by telephone. The follow-up data collection effort for non-respondents to the mail survey included a telephone follow-up of non-respondents for whom a telephone number was available and an overnight mailing to other non-respondents.

The data collection plan for this mail survey followed the traditional method of mailing an advance letter, followed by a survey package, followed by a thank you/reminder letter. These initial contacts were followed by a replacement survey, which was mailed to non-respondents about 2 weeks after the thank you/reminder letter. A final or third wave survey was sent by overnight mail to provide a "last chance" for non-respondents to participate. The third wave mailing was sent 5 weeks after the second wave mailing to help reduce overlap in the returns. The data collection period for the CAHPS® FFS Survey started with the mail-out of the pre-notification letter on October 9, 2000, and ended with the close of the telephone follow-up on February 1, 2001.

The sample of 167,993 beneficiaries selected for the 2000 FFS Survey was drawn from a sampling frame constructed from the August 2000 version of the Medicare Enrollment Data Base (EDB). The frame comprised 30.1 million persons who were enrolled in Medicare FFS for at least the prior 6 months and who resided in the United States or Puerto Rico. However, the frame also included the following beneficiaries who were determined to be ineligible for the survey:

- beneficiaries under the age of 18

- sample members who self-reported that they were not in Medicare FFS

- beneficiaries who died before or during data collection

After selecting the FFS sample, 5,863 beneficiaries were removed from the sample as ineligible because they either had died, indicated they were not in FFS Medicare, or were under 18 years of age. This reduced the number of eligibles in the sample to 162,130 beneficiaries. The frame also included beneficiaries who did not speak English or Spanish and beneficiaries who were mentally or physically incompetent and without access to a proxy. Even though these beneficiaries were systematically excluded from participation in the survey, they were classified as survey eligibles to be consistent with the Medicare CAHPS® Managed Care Survey. A total of 103,551 surveys was completed, resulting in a response rate of 63.9 percent.

The sample was drawn from 280 distinct geographic areas in the United States and Puerto Rico. Approximately 600 sample members were selected from each geographic area. Geographic stratification was used to vary the sampling rates of beneficiaries selected for the FFS Survey in order to achieve the design goals of the study (Elliot et al., 2000). Eight states and the District of Columbia were each assigned only one geographic unit. In the 42 states with two or more geographic units assigned, counties were agglomerated into geographic reporting units according to a hierarchical series of rules. The hierarchy of grouping rules was prioritized first to last as follows: (1) geographic contiguity; (2) managed care contract area boundaries, in order to facilitate comparison with Medicare managed care; (3) MSA boundaries; and (4) HSA boundaries.

Sampling weights enable design-consistent estimation of population parameters by scaling the disproportionalities between the sample and the population. For the FFS Survey, the weights may be viewed as inflation factors that account for the number of beneficiaries in the target population that a sample member represents. The basic component of the FFS sampling weight was the selection probability specified by the sample design. An initial sampling weight was assigned to each selected beneficiary as the inverse of the selection probability and reflects the differential selection rates that were used to select beneficiaries from each state or county. Adjustments were made to compensate for potential biases attributable to differential response and coverage among sample members using available demographic information for all sampled beneficiaries.

7.2   Survey-Specific Response Rates

We begin our detailed examination of possible non-response bias in the CAHPS® FFS Survey by first exploring differences in response rates by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. **Table 7-1** displays three sets of response rates: a rate calculated in which each inverse of the probability of selection is given equal weight (we consider this an unweighted response rate), a rate calculated as the mean of the means of the individual states' response rates, and a rate calculated using the beneficiary. Looking across the three estimates for each category of the characteristics shows them to be fairly close to one another, usually within 1 or 2 percentage points.

Table 7-1

Survey-Specific Response Rates by Demographic and Health Status Characteristics for the 2000 Medicare CAHPS® FFS Survey

Characteristic	Unweighted Response Rate[1] (%)	Mean of the Means Response Rate[2] (%)	Selection Probability Weighted Response Rate[3] (%)	
Age				
Under 65	51	53	51	*
65-74	68	70	68	
75-84	67	68	66	*
85 +	54	55	53	*
Race				
Unknown	53	56	52	*
White	66	67	66	
Black	51	48	50	*
Other	50	54	49	*
Asian	54	55	51	*
Hispanic	50	50	49	*
American Indian	49	53	51	*
Gender				
Male	65	66	64	
Female	63	65	62	*
Medicaid Status				
Not Enrolled	66	68	65	
Enrolled	51	53	50	*
Reason for Medicare Entitlement				
Aged without ESRD	66	67	65	
Aged with ESRD	63	64	62	
Disabled without ESRD	51	53	51	*
Disabled with ESRD	54	55	53	*
ESRD Only	45	45	47	*
Risk Score Decile				
0.36 - 0.45	63	65	63	*
0.46 - 0.53	66	68	65	*
0.54 - 0.57	70	72	69	*
0.58 - 0.70	68	70	67	*
0.71 - 0.73	67	69	66	*
0.74 - 0.87	67	69	66	*
0.88 - 0.91	63	64	63	*
0.92 - 1.07	60	62	59	
1.08 - 1.26	56	59	55	*
1.27 - 6.91	58	59	58	*
Number of Hospitalizations				
Zero	64	66	63	
One	64	65	63	
Two	63	64	62	
Three or More	57	58	56	*

[1] An equal weight whereby all *sampled beneficiaries* are given a weight of 1.
[2] An equal weight whereby all *health plans or states* are given a weight of 1.
[3] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of p<0.05 with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

Pair-wise comparisons of differences in weighted response rates between the categories of each variable and the reference level are made using a two-sided z-test for proportions with a significance level of $p<0.05$ and using the Bonferroni multiple comparison adjustment. We summarize our findings below, focusing on the response rates in the weighted column:

- With very few exceptions, and likely because of the large sample size, the weighted distribution of response rates differs significantly by category of enrollment, demographic, health status, and use measures.

    – The response rates of beneficiaries under age 65 and above age 74 are significantly lower than those for beneficiaries 65 to 74 years of age. This finding is particularly true for the youngest and oldest age groups.

    – The response rates for all of the racial/ethnic groups are significantly lower than for Whites.

    – The response rate for males is significantly higher than for females, although the difference is small.

    – Beneficiaries not dually enrolled in Medicare and Medicaid have a significantly higher response rate than dually enrolled beneficiaries.

    – Beneficiaries entitled to Medicare because they are disabled or because of ESRD only responded at a significantly lower rate than aged beneficiaries without ESRD.

    – Beneficiaries with a PIP-DCG score lower than the category containing 1.00 (in better health because they are below the mean) have a significantly higher response rate than those in the category containing 1.00, and those in categories with PIP-DCG scores higher than the category containing 1.00 (in worse health than average because they are above the mean) have significantly lower response rates.

    – Only beneficiaries who had three or more hospital discharges have a significantly lower response rate than those who had not been hospitalized in the prior year.

