NONRESPONSE TO THE 2008 MEDICARE CAHPS SURVEY

Marc N. Elliott, RAND
A.K. Burkhart, RAND
Amelia M. Haviland, RAND
Alan M. Zaslavsky, Harvard Medical School
Nate Orr, RAND
Lawrence S. Zaborski, Harvard Medical School

INTRODUCTION

Surveys are subject to *unit* and *item* nonresponse. *Unit nonresponse* is the failure of an eligible member of the sample to provide a usable response to the survey as a whole. *Item nonresponse* is the failure of a unit respondent to answer specific survey items that should have been answered. Nonresponse that is related to the quantities being measured can bias survey findings if appropriate adjustments are not made. In this paper, we describe patterns of unit and item nonresponse to the 2008 Medicare CAHPS Survey in order to investigate overall patterns and factors influencing nonresponse. Knowledge about factors associated with nonresponse can then be used to improve sampling design and survey field procedures in order to increase representation of groups with historically low response rates.

A number of studies have examined the characteristics of individuals who are less likely to respond to health surveys. Several studies have found nonrespondents more likely to be male (Burroughs et al 1999; Elliott et al 2005; Mishra et al 1993; Ware and Berwick 1990). Elliott et al 2008 found that both younger and very old adults are less likely than adults of intermediate age to respond to surveys of hospital experiences. Zaslavsky, Zaborski, and Cleary (2002) and Kahn et al (2003) found higher nonresponse for all race/ethnic groups relative to non-Hispanic whites. Kahn et al (2003) found lower response rates for HMO members with lower educational attainment and income. Several studies have found that nonrespondents to health surveys are less healthy than respondents (Cohen and Duffy 2002; Mishra et al 1993; Paganini-Hill et al 1993) and more likely to have declined in health recently (Kahn et al 2003). Limited available information suggests that nonrespondents may have less favorable perceptions of care than respondents, suggesting that a failure to correct for nonresponse may underestimate negative experiences with care (Elliott et al 2005; Elliott et al 2008). Recently, Neuman et al (2007) in a survey of Medicare beneficiaries 65 and older found nonresponse "higher among seniors who were racial/ethnic minorities, had fewer years of education, were age eighty-five and older, or were in poorer health."

There is also evidence that multimode approaches, such as mail with telephone follow-up, reduce nonresponse bias because certain members of the population are more likely to respond to each mode of data collection (Fowler et al 2002; Zaslavsky, Zaborski, and Cleary 2002). For example, older beneficiaries are more likely to respond by mail than by telephone (Elliott et al 2008; Zaslavsky, Zaborski, and Cleary 2002). Because Medicare CAHPS is fielded as a mail survey with telephone follow-up of initial nonrespondents, it is useful to describe which characteristics are associated with preference of one mode over another, bearing in mind that early response and response by mail are confounded in such a design.

In addition to unit nonresponse, item nonresponse erodes usable sample sizes and raises issues of potential bias. Termination or breakoff from telephone interviews can be one of the most important sources of missing data for items. For example, Keeter et al (2006) report 11% breakoff of phone interviews, with higher breakoff rates for respondents 65 and older and those with less educational attainment. "Don't know" responses and evidence of disengagement were predictive of breakoff, consistent with previous suggestions that breakoff may indicate a lack of interest in the survey topic (Groves, Stanley, and Dipko 2004).

METHODS

We describe 2008 overall rates and analyzed predictors of unit and item nonresponse for the 682,088 Medicare beneficiaries who were selected to receive the MCAHPS survey and the 408,020 unit respondents (see definition below).

Survey Methods

All beneficiaries were initially contacted by mail, with phone follow-up for nonrespondents. A bilingual (English and Spanish) postcard was followed by a Spanish-Language survey for beneficiaries residing in Puerto Rico and an English-language survey for others. However, at any point, beneficiaries could request a survey in the other language. A second mailing was sent in 2008. If no mail response was received by a cut-off date, telephone follow-up in Spanish (Puerto Rico) or English (other locations) ensued, with the option for interviews in the other language available throughout the phone follow-up period. Phone follow-up included up to five calls to nonrespondents.

