

MODE EFFECTS IN A SURVEY OF MEDICARE PROSTATE SURGERY PATIENTS

Floyd J. Fowler, Jr., Anthony M. Roman, Zhu Xiao Di: Center for Survey Research
University of Massachusetts at Boston, 100 Morrissey Boulevard, Boston, MA 02125

KEY WORDS: Mode, Medicare, Survey

I. Background

There have been numerous studies of how the mode of survey data collection affects the answers respondents give. Some studies have found that telephone, personal interview, and mail surveys usually yield comparable results when responding samples are comparable. However, there are some differences that have been attributed to mode.

When differences are found, self-administered responses have been likely to yield more socially undesirable responses than responses to interviewers (Hochstim, 1967; Locander, 1976; Aquilino, 1990; Turner, et al., 1992; Fowler, 1993). When telephone and personal interviews are compared, telephone interviews may yield fewer socially undesirable responses, when differences occur (Henson, 1977; Singer, 1981; de Leeuw, 1988).

There are at least three possible explanations for these differences:

- 1) Respondents are less willing to articulate socially undesirable answers to an interviewer than to write them down.
- 2) Respondents answer questions differently in self-administration (when they can see the whole scales and subsequent questions) than in an oral interview.
- 3) There is more time to reflect and consult with others in a mail survey than in an interview.

A survey of Medicare patients about the effects of prostate surgery provided a chance to evaluate these hypotheses.

In the summer of 1992, a national study was conducted of Medicare recipients who had received surgery for prostate cancer. The study was funded by the Agency for Health Care Policy and Research (AHCP, grants #HS05745 and #HS06336) and conducted by the Center for Survey Research (CSR) at the University of

Massachusetts-Boston in cooperation with the Center for the Evaluative Clinical Sciences at the Dartmouth Medical School.

The primary purpose of the study was to estimate the rates of various surgical outcomes and complications.

The secondary goal was to compare how various modes of data collection could be used on a sample of Medicare recipients. Mail, telephone, and face-to-face interviews were all conducted as part of this study. This paper addresses the issues of mode effects. The issues relating to rates of surgical outcomes are contained in Fowler, et al. (1993).

II. Methods

Using a file containing a 5 percent sample of hospitalizations of Medicare enrollees for the years 1988-1990, patients whose claims records indicated they had undergone a radical prostatectomy were selected. Eliminating Massachusetts patients, 840 patients were randomly selected from the remainder of this file, constrained by having 280 patients in each of the years 1988, 1989 and 1990. The 840 sample cases became the U.S. sample and were randomly split into two groups. The first group received a telephone interview with mail being used for those who did not have a telephone, those who could not be reached by telephone after persistent efforts, and those who requested a mail interview due to difficulties with using the telephone. The second group received a mail questionnaire with the telephone used to attempt to interview all those who did not respond to an initial mailing, a postcard reminder, and a follow-up mailing.

Within Massachusetts, all 405 patients receiving radical prostatectomies between 1988 and 1990 were selected. They were split into two random groups. The first group received the same mail with telephone follow-up procedure as one of the U.S. sample groups. Those in the other Massachusetts group were contacted in their homes for face-to-face interviewing. Initial versions of the survey instrument were pretested in person

with men who had had radical prostate surgery, using extensive debriefing questions to evaluate their understanding of the concepts (Lessler and Tourangeau, 1986; Forsythe and Lessler, 1992). Then it was pretested on the telephone using coding of behaviors of respondents and interviewers to further evaluate question wording and the clarity of terms (Fowler, 1993; Fowler, 1992; Oksenberg, Cannell, Kalton, 1992). The final version of the survey instrument was designed for administration over both the telephone and for mail self-reports. The survey instrument was designed to collect data about post surgical patient experiences with incontinence, impotence, and strictures since surgery. Information on treatments received since surgery and on the recurrence of cancer was also obtained. Operationally, for each sample group, a letter from the Administrator of the Health Care Financing Administration (HCFA) was sent to each sample member explaining that they had been selected for a national study of health care outcomes, and that they could participate if they wished. In each instance, the letter was followed about a week later by the first mailing, telephone call or visit to complete the questionnaire.

III. Results Using the U.S. Sample

When considering the U.S. sample assigned for initial telephone contact, the following sample breakdown was obtained. Of the 420 sample cases, 13 were ineligible for interview due to death or having been erroneously identified as having had surgery for prostate cancer. Of the remaining 407 sample cases, 71 were initially done through the mail due to not being able to find a telephone number, while an additional 23 were transferred to the mail after repeated attempts by telephone failed. In the end, 377 completed questionnaires were obtained for an overall response rate of 92.6%. Of the 377 completed questionnaires, 304 were from initial telephone interviews, while 73 were received by mail in the follow-up procedure. A total of 52 of the 73 mail surveys were with respondents for whom a phone number could not be found. Two surveys were done with respondents who requested a mail questionnaire, while the remaining 19 were done with people who could not be reached by telephone. If a mail

follow-up was not a possibility, the overall response rate for this group would have been 74.7%. Therefore, using mail as a follow-up procedure sharply raised the overall group response rate.