7.3   Differences in Characteristics of Respondents and Non-respondents

Table 7-2 presents the distribution of the total eligible sample, respondents, and non-respondents across the categories or levels of the enrollment, demographic, health status, and use measures. The distributions presented are weighted and reveal some meaningful differences in the distributions between respondents and non-respondents. This review provides an overall sense of how different respondents are from non-respondents in terms of their demographic and health status characteristics and is a critical factor in the determination of the potential adverse effects of non-response bias. Statistical significance testing is performed using the chi-square test and $p<0.05$ level of significance. Our findings are summarized below:

110

Table 7-2
Distribution of Demographic and Health Status Characteristics among
2000 Medicare CAHPS® FFS Survey Eligibles, Respondents, and Non-respondents,
Selection Probability Weighted[1]

Characteristic	Eligibles (%)	Respondents (%)	Non-respondents (%)	
Age[2]				
Under 65	13	10	18	*
65-74	43	46	38	
75-84	32	34	30	
85 +	12	10	15	
Race				
Unknown	1	0.4	0	*
White	86	88	80	
Black	9	7	12	
Other	3	2	4	
Asian	1	1	1	
Hispanic	2	1	3	
American Indian	0.2	0.1	0.3	
Gender				
Male	43	44	41	*
Female	57	56	59	
Medicaid Status				
Not Enrolled	86	89	81	*
Enrolled	14	11	19	
Reason for Medicare Entitlement				
Aged without ESRD	87	89	82	*
Aged with ESRD	0.2	0.2	0.2	
Disabled without ESRD	13	10	17	
Disabled with ESRD	0.2	0.2	0.3	
ESRD Only	0.3	0.2	0.4	
Risk Score Decile				
0.36 - 0.45	11	10	11	*
0.46 - 0.53	10	10	9	
0.54 - 0.57	10	11	8	
0.58 - 0.70	11	12	10	
0.71 - 0.73	9	9	8	
0.74 - 0.87	11	11	10	
0.88 - 0.91	9	9	10	
0.92 - 1.07	11	10	12	
1.08 - 1.26	10	8	11	
1.27 - 6.91	10	9	12	
Number of Hospitalizations				
Zero	84	85	84	*
One	10	10	10	
Two	3	3	3	
Three or More	2	2	2	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Statistical significance tested using chi-square distribution differences between respondents and non-respondents. An asterisk (*) denotes significance at <0.05 level.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

- The weighted distribution of response rates differs significantly by category of enrollment, demographic, health status, and use measures.

  – Respondents differ by age from non-respondents, with disproportionately more beneficiaries 65 to 74 years old responding, and fewer responding in the under 65, 75 to 84, and 85 and older age groups.

  – Respondents also differ by race/ethnicity from non-respondents, with disproportionately more White beneficiaries responding and fewer Black, Hispanic, and American Indian beneficiaries responding.

  – With respect to gender, male beneficiaries are slightly more likely to respond than females.

  – When classified by their Medicaid status, Medicare beneficiaries not enrolled in Medicaid are far more likely to respond than those dually enrolled in Medicare and Medicaid.

  – Beneficiaries entitled to Medicare only because they are 65 or older are most likely to respond. The disabled are less likely to respond.

  – When rates of response are arrayed according to health status as measured by PIP-DCG score deciles, with one exception, beneficiaries who are healthier (the categories including PIP-DCG scores below 0.88) are more likely to be respondents, while less well beneficiaries are less likely to respond.

  – Beneficiaries who were not hospitalized in the prior year are slightly more likely to respond than beneficiaries with prior hospitalization.

7.4   Differences in Outcomes by Demographic and Health Status Characteristics

Next, we examine differences in survey-specific outcome scores by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. For the CAHPS® FFS Survey, in **Table 7-3** we report mean physical component summary (PCS) and mental component summary (MCS) scores calculated from the SF-12 as the outcome measures. Although based on fewer items, these outcomes are theoretically similar to those selected for the HOS instrument, which uses the SF-36. If outcome measures, such as physical health, vary by demographic characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias increases. Pair-wise comparisons of differences in weighted means between the various levels of demographic and health status variables and a reference level is made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment.

There are statistically significant differences in the mean PCS and MCS scores according to categories of enrollment, demographic, health status, and use variables. Our findings are summarized below:

Table 7-3

Average Physical and Mental Health Component Scores by Demographic and Health Status Characteristics of Respondents to the Medicare CAHPS® FFS Survey, Selection Probability Weighted[1]

Characteristic	Physical Health Component Score (PCS) (mean)		Mental Health Component Score (MCS) (mean)	
Across all Respondents	38.79		53.05	
Age[2]				
Under 65	29.84	*	44.98	*
65-74	42.29		54.44	
75-84	38.09	*	53.66	*
85 +	32.77	*	52.41	*
Race				
Unknown	40.23		53.42	
White	39.04		53.40	
Black	36.19	*	50.11	*
Other	37.58	*	50.44	*
Asian	39.93		52.77	
Hispanic	35.80	*	48.82	*
American Indian	34.13	*	50.94	
Gender				
Male	39.68		53.39	
Female	38.08	*	52.79	*
Medicaid Status				
Not Enrolled	39.56		53.67	
Enrolled	32.26	*	47.85	*
Reason for Medicare Entitlement				
Aged without ESRD	39.82		53.97	
Aged with ESRD	27.70	*	49.02	*
Disabled without ESRD	29.75	*	44.91	*
Disabled with ESRD	29.47	*	45.68	*
ESRD Only	34.28	*	49.70	*
Risk Score Decile				
0.36 - 0.45	42.67	*	53.09	
0.46 - 0.53	44.29	*	54.42	*
0.54 - 0.57	42.74	*	54.80	*
0.58 - 0.70	41.52	*	53.87	*
0.71 - 0.73	39.78	*	53.76	*
0.74 - 0.87	38.07	*	52.92	
0.88 - 0.91	35.97	*	52.72	
0.92 - 1.07	34.81		52.70	
1.08 - 1.26	32.79	*	50.89	*
1.27 - 6.91	31.41	*	50.16	*
Number of Hospitalizations				
Zero	39.80		53.37	
One	34.16	*	51.84	*
Two	31.61	*	50.81	*
Three or More	29.17	*	48.60	*

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

- Across all respondents, the mean PCS score is 38.79, considerably lower than the norm-based mean of 50 based on a general population. Clearly, the Medicare population is not the general population and has a considerably lower level of self-reported physical health status than the general population.

  – Medicare beneficiaries younger and older than 65 to 74 years of age have much lower mean PCS scores than beneficiaries 65 to 74 years of age.

  – Hispanic, American Indian, and Black Medicare beneficiaries have much lower mean PCS scores than White beneficiaries.

  – Female beneficiaries have a slightly lower mean PCS score than males.

  – Beneficiaries dually enrolled in Medicare and Medicaid have a much lower mean PCS score than beneficiaries in Medicare alone.

  – Beneficiaries entitled to Medicare because of disability or ESRD have considerably lower mean PCS scores than those whose only entitlement to Medicare is because of age.

  – Compared to beneficiaries in the decile of PIP-DCG scores containing 1.00, beneficiaries in categories with lower scores (better health) have progressively higher mean PCS scores, while beneficiaries whose PIP-DCG is in the deciles with scores higher than 1.00 have progressively lower mean PCS scores.