Beneficiaries were considered to be unit respondents if they answered at least one assessment question (evaluations of care received, rather than just their own name, etc.). Unit nonrespondents included beneficiaries who were unable to complete the survey in English or Spanish, were mentally or physically unable to respond, refused to complete the survey, or completed no assessment items. We excluded from all 2008 analyses the 10,808 cases / 1.6% who were ineligible because they were institutionalized, deceased, under 18 years of age, or duplicate cases. We also excluded otherwise eligible 866 cases / 0.1% for which we did not have information regarding urbanicity. These exclusions left a final analytic sample of 669,684, with 406,944 unit responders.

Survey Instrument

The 2008 Medicare CAHPS survey included four versions: Medicare Advantage with no prescription drug coverage (MA), Medicare Advantage with prescription drug coverage (MA-PD), original Medicare Fee-for-Service (FFS) with no prescription drug coverage, and Medicare Fee-for-Service with a free-standing prescription drug plan (FFS-PD). Although all four versions share a nearly identical core set of questions (with phrasings customized to the coverage type as appropriate), each version also included additional questions and response categories related to the enrollees' particular form of coverage.

The 2008 MA Survey contains 65 items, organized into the following sections: *Your Health Plan* (11 items), *Your Healthcare in the Last 6 Months* (11 items), *Your Personal Doctor* (7 items), *Getting Healthcare from Specialists* (4 items), *Your Medicare Rights* (3 items), and *About You* (29 items). For the 2008 MA-PD Survey, the first five sections and the last section are the same as the MA version. In addition, a penultimate section of 17 items, *Your Prescription Drug Plan*, is included. The 2008 FFS Survey contains 61 items. The first and fifth sections—*Your Health Insurance Coverage* (2 items) and *Your Medicare Experience* (9 items)—correspond to the 10-item *Your Health Plan* section in the MA survey. The next three sections (*Your Healthcare in the Last 6 Months, Your Personal Doctor, Getting Healthcare from Specialists*) and the last (*About You*) are the same as the MA version except that it is not asked whether flu shots were obtained via plans. There is no *Your Medicare Rights* section. The FFS-PD Survey includes the 6 sections on the FFS survey plus a 19-item *Medicare Prescription Drug Plan Experiences* section occupying the penultimate position. The 17 items from the MA-PD *Your Prescription Drug Plan* section, plus two additional items, comprise the FFS-PD section on prescription drug plan experiences.

Many of the items in the CAHPS survey are preceded by screener questions, so that only those beneficiaries for whom the item is relevant (i.e., those with relevant needs or experiences) are asked to answer those questions.

Sample Design[1]

In 2008, 69,562 beneficiaries were sampled from those enrolled in original fee-for-service Medicare with <u>no</u> prescription drug plan (Part D) coverage (FFS-only). 235,578 were sampled from those enrolled in FFS Medicare plus a free-standing prescription drug plan (FS-PP). The MA sample was made up of 324,369 beneficiaries with prescription drug coverage (MA-PD) and 52,579 without (MA-only).[2] The 2008 MCAHPS sample was stratified by contract for MA beneficiaries and for FFS beneficiaries with a free-standing prescription drug plan (PDP) and by state for FFS beneficiaries without a PDP. .

Predictors of Unit and Item Nonresponse

Predictor variables in models for unit and item nonresponse were beneficiary characteristics available for both respondents and nonrespondents from administrative sources. These included MA-PD, MA-Only, and FS-PD indicators; beneficiary age (18-24, 25-34, 35-44, 45-54, 55-64, 65-74, 75-80, 81-84, 18-24, 25-34, 35-44, 45-54, 55-64, 65-69, 70-74, 75-79, 80-84, and 85+ in 2008); race/ethnicity (non-Hispanic White, African American, Hispanic, Asian, and Native American/other/unknown); gender; aged without end stage renal disease (ESRD), aged with ESRD, disabled without ESRD, disabled with ESRD, ESRD only; an indicator of Low Income Subsidy (requires income less than or equal to 150% of the Federal Poverty Line and enrollment to receive the benefit) eligibility; urbanicity as defined by the Beale code (ranges from 1 "County in metro area with 1 million population or more" to 9 "Non-metro county completely rural or

---

[1] See chapter one for details.
[2] Although a PPO test sample of 7,200 beneficiaries was conducted at the same time as the main 2008 MCAHPS survey, these beneficiaries were not included in the nonresponse analysis.

less than 2,500 urban population, not adj. to metro area"[3]); and an indicator of residence in one of the U.S. Territories of Puerto Rico or U.S. Virgin Islands. State and contract were also included as covariates.

