RATING MEDICARE: DOES MODE MATTER?

Celia R. Eicheldinger, Vincent G. Iannacchione, and Lisa C. Carpenter RTI International, Research Triangle Park, NC 27709

Abstract

Telephone respondents to the non-response follow-up portion of the 2000 Consumer Assessment of Health Plans Study (CAHPS®) survey of Medicare Fee-For-Service (MFFS) gave Medicare the highest rating 12 percentage points more often than those who completed a self-administered survey. These apparent mode effects remained significant in regression models even after the inclusion of demographic and health utilization measures. However, an evaluation of 'pure' mode effects could not be made because survey mode had not been randomly assigned prior to the follow-up. In the 2002 CAHPS® MFFS survey, a generalized random-block design (GRBD) experiment was embedded in the non-response follow-up portion of the survey to reduce the possibility of confounding. The results of this embedded experiment validated the 2000 CAHPS® survey findings. In addition to the rating of Medicare, mode effects were found for 12 of the 19 questions comprising various CAHPS® Performance Measures. In each case, the percentage of beneficiaries reporting the most positive response was greater among those answering a telephone survey than those who completed a self-administered survey.

Key Words: Survey Mode Effects, CAHPS®, Medicare Fee-For-Service, Telephone Survey, Social Desirability

Introduction

The Consumer Assessment of Health Plans Study (CAHPS®) survey of Medicare Fee-For-Service (MFFS) beneficiaries is conducted annually by the Centers for Medicare & Medicaid Services (CMS) to provide information to Medicare beneficiaries on the quality of health services provided through Medicare. The CAHPS® methodology traditionally has followed a mixedmode approach to survey administration: two mailings with a telephone follow-up. For the CAHPS®-MFFS, the telephone numbers of the nonrespondents to the mail portion of the survey are traced in preparation for the telephone follow-up. are. however, some non-responding There beneficiaries whose telephone number cannot be traced. These individuals are mailed the CAHPS® MFFS survey in an overnight delivery package.

The CAHPS® instrument was designed to produce data that are comparable regardless of whether they are collected by mail or by telephone. During the development of the original CAHPS® survey instrument, Fowler *et al* (1997) concluded that the mode of data collection had little effect on key CAHPS® results. However, this research did not focus on Medicare beneficiaries.

Pugh *et al* (2002) studied the effects of survey mode (overnight delivery of survey versus telephone administered survey) for the non-response follow-up portion of the 2000 CAHPS®-MFFS survey. Significant mode effects were found for the survey responses to three critical survey questions. These questions were:

- 1. How would you rate your experience with Medicare?
- 2. In general, how would you rate your overall health now?
- 3. In general, how would you rate your overall mental health now?

After adjusting for demographic and health-care utilization measures, the percent of telephone respondents that rated their experience with Medicare the highest rating was 9.1 percentage points higher than those answering the survey delivered in an overnight package. Similar results were found for the percent of Medicare beneficiaries rating their overall health and overall mental health poor/fair.

Random assignment of mode was not possible because beneficiaries whose telephone number could not be found received the survey via overnight delivery survey. Thus, mode was confounded with whether or not the CAHPS® team was able to trace the beneficiary's telephone number. Because of this confounding, Pugh *et al* were not able to conclude that the existence of mode effects alone drove the 9.1 percentage point difference. The 2002 CAHPS®-MFFS team decided an experiment designed to study mode effects was necessary to fully understand mode effects in the non-response follow-up portion of the survey.

Experimental Methods

For obvious reasons, randomly assigning telephone administered surveys to beneficiaries without a traced telephone number is impossible. To work around this challenge, two options were available. The first option was to exclude from the

Tuble If Buyout of filode Effects Experiment							
	In Experiment			Not in Experiment			
Sample	Beneficiaries with Traced Telephone Number			Beneficiaries with Traced Telephone Number	Beneficiaries with Untraced Telephone Number		
Sub-Sample	"Traced	'Traced Telephone'' "Untraced Telephone"		N/A	N/A		
Mode	Phone	Overnight	Phone	Overnight	Phone	Overnight	
Sample Size	n=2200	n=2200	n=2200	n=2200	n=65169	n=4260	

 Table 1. Layout of Mode Effects Experiment

experiment non-responding beneficiaries without a traced telephone number thereby randomly assigning mode of survey (overnight delivery versus telephone administered) to the non-respondents with a traced telephone number. While this option allows mode effects to be studied, this option does not allow generalizations to all beneficiaries in the nonresponse follow-up portion of the survey. The second option, the one chosen, addresses this limitation and is described in the following paragraph.