  – Beneficiaries with hospital stays during the prior year have progressively lower mean PSC scores as the number of stays increase when compared to Medicare beneficiaries without a prior hospital stay.

- Across all respondents, the mean MCS score is 53.05, slightly higher than the norm based mean of 50 based on a general population. While we stated above that the Medicare population is not a general population, in the case of the MCS, they have a slightly higher level of self-reported mental health status than the general population.

  – Medicare beneficiaries under the age of 65 have a mean MCS score that is 10 points lower than the mean for beneficiaries age 65 to 74. A 2-point difference in mean MCS or PCS scores is considered clinically meaningful. Medicare beneficiaries older than 65 to 74 years of age also have lower mean MCS scores than beneficiaries 65 to 74 years of age.

  – Hispanic and Black Medicare beneficiaries have slightly lower mean MCS scores than White beneficiaries.

  – Female beneficiaries have a slightly lower mean MCS score than males.

  – Beneficiaries dually enrolled in Medicare and Medicaid have a lower mean MCS score than beneficiaries in Medicare alone.

114

– Beneficiaries entitled to Medicare because of disability or ESRD have considerably lower mean MCS scores than those whose only entitlement to Medicare is because of age.

– Compared to beneficiaries in the decile category of PIP-DCG scores containing 1.00 (the average score), beneficiaries in categories with lower scores (better health) have higher mean MCS scores, while beneficiaries whose PIP-DCG score is in the deciles with higher PIP-DCG scores (worse health) have lower mean MCS scores.

– Beneficiaries with hospital stays during the prior year have progressively lower mean MCS scores as the number of stays increase when compared to Medicare beneficiaries without a prior hospital stay.

In **Table 7-4**, we examine differences in a second set of survey-specific outcome scores by levels of beneficiary demographic and enrollment characteristics, health status, and medical care use rates. For the CAHPS® FFS Survey, we display estimates of the average rating of respondents to two of the nine CAHPS® measures of satisfaction—rating of the Original FFS Medicare and satisfaction with getting care when needed—as the outcome measures. If outcome measures, such as satisfaction with care, vary by demographic characteristics and there are systematic differences in the distribution of characteristics between respondents and non-respondents, then the likelihood of non-response bias existing increases. Pair-wise comparisons of differences in mean satisfaction levels between the various levels of the demographic and enrollment variables and a reference level is made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. Because of the large sample sizes in the CAHPS® FFS Survey, many comparisons are statistically significant. We reserve our comments to the largest differences. Our findings are summarized below:

- With respect to beneficiaries' rating of their satisfaction with Original FFS Medicare, there are a number of variables within whose categories there are meaningful statistically significant differences.

  – Persons under 65 (the disabled) rated satisfaction with Medicare lower than persons in the 65 to 74 age category, while those over 74 self-reported higher rates of satisfaction than beneficiaries age 65 to 74.

  – Women rated Medicare higher than men.

  – Persons who are entitled to Medicare because of their disability (without ESRD) or because of ESRD only rated Medicare lower than those entitled because they are aged (without ESRD).

  – Beneficiaries with a PIP-DCG risk score in a category lower than the one that includes 1.00 (better health) rated Medicare lower than those in the category including 1.00.

115

Table 7-4

Mean CAHPS® Plan Satisfaction Rating and Mean CAHPS® Composite for Getting Care When Needed for 2000 Medicare CAHPS® FFS Survey Respondents for Levels of Demographic, Health Status, and Utilization Measures, Selection Probability Weighted[1]

Characteristic	Self Reported Satisfaction with Plan			Self Reported Satisfaction with Getting Care when Needed		
	Mean	SE	Number of Cases	Mean	SE	Number of Cases
Overall	8.7		97,924	2.8		88,985
Age[2]						
Under 65	8.0	0.024	10,178 *	2.7	0.005	9,784 *
65-74	8.6	0.009	45,026	2.8	0.002	40,292
75-84	8.9	0.009	33,244 *	2.9	0.002	30,338 *
85 +	8.9	0.017	9,476 *	2.8	0.004	8,571
Race						
White	8.7	0.006	86,755	2.8	0.001	78,962
Black	8.7	0.025	6,555	2.8	0.006	5,849
Other	8.6	0.030	4,614	2.7	0.008	4,174
Gender						
Male	8.5	0.009	42,747	2.8	0.002	38,673
Female	8.8	0.007	55,177 *	2.8	0.002	50,312
Medicaid Status						
Not Enrolled	8.7	0.006	86,910	2.8	0.001	78,952
Enrolled	8.7	0.019	11,104	2.7	0.005	10,033 *
Reasons for Medicare Entitlement						
Aged without ESRD	8.8	0.006	87,431	2.8	0.001	78,904
Aged with ESRD	8.8	0.118	220	2.8	0.024	216
Disabled without ESRD	7.9	0.024	9,934 *	2.7	0.005	9,537 *
Disabled with ESRD	8.7	0.136	162	2.8	0.034	152
ESRD Only	8.2	0.173	169 *	2.8	0.025	169
Risk Score Quintile						
0.36 - 0.53	8.3	0.014	19,325 *	2.8	0.003	17,219
0.54 - 0.70	8.6	0.012	22,485 *	2.8	0.003	20,284
0.71 - 0.87	8.8	0.012	19,866 *	2.8	0.003	18,136
0.88 - 1.07	8.9	0.012	18,994	2.8	0.003	17,330
1.08 -6.91	8.8	0.013	17,254 *	2.8	0.003	16,016
Number of Hospitalizations						
Zero	8.7	0.006	82,563	2.8	0.001	74,496
One	8.8	0.017	10,345 *	2.8	0.004	9,718
Two	8.8	0.031	3,231 *	2.8	0.007	3,059
Three or More	8.7	0.043	1,785	2.8	0.010	1,712

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons of differences are made using a two-sided z-test at the significance level of $p<0.05$ with the Bonferroni multiple comparison adjustment. An asterisk (*) denotes those comparisons that exceed the specified significance level. The reference stratum within each set of characteristics is in bold.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

– Persons with one or two hospital stays in the prior year had higher levels of satisfaction than persons with none.

- With respect to beneficiaries' reported level of satisfaction with getting needed care, there are fewer variables within whose categories there are meaningful statistically significant differences.

  – Beneficiaries under 65 years of age reported slightly lower satisfaction with getting needed care than those 65 to 74 years of age, while persons 75 to 84 years of age reported slightly higher satisfaction.

  – Persons dually enrolled in Medicare and Medicaid had a lower level of satisfaction with getting needed care than those not also enrolled in Medicaid.

  – Disabled persons (without ESRD) had a lower level of satisfaction with getting needed care than aged persons (without ESRD).

7.5   Factors that Predict Probability of Response

With the next analysis, we predict the probability of response as a function of demographic and enrollment characteristics of all eligible sampled beneficiaries using a multivariate logistic regression model. We estimate the model unweighted and weighted by the inverse of the probability of the beneficiary being selected for the survey in the sampling unit. **Table 7-5** contrasts the statistically significant odds ratios resulting from a logistic regression model intended to predict response that is not weighted and one that is weighted.