Missing values for administrative variables used to predict nonresponse were uncommon, with no such variable missing in more than 0.1% of cases. These few missing values were imputed as the mean of the contract for nonmissing cases.

Analysis of Unit Nonresponse

We estimated two multivariate mixed regression models of unit nonresponse (Models 1 and 2). To estimate the extent to which unit nonresponse varied by contract, geography (states), and their interaction, both models included contract, state, and their interaction as random effects, treating FFS-only as a single "contract." Beneficiary characteristics (the covariates described above) were added as fixed effects. Model 1 included only main effects of the beneficiary characteristics, whereas model 2 added all two-way interactions among beneficiary characteristics with significant ($p<0.05$) overall main effects in Model 1.

Analysis of Response Mode

To analyze associations between beneficiary characteristics and mode of response, we estimated a logistic regression model to assess characteristics associated with earlier response by mail relative to later response by telephone, including state and plan fixed effects, and five beneficiary characteristics (gender, age, race/ethnicity, eligibility for Low Income Subsidy, and urbanicity).

Analysis of Breakoffs

A "breakoff" was defined as a unit responder (at least one substantive question answered) who did not complete any of the demographic items in the *About You* section, 29 items in 2008. We compare rates of breakoff completes by mode, and because we anticipated that telephone mode would involve the vast majority of breakoffs, we estimated a multinomial logistic regression regarding the resolution of cases that were not first completed by mail: nonbreakoff complete, breakoff complete, or nonresponse. This multinomial logistic regression uses the same predictors as the model of response mode.

Analysis of Item Nonresponse

We first compute descriptive statistics of item eligibility and missingness by mode of response and breakoff status. To more formally model item-nonresponse among unit survey respondents, we created an item-level dataset where each observation was a unique combination of item and unit respondent, for all items a unit respondent was eligible to complete. We analyzed whether each item was answered in a multivariate logistic regression model, accounting for the clustering of items within unit respondents. The same respondent-level covariates used to predict unit nonresponse were used in this model as for the model of response mode. In addition, we included

_____

[3] Top coded at 5 ("Urban population of 20,000 or more", not adjacent to a metro area) for values 5-9 (the least urban).

the following item-level predictors: the item position on the survey (scaled as a proportion from 0=1st item to 1=last item), the item type (demographic, gatekeeper, outcome [reference], or other [primarily candidate items not used in public reporting]), and an indicator of whether an item was related to Part D (prescription drug) coverage.   Unlike the models of unit nonresponse, response mode, and breakoff, this model was restricted to survey (unit) respondents.

The model of item nonresponse was run on four samples: (1) all unit respondents, (2) all mail respondents, (3) all telephone respondents, and (4) nonbreakoff telephone respondents. Because item missingness in sample three is dominated by breakoff, sample four examines item missingness among nonbreakoff cases.

RESULTS

Unit Nonresponse

Of the 682,088 sampled beneficiaries, 10,808-1.6% were ineligible or missing urbanicity information. Of the remaining 671,280, there were 407,543 unit responses for an overall response rate of 60.7%.   The sample from whom we sought responses in 2008 was 42% male. The median age was 70-74, with 3% under age 45 and 12% over age 84. African Americans constituted 11% of those sampled, with Asians, Hispanics[4], and non-White others 3-4% each. Relatively small sample sizes for beneficiaries 18-24 years of age limit power to distinguish nonresponse in that category relative to the reference category. Subgroups with notably lower than average response rates include beneficiaries under age 45: 26-33% (27-41 %) response rates, Asian beneficiaries: 29% (38%), and Hispanic beneficiaries 33% (43%), p<.001 for each.

In 2008, unexplained variation at the state level was below a standard deviation of 3% . The plan-level standard deviation was 5%, suggesting that plans with similar beneficiaries differ somewhat in response rates: the middle 95% having adjusted response rates of 35-63% (51-71%) (mean +/- 2SD).