Please refer to Table 1, above, as the design is described, for a depiction of the experimental set-up. From the pool of non-responding beneficiaries with a traced telephone, a sub-sample of 4400 beneficiaries was created to mimic the demographic characteristics of beneficiaries without a traced telephone number. This sub-sample was referred to as the "untraced telephone" sub-sample (third column in Table 1). A similar process was used for the non-responding beneficiaries with a traced telephone number. Mimicking the characteristics of beneficiaries with a traced telephone number, a second sub-sample titled "traced telephone" (second column in Table 1) was created. After creating these two sub-samples, the mode of survey was randomly assigned to each subsample. The remaining non-responding beneficiaries, those without a traced telephone number, received the overnight delivery survey. The remaining nonresponding beneficiaries (not pictured in Table 1) with a traced telephone number were followed-up by telephone.

The analysis assumes that the "traced telephone" and "untraced telephone" sub-samples were correctly drawn and the groups truly represent beneficiaries with and without a traced telephone number. Although this assumption cannot be formally tested, the results from the overnight delivery survey for beneficiaries with and without a traced telephone number (those not in the experiment) can be compared.

Critical Assumption

The design of this experiment allowed the investigation of mode by sub-sample interaction for response rate, and for, CAHPS® performance measures. This interaction tested the hypothesis that mode effects did not depend on whether or not a

beneficiary had a traceable telephone number. The assumption that the "untraced telephone" sub-sample truly represented beneficiaries without a known telephone number (even though they did have a traced telephone number) was a crucial assumption. Without this assumption, reference to all beneficiaries in the non-response follow-up is not justified; the results are limited to beneficiaries with a traced telephone number.

To evaluate this assumption response rates and responses to the CAHPS® performance measures of beneficiaries with ("untraced telephone" sub-sample) and without (not in the experiment) a traced telephone that received the survey via an overnight delivery package were compared.

The response rates of the "untraced telephone" sub-sample were within three percentage points of the beneficiaries without a traced telephone. For all the CAHPS® performance measures, with the exception of one question, the two groups answered the questions in a similar manner; the critical assumption was supported, thus allowing the interaction of sub-sample by mode of survey to be tested. The CAHPS® survey question, "How would you rate your overall health?" did not support this assumption. However, significant differences due to mode regardless of sub-sample were not found for this question.

Analysis Methods

Logistic regression models were run to investigate mode effects and mode by sub-sample interactions for the response rate, CAHPS® performance measures, self-reported general health, and self-reported mental health. The CAHPS® performance measures included three questions asking beneficiaries to rate their experience with Medicare, their health care, and their personal doctor. The measures also included five composites, made up of two to four separate questions, assessing a beneficiary's experiences with 1) getting needed care, 2) experience with customer services, 3) doctor or other health provider communication, 4) doctor's office staff helpfulness, and 5) getting care quickly.

The outcome variables for the CAHPS® performance measures in the logistic models were dichotomous variables indicating if a beneficiary

reported the most positive response to the survey question. For example, the dichotomous variable asking the beneficiary to rate his or her Medicare would indicate if a respondent answered a "ten" (the highest rating) response. Similarly for the questions comprising the composites the dichotomous variable indicates if the beneficiary answered "Not a Problem" or "Always" (the most positive response). For the self-reported mental and general health questions, the dichotomous variable indicated if a survey respondent reported "excellent" or "very good" health.

Covariates were included in the logistic models to ensure significant mode effects were the result of survey mode and not other characteristics of the beneficiaries. The covariates used in the response rate analysis differed from the covariates used in the CAHPS® Performance Measures analysis. The covariates for the response rate analysis reflected variables that were available for both nonrespondents and respondents. These variables included gender, race, and age and were available through the Medicare Enrollment Database (EDB). The current set of case-mix variables used by the CAHPS® MFFS team were included in the list of covariates used in the CAHPS® Performance Measures analysis. These case-mix variables are education, age, self-perceived mental and general health status and an indicator variable signifying if a beneficiary received proxy help when answering the survey. In addition to the case-mix variables five other variables were added to the logistic models; gender, race, ethnicity, and, two indicator variables: if a beneficiary was enrolled in Medicaid and if the beneficiary was previously enrolled in Managed care.

Segmentation analyses were conducted to identify potential interactions for the CAHPS® performance measures that exhibited significant mode effects. In many cases, mode of survey was the most discriminating variable, but in other cases a covariate was the most discriminating variable. The segmentation analyses were also used to determine two-way interactions that needed to be added to the logistic models. All of the logistic models for the CAHPS® measures with significant mode effects were re-run with the two-way interactions. If the interactions were not significant they were removed from the models; if they were significant, the twoway interactions remained. All of the results presented in the results sections reflect the final logistic models.