- There are five fewer statistically significant categories in the model without weights.

- The direction and magnitude of the odds ratios for the included beneficiary-level variables are consistent with the descriptive comparisons between respondents and non-respondents. The results from the weighted regression model are as follows:

  – Persons under age 65 have about 45 percent of the odds of responding as persons 65 to 74 years of age.

  – Persons age 85 and over have about 40 percent lower odds of responding as those age 65 to 74.

  – Males have about 5 percent higher odds of responding than females.

  – All of the race/ethnic categories have from 30 percent to 40 percent lower odds of responding than do Whites.

  – Beneficiaries dually enrolled in Medicare and Medicaid have about 30 percent lower odds of responding than those only enrolled in Medicare.

117

Table 7-5
Logistic Regression of Probability of Response to the 2000 Medicare CAHPS® FFS Survey

Characteristic	Unweighted Regression Odds Ratio[1]	Selection Probability Weighted Regression Odds Ratio[2]
Beneficiary Characteristics		
Under 65	0.550	0.555
75 to 84	0.907	0.914
85 +	0.606	0.611
Black	0.627	0.621
Unknown or Other Race	0.620	0.603
Asian	0.699	0.657
Hispanic	0.596	0.588
American Indian	0.605	0.672
Male	1.054	1.056
Medicaid	0.699	0.692
ESRD	1.034	1.083
Risk Score Decile		
0.36 - 0.45	0.980	0.964
0.46 - 0.53	1.033	1.010
0.54 - 0.57	1.145	1.136
0.58 - 0.70	1.166	1.162
0.71 - 0.73	1.129	1.108
0.74 - 0.87	1.258	1.240
0.88 - 0.91	1.067	1.064
1.08 - 1.26	1.032	1.014
1.27 - 6.91	0.878	0.881
Number of Hospitalizations		
One	1.156	1.160
Two	1.190	1.177
Three or More	1.014	1.015
No. of Observations	162130	162130
Overall Chi-Sq (p-value)	5542***	1021829***

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Highlighted odds ratios are statistically significant at the <0.05 level of significance. Asterisks (***) denote p<0.001 level of significance.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

– Persons with PIP-DCG risk scores lower (healthier) than the category containing 1.00 have slightly higher odds of responding than persons in the categories at or above 1.00 (less healthy).

– Beneficiaries with one or two hospital stays during the prior year have slightly more than 15 percent higher odds of responding than those with none.

7.6   Probable Degree of Non-response Bias

In **Table 7-6**, we indirectly examine the degree of bias present in estimates of health status and medical care usage by comparing means of these measures for respondents to those obtained for total sample of eligible beneficiaries, including non-respondents. Pair-wise comparisons of differences in weighted mean estimates between eligibles and respondents are made using a two-sided z-test for differences in means or proportions at the significance level of $p<0.01$ to account for multiple comparisons. We consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

- Differences in the mean health status and medical use statistics between eligibles and respondents reflect the differences previously observed in the underlying distribution of characteristics of respondents and non-respondents, suggesting that respondents, on average, have a modestly higher level of health status than the surveyed population. We draw this conclusion given that we have previously observed a negative correlation between PCS and PIP-DCG risk scores.

    – The mean PIP-DCG score is 2 percent lower for respondents than for survey eligibles, implying modestly better health status. Thus, health status estimates derived for respondents only will overstate the average health status of all surveyed beneficiaries.

    – Similarly, the mean hierarchical coexisting condition (HCC) risk score for respondents is also modestly lower than scores calculated for all eligible beneficiaries. The HCC score uses diagnoses from all types of claims, rather than just the principal diagnosis from inpatient hospitalizations. As with the PIP-DCG score, the HCC risk score is centered around 1.0, with lower scores indicating better health status.

    – In contrast, the Charlson comorbidity index is a construct where increasing numbers of comorbidities or increasing severity of comorbidity yields increasing scores. Thus, it is surprising to see that respondents have poorer health status, on average, using this particular claims-based measure of health status. An evaluation of the distribution of scores between respondents and non-respondents showed that roughly 50 percent of both respondents and nonrespondents had no comorbidities during the year prior to survey; thus, over 50 percent of both groups had a Charlson cormorbidity index of zero. In fact, the distribution of scores up to the 99th percentile was virtually identical for both groups of beneficiaries. Five respondents have scores in excess 15, ranging up to 22; while the top five

119

Table 7-6

Average 1999 Claims-Based Health Status and Medical Care Use among Survey Eligibles, Respondents, and Non-respondents to the 2000 Medicare CAHPS® FFS Survey, Selection Probability Weighted[1]

Analytic Variable	Eligibles	Respondents	Non-respondents	Degree of Bias Difference in Means[2]
Health Status				
PIP-DCG Risk Score	0.97	0.95	1.02	-0.02 *
HCC Risk Score	0.92	0.91	0.93	-0.01 *
Charlson Comorbidity Index	0.98	1.02	0.9	0.04 *
Mean 1999 Payments for Medical Services				
Hospital	$11,324	$10,822	$12,151	-502 *
Professional Services	732	723	750	-9
Nursing Home	273	229	301	-44 *
Durable Medical Equipment	651	605	749	-46
Home Health	2,776	2,647	2,982	-129
All Other Services	3,608	3,271	4,298	-337 *
Mean Number of Hospitalizations in 2000	0.25	0.24	0.27	-0.01 *
Mean Number of Inpatient Days in 2000	9.33	8.32	10.98	-1.01 *

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of $p<0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

non-respondents have scores that are 15 or 16. This results in a small number of respondents skewing the data and raising the average of the respondents above the non-respondents when in actuality the distributions of scores look quite similar between the two sets of beneficiaries.

– Mean 1999 payments for medical services are also underestimated when considering only respondents. Eligibles for the 2000 Medicare CAHPS® FFS Survey had roughly, on average, $500 more in expenditures in 1999 than respondents. The degree of bias varies across types of providers, with average physician estimates being quite similar between eligibles and respondents and least similar for the category of "All Other Services."

– Mean number of hospitalizations and inpatient days are both lower for respondents than for survey eligibles, once again understating the health status of the eligible survey population when deriving estimates from respondents only.

We also explore differences in mean PIP-DCG scores between respondents and survey eligibles by demographic and enrollment characteristics in **Table 7-7.**

- Differences in the mean health status (as represented by the PIP-DCG risk score) between survey eligibles and respondents display a general trend in which health status estimates for respondents derived using the PIP-DCG risk score are modestly lower (better health) than those derived for survey eligibles across most major subpopulations of Medicare beneficiaries.

  – As noted earlier, health status estimates derived from claims for respondents only relative to the entire pool of sample eligibles tend to modestly overestimate the health of Medicare FFS beneficiaries. The difference is statistically significant for beneficiaries of both genders, ages 65 to 84, and dually enrolled in Medicare and Medicaid, and for Whites.

  – There is one noted exception. Respondents who are entitled to Medicare because of a disability produce an average health status estimate that is 2 percent worse than an estimate derived for all survey eligibles.