Among beneficiary fixed effects, the strongest predictors of unit response were age, race/ethnicity, urbanicity, and eligibility for low income subsidy in 2008. Adjusted response rates in 2008 were highest for beneficiaries 55-79 years of age and lowest for those 18-44 and over 84, with 18-24 having adjusted rates 25% lower than beneficiaries age 65-69. Beneficiaries over 84 had adjusted response rates 9% lower than beneficiaries 65-69 in 2008.

Non-Hispanic Whites responded at adjusted rates 6-18% higher than Asians, African-Americans, Hispanics, and others, p<.001 for each, with the lowest response rates for Asians in each year. Those living in the most urban settings were nearly 8% less likely to respond than those living in the most rural settings. p<.001 for each year. Beneficiaries eligible for low income subsidy were 6% less likely to respond than other beneficiaries, p<.001. Males were somewhat less likely to respond than females: 60% vs. 62%, p<.001.

---

[4] Race/ethnicity is based on CMS administrative variables, which are known to undercount Hispanic ethnicity (Arday et al 2000)

Mode of Response

The results of a logistic regression model predicting mail response 47.0% completed via the mail versus telephone response, 13.7% completed during this follow-up, with an odds ratio greater than 1 representing a tendency to respond earlier by mail rather than later by telephone. Age and race/ethnicity had the strongest impact on mode of response. Consistent with previous findings, older beneficiaries were more likely to respond by mail rather than phone than the 65-74/65-69 year old reference groups, and younger beneficiaries less likely to do so. Compared to non-Hispanic White respondents, Asian respondents were considerably more likely to respond by mail (OR= 2.07), whereas African American (OR= 0.61) and Hispanic (OR=0.67) respondents were less likely. Males were somewhat more likely to respond by mail, while those dually eligible for Medicaid/low income subsidy were more likely to respond by phone (p<.05 for all differences cited in this paragraph).

Breakoffs

In 2008 a total 13,542 complete cases met the definition of a breakoffs: 3.3% of all completes. By mode, 92.4% of breakoffs occurred in the telephone mode, where breakoffs were much more common --13.6% -- than in mail mode-- 0.3% , p<0.0001).

Relative to the reference age of 65-69, respondents who were 25-34 had odds ratios significantly greater than 1 for both complete response and nonresponse. A similar, though less consistently significant pattern, was evident for ages 18-34. Respondents 70 and older had odds ratios for both complete response and nonresponse that were consistently less than one. Taken together, these findings imply that breakoffs are an increasingly common resolution of these cases (relative to both nonresponse and nonbreakoff complete) with increasing age, consistent with Keeter et al (2006). The tendency to breakoff rather than complete was especially pronounced for those 85 and older (OR=0.32). African-Americans had odds ratios significantly less than one for both alternatives to breakoffs, suggesting a stronger tendency than Whites to break off, whereas Hispanics showed the opposite pattern.

Item Nonresponse

In 2008 the average beneficiary was eligible to answer 76.9 items, with a standard deviation of 9.3 and failed to answer 9.2 of them, with a standard deviation of 16.1. The proportion of inappropriate missing responses averaged 11%, with a standard deviation of 19%. The median nonresponse rate was only 3%, with a 75th percentile of 14%, and a 95th percentile of 49%. Moreover, 5% of respondents accounted for 34% of all missing values, 25% of respondents account for 81% of missing values.

Rates of item eligibility varied little by mode. Item missingness was both higher by phone than mail (15% vs. 10%) and more variable (standard deviations of 29% vs. 14%, p<0.0001 for all. Item nonresponse by telephone was dominated by the 14% of complete cases that were breakoffs, which averaged 85% missing. Nonbreakoff phone completes had a lower rate of item missingness -- 4% -- than mail, p<0.0001.

As anticipated, item missingness was much higher for items appearing later in the survey (p<0.001 for all comparisons in this paragraph). Odds for item topics must be interpreted with their survey position in mind, so that the small odds ratio for demographics items (OR=0.28 vs. outcome items), which appear last on the survey, reflects much lower than expected missingness for items appearing so late. Miscellaneous/test items showed the highest rates of item missingness for their position (OR=1.41 vs. outcomes), with gatekeepers having lower missingness than the outcomes for which they sometimes determined eligibility (OR=0.58), and prescription drug subject matter being associated with lower odds of missingness than items on other topics (OR=0.65).