When considering the U.S. sample using a mail questionnaire with telephone follow-up, a similar picture developed. Of the 420 sample cases in this group, 18 were ineligible due to death or not having received radical prostate surgery. Of the remaining 402 eligible sample cases, 67 were eventually attempted by telephone. In the end, 367 completed questionnaires were obtained for an overall group response rate of 91.3%. Of the 367 completed questionnaires, 341 were done by mail and only 26 were done by telephone follow-up. This implies an 84.8% response rate would have been attained if the telephone were not available for follow-up.

For analytic purposes, both sample groups were combined in order to analyze effects due to mode of interview. Since each sample group was clearly dominated by its initial mode, combining them and looking at direct mode effects was not considered to be problematic.

Initially, it was planned to examine mode effects as they were related to the following specific situations:

- A) Whether the existence of contingent questions caused respondents to answer differently because telephone respondents were unaware of the additional questions while self-respondents could see what they were.
- B) Whether respondents answered sensitive questions differently in a self-administered form than when talking to an interviewer.
- C) Whether respondents answered questions concerned with current situations differently in a self-administered form when talking to interviewers.
- D) Whether questions involving the use of response scales were answered differently under self-administration when respondents could see all the categories.

It became apparent that one of these hypotheses could not be investigated. Since questions with contingent questions were almost

always sensitive in nature, it was impossible to separate that effect from that of sensitivity. Meanwhile, sensitive questions took on many forms. Therefore, this paper will address the issues of sensitivity, recency of event, and the use of response scales. Table 1 breaks down the 53 core questions in the questionnaire by these attributes.

Basically, the nature of the observed mode effects is that mail questionnaires tended to have higher reports of problems and negative surgical results if the questions pertained to sensitive topics or to the current time period. Table 2 gives examples of this phenomenon. When asked if they had ever had a problem with incontinence, there were no differences between mail and telephone respondents. When asked if they had a current problem with incontinence, mail respondents had a 10 percentage point higher reporting of the problem. When asked if the prostate surgery reduced their ability to have erections, there were no differences between mail and telephone respondents. When asked if they had had any erections since surgery, a nearly 7 percentage point higher estimate of that problem was obtained from the mail. Examining this further, when asked relatively nonsensitive questions about receiving post-surgical treatments for strictures, sexual functioning or incontinence, no differences were found. There were also no differences in the percent reporting having had a recurrence of cancer. Even when asking a very sensitive question about erections in the more distant period before surgery, there were no mode differences. But when the time frame became current, more feelings of physical discomfort, limitations of activity, problems with frequent urination, and even problems with not feeling calm were found in the mail. The magnitude of the differences ranged from 7 percentage points up to almost 15 percentage points. And the mail always produced more reporting of problems. In fact, there were 53 questions in the questionnaire that asked about post surgical experiences. Of these 53 questions, 25 showed a significant mode effect at the .05 level of significance. Of these 25 significant differences, 24 can be attributed to the fact that mail produced more reporting of problems. The only other

difference came on the question asking how respondents felt now about how their surgery worked out. Mail questionnaires produced more extreme negative and positive answers to this question.

These results hold whether the question is answered by a scale or in a simple YES-NO. As Table 1 demonstrates, 80% of all subjective scale questions that were both sensitive and concerned with a recent time period were significantly different at the .05 level of significance. Comparatively, 71% of all "YES/NO" questions that were sensitive and recent had significant differences. If questions were about a recent time period but not sensitive, 56% of all scale questions and 50% of all "YES/NO" questions produced significant differences. No questions that were concerned about nonrecent events produced significant differences, whether they were sensitive or not or whether they were a scale or a "YES/NO". In instances where significant differences occurred, more problems or negative responses are reported using the mail version of the questionnaire. Results similar to this have been reported by Turner, et al. (1992) when examining the use of illicit drugs using self-administered forms.

IV. Results from Massachusetts Sample

Within Massachusetts, half of the sample received a mail questionnaire with telephone used as a follow-up for nonrespondents, while the other half received face-to-face interviews. One feature of the face-to-face interviewing that was employed was that the most sensitive questions (i.e., those involving urinary or sexual function) were administered by a booklet in which the interviewer would not read the questions, but the respondent would fill out the booklet in the interviewer's presence and simply close it and hand it back to the interviewer upon completion. It was felt that this might help reduce any effects due to the sensitivity of the questions. Table 3 displays selected results from the mail vs. face-to-face comparisons. In the 23 questions which were asked in the self-administered booklet during the face-to-face interview, there were no significant differences found between the booklet response and those obtained through the mail. For the remaining 28 post-surgical questions, which were

asked directly by interviewers during the face-to-face interview, there were 8 significant differences (all at the .05 level of significance). Of those 8 differences, 5 were of the same nature as the telephone comparisons. Namely, the mail responses produced more reporting of problems. Of the remaining 3 significant differences, 2 were caused by face-to-face respondents using more extreme positive and negative responses to a scale. Finally, 1 was due to more reporting of a problem through face-to-face interviewing.