And, finally, in **Table 7-8** we examine the differences between the pool of sample eligibles and respondents by state response rate deciles to investigate whether there is a state response rate below which respondents are an unrepresentative sample of survey eligibles. We compare eligibles and respondents with response rate deciles on average age; the proportion that are female, White, enrolled in Medicaid, and aged without ESRD; as well as average PIP-DCG risk score, mean number of hospitalizations, and mean number of inpatient days.

- A comparison of the differences between eligibles and respondents by state-level response rate deciles does not immediately suggest that there is a response rate below which respondents are an unrepresentative sample of survey eligibles. In fact, there are only small observed differences between eligibles and respondents for the states with the lowest level of response. Although we observe statistically significant

Table 7-7

Average PIP-DCG Score for Eligibles and Respondents by Beneficiary and Enrollment Characteristics, 2000 Medicare CAHPS® FFS Survey, Selection Probability Weighted[1]

Characteristic	Eligibles Mean	Respondents Mean	Degree of Bias Difference in Means[2]	
Total	0.97	0.95	-0.02	*
Age				
Under 65	0.90	0.92	0.02	*
65-74	0.76	0.74	-0.02	*
75-84	1.11	1.09	-0.02	*
85 +	1.44	1.43	-0.01	
Race				
Unknown	0.92	0.84	-0.08	
White	0.96	0.93	-0.03	*
Black	1.07	1.05	-0.02	
Other	0.95	0.93	-0.02	
Asian	1.24	1.19	-0.05	
Hispanic	1.14	1.12	-0.02	
American Indian	1.16	1.06	-0.10	
Gender				
Male	0.99	0.98	-0.01	*
Female	0.96	0.92	-0.04	*
Medicaid Status				
Not Enrolled	0.90	0.89	-0.01	*
Enrolled	1.42	1.38	-0.04	*
Reason for Medicare Entitlement				
Aged without ESRD	0.98	0.95	-0.03	*
Aged with ESRD	1.96	1.82	-0.14	
Disabled without ESRD	0.88	0.90	0.02	*
Disabled with ESRD	1.61	1.70	0.09	
ESRD Only	1.07	1.19	0.12	

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of p<0.01 to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

Table 7-8

Average Demographic and Health Status Characteristics of Eligibles and Respondents by Decile of State Response Level to the 2000 Medicare CAHPS® FFS Survey, Selection Probability Weighted[1]

Analytic Variables	Level of State Response		
	41-50%	61-70%	71-80%
Number of States (Total = 52)	10	32	10
Number of Beneficiaries (Total=162,130)	37,394	110,769	13,967
Demographics			
Average Age[2]			
Eligibles	73	73	74
Respondents	73	73	74
Percent Female			
Eligibles	58	57	57
Respondents	57	57	57
Percent White			
Eligibles	79*	86*	96*
Respondents	83	88	97
Percent Medicaid Enrolled			
Eligibles	18*	14*	9*
Respondents	14	11	8
Percent Aged without ESRD			
Eligibles	86*	87*	90*
Respondents	89	89	92
Health Status and Use			
Average PIP-DCG Risk Score			
Eligibles	0.99*	0.97*	0.92*
Respondents	0.96	0.95	0.90
Average Number of Hospitalizations			
Eligibles	0.24	0.25*	0.22
Respondents	0.23	0.24	0.21
Average Number of Inpatient Days			
Eligibles	10.6	9.1*	7.5
Respondents	9.1	8.3	6.9

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Pairwise comparisons between eligibles and respondents are made within decile of response rate category using a two-sided z-test of differences at the significance level of $p < 0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey.

differences between eligibles and respondents for some subpopulations, the level of difference is small. The statistical difference is likely a function of the very large sample size for this survey.

- There do not appear to be differences in the mean ages or percent female within all levels of state response rate.

  – As the state-level response rate increases, the proportion of eligibles who are White increases as well, and the difference in the proportion of Whites between respondents and eligibles declines. The same is true for the proportion of beneficiaries who are eligible for Medicare because they are aged without ESRD.

  – The percent of eligibles dually enrolled in Medicare and Medicaid declines as state-level response rates increase. As noted with race, the difference in the proportion of dual enrollees between respondents and eligibles declines as response rates increase.

  – A similar pattern emerges from the analysis of mean PIP-DCG risk scores—as the response rate at the state level increases, the mean risk score for eligibles and respondents both decline, but the respondents are significantly less healthy across the board. Mean number of hospitalizations and hospital inpatient days follow a similar pattern, but the difference between respondents and eligibles is only significant for one of the levels of state response rate.

CHAPTER 8
ANALYSIS OF NON-RESPONSE BIAS

8.1   Introduction

The previous six chapters of this report focused on analyzing non-response to six different surveys using the response definition specific to each survey. The six surveys are as follows:

- 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS)

- 2000 Cohort 3 Baseline Medicare HOS

- 2000 CAHPS® Medicare+Choice (M+C) Enrollee Survey

- 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey

- 2000 CAHPS® M+C Disenrollment Assessment Survey

- 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only)

This chapter reports on the analysis of non-response to the six surveys using a common definition of response. Each survey contains the question, *In general, would you say your health is excellent, very good, good, fair, or poor?*

Table 8-1 displays the number of health plans (or states in the case of the CAHPS® FFS Survey) that participated in each of the six surveys. At the health plan or state level, the mean number of eligible beneficiaries across the six surveys ranged from a high of 3,118 for the 2000 CAHPS® FFS Survey to a low of 46 for the 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter).

Average response rates using the general health status question across the six surveys ranged from a low of 39 percent for the 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter) to a high of 86 percent for the 2000 Cohort 1 Follow-up HOS. For five of the surveys, the average response rates for the general health status question are quite similar to those observed for survey-specific average response rates. The 2000 CAHPS® M+C Disenrollment Reasons Survey exhibited the largest reduction in average response rate from using the general heath status question as opposed to using the definition of a respondent answering affirmatively to one of the preprinted reasons for disenrolling from the plan; average response rate fell from 58 percent to 40 percent. There is limited variation in response rates by method used to weight plan- or state-specific estimates of response.

8.2   Survey-Specific Response Rates

As was done in the other chapters, we explore differences in response rates by beneficiary demographic and enrollment characteristics, health status, and medical care use rates. **Table 8-2** displays selection probability weighted response rates. Statistical significance testing

Table 8-1
Comparison of Plan Participation, Mean Number of Sampled Beneficiaries per Participating Plan, and Response Rate to General Health Status Question

	Cohort 3 Baseline Medicare HOS	Cohort 1 Follow-up Medicare HOS	CAHPS® M+C Enrollee Survey	Medicare CAHPS® FFS Survey	CAHPS® M+C Disenrollment Assessment Survey	CAHPS® M+C Disenrollment Reasons Survey
Number of Health Plans or States (CAHPS® FFS)	306	225	292	52	279	271
Mean Number of Sampled Beneficiaries per Plan (State)	956	391	743	3,118	80	46
Response Rates						
Unweighted Response Rate (%)[1]	73	86	81	63	51	40
Mean of the Means (%)[2]	74	86	81	64	51	40
Enrollment Weighted Mean Response Rate (%)[3]	71	86	79	62	51	39

[1] An equal weight whereby all *sampled beneficiaries* are given a weight of 1.
[2] An equal weight whereby all *health plans or states* are given a weight of 1.
[3] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS), 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS), 2000 CAHPS® M+C Enrollee Survey, 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey, 2000 CAHPS® M+C Disenrollment Assessment Survey, and 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only).