Because item missingness was substantially higher by telephone than by mail, some of these patterns may be confounded by mode and breakoff effects. In addition, there may be different relationships between beneficiary characteristics and missingness within the two modes. The tendency for increasing missingness with age holds strongly in mail mode, but even more so in telephone mode, where those 85 and older are missing at almost three times the rate of those 65-69 in 2008 (OR=2.92). The tendency of Asians to have higher item missingness than Whites was essentially limited to phone mode (and persisted in nonbreakoff cases), whereas the tendency for greater Hispanic than non-Hispanic White missingness was more notable in mail than phone mode. African-Americans had more missing values than non-Hispanic Whites in both modes, but especially in mail mode. Item position and item topic response patterns were fairly similar between mail and nonbreakoff telephone cases.

DISCUSSION

In 2008 response rates were lowest among non-Whites, beneficiaries under 45 or over 84, and low income beneficiaries, especially beneficiaries who were Asian, Hispanic, or under age 45. These patterns were largely consistent with previous findings regarding unit nonresponse for health surveys. Low response rates for Asian beneficiaries may in part reflect the survey being offered only in English and Spanish. In general, the effects of these "risk factors" for low response propensity were less than fully additive.

The second mailing had a dramatic effect, especially on response by mail which was 47% in 2008. The overall response rate by phone was 1% lower in 2008 (14%) but this corresponded to 26% completion by telephone overall in 2008. Consistent with previous findings (e.g. Fowler et al 2002; Zaslavsky et al, 2002), older beneficiaries were more likely to respond by mail rather than phone. Additional follow-up in 2008 seemed to particularly result in more mail completions among older beneficiaries and more telephone completions among younger beneficiaries. Asian respondents were the most likely to respond by mail, whereas African and Hispanic respondents were the least likely to respond by mail.

The tendency to complete some of the survey but breakoff before completion increased with age and was particularly high for African-Americans and low for Hispanics. Telephone breakoffs totaled 14% of all telephone completes. Telephone breakoffs in 2008 were more pronounced in later call-backs than in earlier ones, with one-third of the additional completes achieved through the second half of telephone calls being breakoffs.

7

Item missingness was higher by telephone than mail, but telephone item missingness was dominated by breakoff cases; item missingness in nonbreakoff telephone cases was less than for mail cases. Item missingness was highest among beneficiaries who were older and non-White, consistent with findings by Elliott et al (2005) for recent hospital inpatients. Item missingness was greater for items appearing later in the survey, although demographic items were answered more often than would be expected on the basis of their position at the end of the survey. Older and Asian beneficiaries had particularly high item missingness via telephone. Higher item nonresponse in may in part reflect an older mix of unit respondents in 2008, reflecting larger increases in response rates for beneficiaries 55 and older than for beneficiaries under 55 (given that older age is the strongest predictor of item nonresponse).[5]

CONCLUSIONS

Non-White and older beneficiaries, especially those 85 and older, have the highest rates of unit and item nonresponse, with especially high differences in item nonresponse in telephone mode. To increase overall response rates and representation of disadvantaged groups, CMS could consider tailored follow-up procedures that allow for extra mail follow-up for beneficiaries over 85 and additional telephone follow-up for younger beneficiaries. Telephone protocols may be confusing or difficult for some beneficiaries, especially older beneficiaries. These protocols should be examined and possibly modified to reduce item nonresponse and breakoffs. At the same time, efforts to reduce breakoffs should continue to respect beneficiaries' rights to decline at any point in the survey. Higher breakoff rates for older beneficiaries suggests that protocols that offer pauses or breaks, as well as call-backs, to older respondents may result in higher rates of item completes. Higher breakoffs for African-American beneficiaries suggest future research investigating whether additional efforts to build rapport with African-American respondents might decrease breakoffs. Findings by Grove et al (2004) and Keeter et al (2006) suggest that revisiting pre-notification materials and introductory telephone scripts in a way that makes the topic more compelling and engaging, especially for African-American beneficiaries (perhaps involving focus group work), might decrease breakoffs and increase participation for some key subgroups.