Experimental Results Response Rates

The experiment started with 8800 nonresponding beneficiaries. However, some beneficiaries completed the survey before the nonresponse experiment began and those beneficiaries were removed from the experiment. As a result, 3610 non-responding beneficiaries received a survey via overnight delivery (1790 from the "untraced telephone" sub-sample and 1820 from the "traced telephone" sub-sample) and 4096 received a telephone administered survey (2052 from the "untraced telephone" sub-sample and 2044 from the "traced telephone" sub-sample).

The response rates for the experiment are listed in Table 2, below. The response rates for the four different groups vary slightly, from 22.2 percent for the "untraced telephone" overnight delivery group to 27.0 percent for the "traced telephone" and "untraced telephone" phone follow-up survey groups. Regardless of sub-sample, the response rates for the telephone survey were slightly higher; 4.8 percentage point improvement for the "untraced telephone" subsample and 2.0 percentage point improvement for the "traced telephone" sub-sample.

Response Rate¹ Results (SE's)					
Mode	"Untraced Telephone"	"Traced Telephone"	Overall		
Overnight	22.2%	25.1%	23.7%		
Delivery	(0.01)	(0.01)	(0.01)		
Phone	27.0%	27.0%	27.1%		
Follow-up	(0.01)	(0.01)	(0.01)		
Difference	-4.8% (0.01)	-2.0% (0.01)	-3.4% (0.01)		

 Table 2. Comparison of Response Rates by Mode

¹Not adjusted for the covariates in the logistic model.

Logistic regression models were run to determine if significant differences in response rate due to mode of survey existed. The interaction between mode of survey and sub-sample was not found to be significant (p-value=0.15). Mode of survey, however, was significant with p-value = 0.0004. Clearly with a sample size of 7706 there was sufficient power to detect small differences.

CAHPS® Performance Measures

Table 3 lists the CAHPS® Performance Measures and the self-reported health questions. The adjusted percent of beneficiaries rating their Medicare, health care, and personal doctor a "10" (the highest rating), broken down by mode of survey, are also found in Table 3, along with the adjusted percent of beneficiaries responding "Not a Problem" or "Always" (the most positive response) to the questions comprising the composites. A 'NS' in the columns, where the adjusted percents are located, denotes survey questions without significant mode

CAHPS® Measures	Survey Questions	Percentages of Most Positive Responses ² (SE)		
		Overnight Delivery	Telephone	
CAHPS® Ratings	How would you rate your experience with Medicare?		54% (0.02)	
	How would you rate your health care from all doctors and other health providers?		53% (0.02)	
	How would you rate your personal doctor or nurse now?		55% (0.02)	
Health Ratings	In general, how would you rate your overall health now?		31% (0.01)	
Treatti Katings	In general, how would you rate your overall mental health now?	NS	NS	
	How much of a problem, if any, was it to get a personal doctor or nurse you are happy with?		89% (0.01)	
Questions Comprising the	How much of a problem, if any, was it to see a specialist that you needed to see?		NS	
Needed Care Composite	How much of a problem, if any, was it to get the care you or a doctor believed necessary?	NS	NS	
	How much of a problem, if any, were delays in health care while you waited for approval from Medicare?		NS	
Questions Comprising the CAHPS® Customer Service Composite	How much of a problem, if any, did you have with this paperwork for Medicare?	NS	NS	
	How much of a problem, if any, was it to find or understand information in the written materials?	NS	NS	
	How much of a problem, if any, was it to get the help you needed when you called Medicare customer service?	NS	NS	
Questions Comprising the CAHPS® Communication Composite	How often did doctors or other health providers listen carefully to you?	68% (0.02)	80% (0.02)	
	How often did doctors or other health providers explain things in a way you could understand?	63% (0.02)	77% (0.02)	
	How often did doctors or other health providers show respect for what you had to say?	66% (0.02)	80% (0.02)	
	How often did doctors or other health providers spend enough time with you?	54% (0.02)	73% (0.02)	
Questions Comprising the	How often did office staff at a doctor's office or clinic treat you with courtesy and respect?	83% (0.02)	89% (0.01)	
CAHPS® Staff Helpfulness Composite	How often were office staff at a doctor's office or clinic as helpful as you thought they should be?	70% (0.02)	79% (0.02)	
	When you called during regular office hours, how often did you get the help or advice you needed?	67% (0.02)	77% (0.02)	
Questions Comprising the CAHPS® Getting Care Quickly Composite	How often did you get an appointment for regular or routine health care as soon as you wanted?	55% (0.02)	72% (0.02)	
	When you needed care right away for an illness or injury, how often did you get care as soon as you wanted?	NS	NS	
	How often did you wait in the doctor's office or clinic more than 15 minutes past your appointment time to see the person you went to see?	19% (0.02)	35% (0.02)	
² Percentage estimates are the predictive margins of logistic models that included covariates.				