Table 8-2
General Health Status Response Rates by Demographic and Health Status Characteristics, Selection Probability Weighted[1]

Characteristic	Cohort 3 Baseline Medicare HOS (%)	Cohort 1 Follow-up Medicare HOS (%)	CAHPS® M+C Enrollee Survey (%)	Medicare CAHPS® FFS Survey (%)	CAHPS® M+C Disenrollment Assessment Survey (%)	CAHPS® M+C Disenrollment Reasons Survey (%)
Age[2]	*	*	*	*	*	*
Under 65	65	82	75	49	45	45
65-74	73	88	83	67	55	42
75-84	73	87	79	65	52	36
85 +	64	81	67	52	38	28
Race	*	*	*	*	*	*
Unknown	63	81	12	51	2	41
White	73	87	80	64	53	39
Black	62	78	71	49	42	38
Other	68	85	8	48	2	37
Asian	75	88	80	50	50	36
Hispanic	63	83	87	48	66	28
American Indian	62	71	91	50	84	55
Gender	*			*	*	*
Male	71	86	79	63	52	40
Female	72	86	79	61	51	38
Medicaid Status	*	*	*	*	*	*
Not Enrolled	72	87	80	64	53	40
Enrolled	62	79	66	48	38	32
Institutionalized Status	*	*				
Community Dwelling	72	87	N/A	N/A	N/A	N/A
Long-term Institutionalized	33	51	N/A	N/A	N/A	N/A
Nursing Home Certifiable	70	77	N/A	N/A	N/A	N/A
Reason for Medicare Entitlement	*	*	*	*	*	*
Aged without ESRD	72	87	80	64	52	38
Aged with ESRD	83	83	66	62	40	27
Disabled without ESRD	65	82	75	49	45	44
Disabled with ESRD	35	0	79	51	7	50
ESRD Only	91	100	83	46	0	75
Risk Score Decile	*	*	*	*	*	*
0.36 - 0.45	74	88	84	61	55	43
0.46 - 0.53	73	88	83	64	55	46
0.54 - 0.57	74	89	83	68	55	43
0.58 - 0.70	72	88	82	66	54	39
0.71 - 0.73	72	87	82	65	55	36
0.74 - 0.87	74	86	81	65	53	37
0.88 - 0.91	73	87	79	61	51	38
0.92 - 1.07	72	86	77	58	45	36
1.08 - 1.26	65	83	70	54	46	37
1.27 - 6.91	66	81	71	56	43	32
Number of Hospitalizations	*	*	*	*	*	*
Zero	72	87	80	62	52	39
One	70	85	77	62	49	40
Two	68	82	76	61	45	28
Three or More	64	80	71	55	42	44

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Statistical significance tested using chi-square distribution differences between respondents and non-respondents. An asterisk (*) denotes significance at <0.05 level.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS), 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS), 2000 CAHPS® M+C Enrollee Survey, 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey, 2000 CAHPS® M+C Disenrollment Assessment Survey, and 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only).

is performed using the chi-square test and p<0.05 level of significance. We summarize our findings below:

- Recognizing that there are different levels of response across the six surveys, similar patterns in relative response emerge across the surveys for the key stratifying variables displayed in **Table 8-2**.

- With very few exceptions, the selection probability weighted response rates differ significantly within categories of enrollment, demographics, health status, and hospital use measures.

  - With the exception of the 2000 CAHPS® M+C Disenrollment Reasons Survey, the response rates of beneficiaries under age 65 and above age 84 are significantly lower than response rates for beneficiaries 65 to 74 years of age.

  - The response rates for Blacks are significantly lower than for Whites. Beneficiaries of Hispanic and American Indian race/ethnicity have response rates that are significantly higher than Whites in some but not all of the surveys. Asians have response rates quite close to those reported for Whites, with the exception of the 2000 CAHPS® FFS Survey.

  - Beneficiaries dually enrolled in Medicare and Medicaid have significantly lower response rates for all six surveys.

  - With the exception of the 2000 CAHPS® M+C Disenrollment Reasons Survey, beneficiaries entitled to Medicare because they are disabled (without ESRD) respond at a significantly lower rate than aged beneficiaries without ESRD.

  - With the exception of both 2000 CAHPS® M+C Disenrollment Surveys, beneficiaries with a PIP-DCG risk score above the category containing average health status (PIP-DCG score of 1.0), or in poorer health status, have a significantly lower response rate than those with average health status. Generally, response rates increase as health status improves (i.e., declining risk score).

  - With the exception of the 2000 CAHPS® M+C Disenrollment Reasons Survey, response rates decline as number of hospitalizations in the year prior to survey increases. We use hospitalizations as a proxy for health status.

8.3   Factors that Predict the Likelihood of Response

We predict the likelihood of response as a function of sociodemographic and health status characteristics of all sampled beneficiaries using the same multivariate regression model employed in the prior chapters. We estimate the model weighted by the inverse of the likelihood of the beneficiary being selected for survey. **Table 8-3** displays the odds ratio for each of the regression models using response to the general health status question as the dependent variable.

- There are a number of general patterns that emerge across the six surveys.

Table 8-3
Logistic Regression of Likelihood of Response to the General Health Status Question, Selection Probability Weighted[1]

Characteristic	Odds Ratio[2]					
	Cohort 3 Baseline Medicare HOS	Cohort 1 Follow-up Medicare HOS	CAHPS® M+C Enrollee Survey	Medicare CAHPS® FFS Survey	CAHPS® M+C Disenrollment Assessment Survey	CAHPS® M+C Disenrollment Reasons Survey
Beneficiary Characteristics						
Under 65	0.751	0.7	0.696	0.55	0.757	1.202
75 to 84	0.986	0.994	0.797	0.904	0.966	0.772
85 +	0.75	0.773	0.549	0.604	0.603	0.571
Black	0.628	0.528	0.652	0.618	0.694	0.951
Unknown or Other Race	0.786	0.815	0.022	0.6	0.017	0.876
Asian	1.142	1.114	0.998	0.666	0.904	0.951
Hispanic	0.676	0.729	1.759	0.6	1.933	0.616
American Indian	0.653	0.406	2.651	0.686	5.065	1.784
Male	0.952	0.99	0.997	1.062	1.026	1.029
Medicaid	0.847	0.857	0.593	0.695	0.58	0.706
ESRD	1.674	1.095	0.743	1.099	0.47	0.827
Risk Score Decile						
0.36 - 0.45	1.042	1.154	1.275	0.977	1.233	0.901
0.46 - 0.53	1.027	1.128	1.14	1.029	1.243	0.987
0.54 - 0.57	1.009	1.203	1.144	1.15	1.158	0.948
0.58 - 0.70	1.014	1.114	1.243	1.173	1.168	0.803
0.71 - 0.73	0.969	1.077	1.071	1.126	1.183	0.824
0.74 - 0.87	1.056	1.139	1.247	1.246	1.179	0.844
0.88 - 0.91	1.001	1.03	1.075	1.074	1.002	0.981
1.08 - 1.26	0.886	0.924	0.874	1.009	1.128	1.047
1.27 - 6.91	0.864	0.854	0.766	0.888	0.973	0.768
Number of Hospitalizations						
One	1.029	0.968	1.093	1.165	1.033	1.149
Two	0.983	0.874	1.109	1.168	0.965	0.738
Three or More	0.839	0.784	0.893	1.02	0.931	1.37
Institutionalized Status						
Long-term Institutionalized	0.247	0.218	N/A	N/A	N/A	N/A
Nursing Home Certifiable	1.073	0.599	N/A	N/A	N/A	N/A
No. of Observations	291,221	88,129	216,919	162,130	22,272	12,658
Overall Chi-Sq (p-value)	83867***	1338***	336031**	1040236***	11487***	1706***