The second mailing in 2008 appeared to yield large gains. The evidence is less clear regarding gains from additional telephone follow-up, though it appeared to be of some help with the youngest beneficiaries, a very low response-rate group. CMS may want to consider limiting telephone follow-up for beneficiaries 55 and older. Additional effort in 2008 yielded far greater gains in MA than in FFS. MA beneficiaries may find the survey more compelling or relevant than FFS beneficiaries. If so, CMS may want to consider changes to introductory FFS survey materials to increase beneficiary interest. Increased survey effort also yielded smaller gains for Asian and Hispanic beneficiaries than for non-Hispanic White and African-American beneficiaries. This may reflect language barriers (especially for Asians), such that some beneficiaries may not respond without additional translations or tailored efforts. To better assess the impacts of additional survey effort, future work should consider within-year comparisons of

---

[5] The mix of unit respondents was very slightly lower in Asians and Hispanics (who tend to have higher-than-average item missingness) in 2008, but this effect on item missingness was much smaller than the age effect.

respondent characteristics, breakoffs, and item missingness by mail wave and telephone call number.

Survey procedures forecast response rates by strata, using estimates based on historical response rates by contract and state in order to optimize sample allocations. Additionally, O'Malley and Zaslavsky (2007) proposed a small area estimation approach to telephone subsampling that seeks to reduce the total number of telephone interviews, costly relative to mail surveys, by borrowing strength within plans from the highly correlated mail estimates. Such an approach might have the additional advantage of reducing the consequences of high item missingness via telephone mode.

Table 1: 2008 Response Rates by Beneficiary Type

	YEAR	MA	FFS with free-standing PDP	FFS-ONLY	TOTAL
Contracts	2008	575[a]	78	N/A	452[b,c]
Benes surveyed	2008	376,948	235,578	69,562	682,088
Eligible	2008	373,334	229,159	68,787	671,280[c]
Completes	2008	242,202	125,764	39,577	407,543[c]
Response Rate	**2008**	**64.9%**	**54.9%**	**57.5%**	**60.7%[c]**

Note: This table excludes cases with missing urbanicity (0.13% in 2008).

a. Includes all beneficiaries surveyed. Among completes only, there were 568 MA contracts in 2008.

b. Total does not equal sum of columns because many contracts enrolled both MA-PD and MA-only beneficiaries.

c. Two contracts (H1111 and H5434) were fielded late and lacked complete administrative information. They were therefore excluded from subsequent analyses. Without these two contracts there were 669,684 eligibles, 406,944 completes, and an overall response rate of 60.8% in 2008.

Table 2: Composition of 2008 Sample and Response Rates by Beneficiary Characteristics

Characteristic	Sampled N	Percent of all sampled who are in this category	Response rate within this category (%)
Male	279,549	41.7	59.5
Female	390,135	58.3	61.7
U.S. Territory	19,882	3.0	59.6
Age			
18-24	663	0.1	27.0
25-34	5,773	0.9	33.0
35-44	15,119	2.4	41.3
45-54	32,611	5.1	50.9
55-64	48,455	7.6	59.7
65-69	*142,558*	*22.5*	*63.7*
70-74	*142,413*	*22.5*	*64.0*
75-79	*97,038*	*15.3*	*64.1*
80-84	*73,617*	*11.6*	*62.4*
85+	75,673	11.9	54.1
Race/Ethnicity			
Asian	16,810	2.5	37.8
African American	75,670	11.3	52.9
Hispanic	25,428	3.8	42.7

Response rates varied by each of these characteristics (chi^2 test, p<.001)

Table 3: Item-nonresponse Rates by Mode and Breakoff Status, 2008

	All Respondents	All Mail Respondents	All Telephone Respondents	Telephone Nonbreak-offs	Telephone breakoffs
N	406,944	314,932	92,012	79,506	12,506
# Eligible Items, of 100 total: **mean (SD)**	76.9 (9.3)	77.4 (9.07)	75.2 (9.7)	73.4 (8.6)	86.6 (8.8)
# non-legitimate missing: **mean (SD)**	9.2 (16.1)	8.1 (11.8)	12.8 (25.4)	3.2 (6.0)	73.8 (14.7)
Proportion of eligible items missing: **mean (SD)**	0.11 (0.19)	0.10 (0.14)	0.15 (0.29)	0.04 (0.08)	0.85 (0.12)