Table 3. Performance Measure Results

effects. The adjusted percents are the margins from the logistic regression analysis. The margins are adjusted for the covariates included in the logistic models.

Mode effects did not exist for the percent of beneficiaries that reported their mental health excellent or very good. Likewise, beneficiaries' responses to the questions comprising the customer service and getting needed care composites, with the exception of one question, did not reveal significant mode effects. The percent of beneficiaries that rated their personal doctor, Medicare, overall health care, and general health perception the highest rating, however, were statistically significant. Questions comprising the doctor communication, staff helpfulness, and getting care quickly composites, with the exception of one question, were all significant.

In all cases of significant mode effects, the telephone administered survey resulted in a higher percent of beneficiaries reporting the most positive response. For instance, the adjusted percent of beneficiaries, among those answering the selfadministered survey arriving via overnight delivery package, rating their Medicare the most positive response was 40 percent compared to 54 percent for those answering a telephone administered survey; mode of survey resulted in a 14 percentage point difference. In the following three sections a more detailed look at the effects of mode on the questions asking beneficiaries to rate their Medicare, health care, and personal doctor, as well the effects of mode on the questions comprising the communication composite are presented.

Rate Medicare

The rate Medicare question was the only outcome exhibiting a significant sub-sample by mode of survey interaction. The mode effects for the "traced telephone" group were much more dramatic than for the "untraced telephone" group (Table 4). The difference in the percent of beneficiaries rating Medicare the highest due to mode for the "traced telephone" group was 18 percent compared to 10 percent in the "untraced telephone" group. The difference, due to mode, for the "traced telephone" group was 8 percentage points lower than the difference, due to mode, for the "traced telephone" group. Thus, for the rate Medicare question, the extent of the mode effects was conditional on whether or not the beneficiary's telephone number was traced.

The segmentation analysis for the rate Medicare question revealed that perceived general health was the most discriminating variable. Among beneficiaries reporting poor/fair health, only 40 percent rated Medicare a ten versus 52 percent among beneficiaries in good/very good/excellent health. Not surprisingly, nested in the perceived general health variable, mode was the most discriminating variable. The interaction, however, of health status and mode was not significant; telephone surveys resulted in a higher percent of beneficiaries reporting Medicare a ten regardless of perceived health status.

Percent ¹ Rating Medicare The Most Positive						
Response (SE)						
Mode	"Untraced Telephone" Group	"Traced Telephone" Group	Overall			
Overnight Delivery	42% (0.02)	38% (0.02)	40% (0.02)			
Phone Follow-up	52% (0.02)	56% (0.02)	54% (0.02)			
Difference	-10% (0.02)	-18% (0.02)	-14% (0.02)			

 Table 4. Rating of Medicare Results

1Adjusted for the covariates in the logistic model.

Rate Health Care and Rate Personal Doctor

These two CAHPS® performance measures, respectively, ask beneficiaries to rate the health care they received from doctors and other health providers and to rate their personal doctor or nurse. The results for both measures revealed significant mode effects. The percent of beneficiaries that rated their health care a ten was 44 percent for the group answering a self-administered survey compared to 53 percent for the group answering a telephone administered survey (see Table 3). Similarly, the percent of beneficiaries rating their personal doctor a ten was six percentage points higher for those answering the telephone survey.

Although mode was significant for both of these ratings questions, mode was not the most discriminating variable. The percent of beneficiaries that rated their health care a ten was more greatly affected by perceived health status. The percent of beneficiaries that rated their personal doctor a ten was more greatly affected by whether or not they had a proxy help them answer the survey questions.

Doctor Communication Composite

Mode effects for the survey questions comprising the doctor communication composite were surprisingly large (see Table 3). The percent of beneficiaries, interviewed by telephone, reporting "Always" when asked if doctors or other health care providers spent enough time with them was 73 percent compared to 54 percent for the overnight delivery survey group. Again, administering a telephone survey versus a self-administered overnight mailed survey resulted in a 19 percentage point increase in the percent of most positive responses. The other questions in the doctor communication composite exhibited a similar pattern with differences ranging from 12 to 14 percentage points. Naturally, according to the segmentation analysis, mode of survey was the most important variable in explaining the percent of beneficiaries that reported "Always" to the questions comprising the communication composite.