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.

[2] Highlighted odds ratios are statistically significant at the <0.05 level of significance. Asterisks (***) denote p<0.001 level of significance.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS), 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS), 2000 CAHPS® M+C Enrollee Survey, 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey, 2000 CAHPS® M+C Disenrollment Assessment Survey, and 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only).

- With the exception of the 2000 CAHPS® M+C Disenrollment Reasons Survey, beneficiaries under the age of 65 and age 75 and older are less likely to respond than beneficiaries age 65 to 74.

- Blacks are consistently less likely than Whites to respond to any of the surveys. There is no consistent pattern of response relative to Whites across the surveys observed for the other racial minorities/ethnicities.

- Nor is there a consistent pattern of response for men relative to women across the six surveys.

- Beneficiaries dually enrolled in Medicare and Medicaid are consistently less likely to respond than beneficiaries not enrolled in Medicaid.

- After controlling for health status, race, and age, there is no consistent pattern of response for beneficiaries with ESRD relative to beneficiaries without ESRD.

- Compared to beneficiaries with average health status, beneficiaries with a high level of health status are generally more likely to respond, while beneficiaries in poor health status are less likely to respond. These findings do not hold for the 2000 CAHPS® M+C Disenrollment Reasons Survey: beneficiaries in good health are less likely to respond to the survey as compared to beneficiaries in average health; beneficiaries in poor health status show no consistent pattern of response.

- There is no consistent pattern of response for beneficiaries with increasing levels of hospitalization in the year prior to survey relative to beneficiaries who were not hospitalized.

- For the two Medicare HOS cohorts, beneficiaries residing in long-term institutionalized settings were significantly less likely to respond relative to community-based beneficiaries.

8.4   Probable Degree of Non-response Bias

We indirectly explore the degree of bias that may be present in estimates of satisfaction by using health status and medical care usage as proxies. In addition, we directly explore the degree of bias that may be present in estimates of health status in the two Medicare HOS and Medicare CAHPS® FFS Surveys using the PIP-DCG score. We compare means of these variables for respondents to those obtained for eligible beneficiaries, including non-respondents. Pairwise comparisons of differences in weighted mean estimates between eligibles and respondents are made using a two-sided z-test for differences in means or proportions at the significance level of $p<0.01$ to account for multiple comparisons. We consider estimates derived for the eligible population to reflect the true population value. Thus, the difference between mean values for respondents and the eligible population is the degree of bias that is present.

Differences in the mean health status between survey eligibles and respondents display a general trend in which health status estimates derived using the PIP-DCG risk score are often lower (healthier) than those derived for survey eligibles for each of the surveys (**Table 8-4**).

Table 8-4

Average Age, Health Status, and Hospital Use among Eligibles, Respondents, and Non-respondents to the General Health Status Question, Selection Probability Weighted[1]

Analytic Variable	Mean Values			Degree of Bias
	Eligibles	Respondents	Non-respondents	Difference in Means[2]
Cohort 3 Baseline Medicare HOS				
Mean PIP-DCG Risk Score	0.90	0.88	0.94	-0.02 *
Mean Number of Hospitalizations	0.2	0.2	0.2	0.00 *
Mean Number of Inpatient Days	7.2	6.8	7.9	-0.40 *
Cohort 1 Follow-up Medicare HOS				
Mean PIP-DCG Risk Score	0.91	0.90	0.99	-0.01 *
Mean Number of Hospitalizations	0.2	0.2	0.3	0.00 *
Mean Number of Inpatient Days	7.1	6.8	8.4	-0.30
CAHPS® M+C Enrollee Survey				
Mean PIP-DCG Risk Score	0.88	0.86	0.94	-0.02 *
Mean Number of Hospitalizations	0.2	0.2	0.2	0.00 *
Mean Number of Inpatient Days	7.1	6.6	8.7	-0.50 *
Medicare CAHPS® FFS Survey				
Mean PIP-DCG Risk Score	0.97	0.94	1.02	-0.03 *
Mean Number of Hospitalizations	0.3	0.2	0.3	-0.10 *
Mean Number of Inpatient Days	9.3	8.3	10.9	-1.00 *
CAHPS® M+C Disenrollment Assessment Survey				
Mean PIP-DCG Risk Score	0.91	0.87	0.95	-0.04 *
Mean Number of Hospitalizations	0.2	0.2	0.2	0.00 *
Mean Number of Inpatient Days	8.9	9.9	8.0	1.00
CAHPS® M+C Disenrollment Reasons Survey				
Mean PIP-DCG Risk Score	0.91	0.87	0.94	-0.04 *
Mean Number of Hospitalizations	0.2	0.2	0.2	0.00
Mean Number of Inpatient Days	8.8	7.9	9.3	-0.90

[1] A selection probability weight whereby all beneficiaries are given a weight based upon the likelihood of selection. The formula for calculating the selection probability weight is the inverse of the number of beneficiaries sampled divided by the number of beneficiaries eligible for selection from the health plan or state.
[2] Pairwise comparisons between eligibles and respondents are made using a two-sided z-test of differences at the significance level of $p < 0.01$ to account for multiple comparisons. An asterisk (*) denotes those comparisons that exceed the specified significance level.

Source: RTI analysis of the 2000 Cohort 1 Follow-up Medicare Health Outcomes Survey (HOS), 2000 Cohort 3 Baseline Medicare Health Outcomes Survey (HOS), 2000 CAHPS® M+C Enrollee Survey, 2000 Medicare CAHPS® Fee-for-Service (FFS) Survey, 2000 CAHPS® M+C Disenrollment Assessment Survey, and 2000 CAHPS® M+C Disenrollment Reasons Survey (3rd quarter only).

This suggests that health status estimates derived from respondents only tend to modestly overestimate the health of beneficiaries participating in all six surveys. Thus, modest non-response bias appears to be present, ranging from 1 percent to 4 percent.

Further, there are virtually no differences in mean number of hospitalizations between eligibles and respondents and only modest differences in mean number of inpatient days. The pattern is very similar across all six surveys.