Discussion

The results of the embedded experiment for the 2002 non-response follow-up portion of the CAHPS® FFS survey indicate that, compared to those who completed a self-administered mailed questionnaire, respondents who answered by telephone were more likely to report the most positive response to 13 of the 20 questions comprising the CAHPS® performance measures. Regression models were developed to determine whether these apparent mode effects could be explained by either differences in the response propensities of persons selected for the telephone and overnight mail follow-up samples, or by differences in the demographic and health status characteristics of those who responded to the follow-up. The models indicated that mode, although statistically significant, did not substantively affect a person's propensity to respond to the follow-up, and yet mode remained a significant factor in explaining the results for the questions comprising the CAHPS® performance measures even after adjusting for demographic differences between telephone and overnight respondents.

The results of the experiment lead to the speculation that the telephone mode may induce social desirability (i.e., instinctive desire to please someone such as a telephone interviewer) in the follow-up sample estimates. Mode effects between telephone and mail survey responses are frequently observed and can be quite pronounced (DeMaio, 1984; Dillman, Sangster, Tarnai, and Rockwood, 1996). For the CAHPS® MFFS survey, social desirability could occur if a respondent believes that the telephone interviewer represents the Medicare program and that the interviewer holds positive beliefs about Medicare. Then, out of a desire to please the interviewer, the respondent, when asked to rate Medicare, can be expected to answer using the more positive end of the scale.

If social desirability is widespread among Medicare FFS beneficiaries, a dramatic, but superficial, increase in the occurrence of most positive responses to the CAHPS® performance measures could be obtained, for a majority of the items, by converting this predominantly mail survey to a telephone survey. Of course, the extrapolation of social desirability to the Medicare FFS population needs to be tempered by the fact that this experiment was restricted to non-respondents to the mail portion of the survey. Further research is needed to determine whether mode effects extend throughout the Medicare FFS population.

Finally, any evaluation of mode effects needs to take into account the advantages of multiple modes of data collection. This experiment found little difference in the follow-up response rates between telephone and overnight delivery. However, Zaslavsky et al (2003) found that a telephone followup to a mail survey of Medicare Managed Care beneficiaries improved the socio-demographic representation of the respondent sample. These findings suggest a need for balance between the advantages of a multiple mode survey and the disadvantages of mode effects. Such a balance might be attained with the addition of mode as a case-mix variable. Currently, CAHPS® MFFS Performance Measures are adjusted by case-mix variables that include age, education, mental and general health status, and whether or not a proxy answered the survey for the sample member (RTI and RAND 2003). An additional case-mix variable for mode of survey would account for differences in survey responses attributable to mode while allowing for a multi-mode survey. Further research is needed to determine whether this approach can mitigate mode effects without adversely affecting the representation of demographic groups that benefit from a multimode survey.

References

- DeMaio, T.J. (1984). Social desirability and survey measurement: A review. C.F. Turner & E. Martin (Eds.), *Surveying subjective phenomena: Volume* 2 (pp.257-282).
- Dillman, D.A., Sangster, R.L., Tarnai, J., & Rockwood, T. (1996). Understanding differences in people's answers to telephone and mail surveys. M.T. Braverman & J.K. Slater (Eds.), *New directions for evaluation series*, 70 (Advances in survey research), pp. 45-62. San Francisco: Jossey-Bass.
- Zaslavsky AM, Zaborski L, Cleary PD, (2002) "Factors affecting response rates to the Consumer Assessment of Health Plans

Study (CAHPS®) survey." *Medical Care*, 40(6): 485-499.

- Fowler, FJ, Gallagher PM, Nederend S. (1999). "Comparing telephone and mail responses to the CAHPS® survey instrument." *Medical Care.* 37(3): MS41-MS49.
- Pugh, N., V.G. Iannacchione, T.X. Lance, and L.D. Dimitropoulos (2002). "Evaluating Mode Effects in the Medicare CAHPS® Fee-For-Service Survey." Presented (by Pugh, N.) at the Joint Statistical Meetings, New York, NY.
- RTI and RAND (2003). Implementation of Medicare CAHPS® Fee-For-Service Survey Final Report for the 2001 Survey, Prepared for the Centers for Medicare & Medicaid Services, Contract No. 500-95-0061.