8.5   Summary and Recommendations

The goal of this study was to examine the potential degree of non-response bias in two major survey efforts, the Medicare HOS and the Medicare CAHPS® Surveys, that collect information from five different Medicare beneficiary populations. Survey non-response is important because it may introduce bias and threaten the validity of estimates from sample surveys. The unique contribution of this project to non-response bias research has been to create and analyze two claims-based measures for both survey respondents and non-respondents. One of the measures is of health status—the PIP-DCG risk score—and the other is a measure of hospital utilization. To the extent that a claims-based measure of health status is a reasonable proxy for self-reported health status (a measure obtained in the HOS and the CAHPS® Surveys), we have directly assessed the degree of non-response bias in estimates of health status when they are based on respondents only. Similarly, if health status is correlated with measures of satisfaction and experiences with care, we have indirectly assessed the degree of non-response bias for satisfaction estimates from the CAHPS® Surveys based on respondents alone. Further, to the extent that hospital utilization is a reflection of poorer health status, we have assessed the extent of non-response bias in hospitalization estimates for respondents relative to all eligible beneficiaries in the study samples.

As discussed earlier in this report, there are two types of non-response in surveys. One type occurs when a selected sample member does not respond at all to the survey. The second occurs when a selected sample member responds to some items but fails to answer all of them. Typically, the first type is referred to as survey non-response and the second as missing data or item non-response. Non-response bias is the systematic difference between the outcome scores for survey respondents and the (unknown) scores that would have been obtained if all subjects had completed the entire survey. The degree of bias is determined by two factors: (1) the difference in characteristics of interest (e.g., health status) between respondents and non-respondents, and (2) the non-response rate.

Across all six surveys, survey-specific response rates ranged from the mid-fifties for the two M+C Disenrollment Surveys to a high of 85 percent for the Follow-up Medicare HOS Survey. Mean PIP-DCG risk scores were 6 percent to 9 percent lower for respondents than for non-respondents. Thus, respondents are generally healthier than non-respondents. Similar levels and patterns of differences in utilization measures were also observed across the surveys. However, the degree of non-response bias at the survey level, for the range of response rates observed across the six surveys, is relatively modest. Mean PIP-DCG risk scores were 2 percent to 4 percent lower for respondents than for survey eligibles.

We did observe a general pattern that certain subpopulations consistently had low response rates and poor health status. Beneficiaries under the age of 65 and age 85 and older are less likely to respond than beneficiaries age 65 to 74. Blacks are consistently less likely than Whites to respond to any of the surveys. Beneficiaries dually enrolled in Medicare and Medicaid are consistently less likely to respond than beneficiaries not enrolled in Medicaid. And, for the two Medicare HOS cohorts, beneficiaries residing in long-term institutionalized settings are significantly less likely to respond relative to community-based beneficiaries. Further, beneficiaries without these characteristics but in poor health are also less likely to respond than beneficiaries of average health status.

Because many of these special populations represent a small proportion of all sampled beneficiaries within each of the surveys, the influence of their significantly lower rate of response and health status on the overall response rate and mean health status estimate at the survey level is muted. Using the Cohort 3 Baseline Medicare HOS as an example, dual Medicare and Medicaid enrollees had a 60 percent response rate as compared to a 70 percent response rate for non-dual enrollees. Dual Medicare and Medicaid enrollees had an average PIP-DCG risk score of 1.35 as compared to a PIP-DCG risk score of 0.88 for beneficiaries not dually enrolled. Because dual Medicare and Medicaid enrollees represent only 5 percent of all sampled beneficiaries, the overall survey response rate is modestly influenced by only one-half of a percent (69.5% = (60%*0.05 + 70%*0.95)). A similarly small influence on average health status is also observed; the mean PIP-DCG risk score derived for all survey eligibles is only 2 percent (90.35 = (1.35*0.05 + 0.88*0.95)) higher than the non-dual Medicare and Medicaid enrollees' mean score. Given the limited influence that dual Medicare and Medicaid enrollees exert on survey-level mean health status estimates, any significant difference in health status between respondents and non-respondents will simply not have any appreciable influence on survey-level estimates.

Of more concern would be within subpopulation analyses as well as analyses focusing on health plans with large proportions of these special populations. Using the Cohort 3 Baseline Medicare HOS again as an example, dual Medicare and Medicaid enrollees who responded had an average PIP-DCG risk score of 1.31 as compared to a PIP-DCG risk score of 1.40 for dual enrolled beneficiaries who did not respond. Because 40 percent of dual enrollees are non-respondents, the mean PIP-DCG risk score for dual enrollees would be underestimated by 4 percentage points.

The four Medicare CAHPS® surveys analyzed in this study adjust all survey-derived estimates for non-response, taking the general approach of using predicted response propensities to adjust initial design-based weights (the inverse of the selection probability) upward for respondents so that they represent both respondents and non-respondents. Sampling weights enable design-consistent estimation of population parameters by scaling the disproportionalities between the sample and the population using available demographic information for all sampled beneficiaries. Sampling weights are not constructed for the Medicare HOS.

Given the modest degree of nonresponse bias observed in this study among the Medicare surveys, efforts to enhance the current Medicare CAHPS® sampling weights by including measures of health status or medical service use as a proxy for health status do not appear warranted. Consideration could be given to the construction of selection probability weights to

133

scale the disproportionalities between the sample and the population for the Medicare Health Outcomes Survey. As with the Medicare CAHPS® sampling weights, demographic information readily available would appear to be reasonable weighting variables. Care should be exercised when conducting analyses within subpopulations that experience high rates of nonresponse and exhibit significant differences between respondents and non-respondents in the analytic variable of interest. One needs to recognize that significant non-response bias could exist.

REFERENCES

Elliot, M. (October 2001). "Analysis of Case-Mix Strategies and Recommendations." Chapter 1 in V. Iannacchione et al., *Implementation of the Medicare CAHPS Fee-for-Service Survey.* Final Report for Year 1, prepared for the Centers for Medicare & Medicaid Services.

Folsom, R.E. (1991). "Exponential and Logistic Weight Adjustments for Sampling and Nonresponse Error Reduction." Proceedings of the *American Statistical Association, Social Statistics Section.*

Iannacchione, V.G., J.G. Milne, and R.E. Folsom (1991). "Response Probability Weight Adjustments Using Logistic Regression." Proceedings of the *American Statistical Association, Section on Survey Research Methods,* pp. 637-642.

McCall, N.T., J. Harlow, and D.A. Dayhoff (Spring 2001). "Rates of Hospitalization for Ambulatory Care Sensitive Conditions in the Medicare+Choice Population." *Health Care Financing Review* 22(3):127-145.

National Committee for Quality Assurance (NCQA) (2000). *HEDIS 2000, Volume 6: The Medicare Health Outcomes Survey Manual.* Washington, DC: NCQA.

Pope, G.C., C.F. Liu, R.P. Ellis, et al. (February 1999). *Principal Inpatient Diagnostic Cost Group Models for Medicare Risk Adjustment.* Final Report, HCFA Contract No. 500-95-0048.