This use of a self-administered booklet during face-to-face interviewing did seem to reduce differences due to the sensitivity of questions. Overall, when mode differences did exist between mail and face-to-face responses, they tended to mirror the effects of mail to telephone comparisons. Still, differences with face-to-face interviews were not as frequent and not as consistently in one direction.

VI. Summary

Overall, the results of this study seem to paint a rather clear picture. If more reporting of problems is considered to be a less biased, more accurate picture of reality, then the mail survey methodology leads to the most accurate estimates. With an almost 85% response rate using the mail alone, the mail methodology produced an even higher initial response rate than the telephone alone (74.7%). Using dual modes of data collection increased response rates to over 90%. When considering that the mail survey was also less expensive, there appears to be no doubt that the mail methodology is the one of choice, for this highly cooperative sample.

When comparing telephone and face-to-face interviewing, face-to-face interviewing is, of course, more expensive. However, if self-administered booklets are used to administer sensitive questions during a face-to-face interview, this methodology may produce more accurate estimates (if more problem reporting is considered more accurate) than telephone interviews for which self-administration is not an option.

The main mode effect differences regarding the recency and sensitivity of questions producing more reports of problems in the mail is a result that should carry over to other surveys. Also, the

fact that a self-administered booklet produced more comparable results to a mail survey should also generalize well.

These results indicate that:

- 1) Respondents are less willing to articulate the presence of current problems or negative states to interviewers than to write them down. Self-administration in a personal interview setting works as well as a mail survey. Overall, sensitivity and recency combined in a single question is a combination which will lead to mode differences about 70-80% of the time. If questions are recent but not sensitive, mode effects appear about 50-55% of the time. Questions about nonrecent events do not produce significant differences.
- 2) The form of the question (i.e. whether it is a scale or a "YES/NO" does not seem to be an issue.

Overall, if a survey of Medicare recipients is planned, it is believed that a mail survey should be strongly considered as an option. The types of questions that are being planned will give you a good indication of whether to expect mode effects or not.

References

- Aquilino, W. S., & Losciuto, L. A. (1990). Effects of interview on self-reported drug use. Public Opinion Quarterly, 54(3), 362-391.
- de Leeuw, E. D., & Van der Zouwen, J. (1988). Data quality in telephone and face to face surveys: A comparative meta-analysis. In R. M. Groves, P. Biemer, L. Lyberg, J. Massey, W. Nicholls, & J. Waksburg (Eds.), Telephone survey methodology. New York: John Wiley and Sons.
- Forsyth, B. H. & Lessler, J. T. (1992). Cognitive laboratory methods: A taxonomy. In P. N. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz, & S. Sudman (Eds.), Measurement errors in surveys (pp. 393-418). New York: John Wiley.
- Fowler, Floyd, J. (1993). Survey Research Methods. Newbury Park, CA: Sage Publications.

- Fowler, F. J. (1992). How unclear terms affect survey data. Public Opinion Quarterly 56 (2), 218-231.
- Fowler, F.J., Roman, A.M., Lu-Yao, G., Wasson, J., Wennberg, J.E., Barry, M.J. (1993). Patient-reported estimates of complications and cancer recurrence following radical prostatectomy: The national experience (1988-1990). Journal of Urology, (forthcoming).
- Henson, R., Roth, A., Cannell, C. (1977). Personal versus telephone interviews: The effect of telephone reinterviews on reporting of psychiatric symptomology. In C. Cannell, & et al (Ed.), Experiments in interviewing techniques: Field experiments in health reporting, 1971-1977. Hyattsville: NCHSR.
- Hochstim, J. (1967). A critical comparison of three strategies of collecting data from households. Journal of the American Statistical Association, 62, 976-989.
- Lessler, J., & Tourangeau, R. (1989, May). Questionnaire Design in the Cognitive Research Laboratory. Vital and Health Statistics (Washington DC), 6(1): US Government Printing Office.
- Locander, W., Sudman, S., & Bradburn, N. (1976). An investigation of interview method, threat and response distortion. Journal of the American Statistical Association, 71, 269-275.
- Oksenberg, Lois, Charles F. Cannell, and Graham Kalton. (1991). "New Strategies for Testing Survey Questions." Journal of Official Statistics, 7:349-365.
- Singer, E. (1981). Telephone interviewing and black box. In Health Survey Research Methods: Third Biennial Conference. Washington, DC: Government Printing Office, 120-127.
- Turner, C., Lessler, J., George, B., Hubbard, M., Watt, M. (1992). Effects of mode of administration and wording on reporting of drug use. In C.F. Turner, J.T. Lessler, J.C. Gfroerer (Eds.) Survey Measurement of Drug Use: Methodological Studies. Washington, D.C.: Government Printing Office. (177-220).

TABLE 1: Results for 53 Core Questions

Question Type	Subjective vs. Factual	Sensitive	Recent	Number of Questions	Number of Significant Differences*
Scale	Subjective	Yes	Yes	10	8
Scale	Subjective	No	Yes	18	10
Scale	Subjective	Yes	No	3	0
Scale	Factual	Yes	Yes	4	1
Yes/No	Factual	Yes	Yes	7	5
Yes/No	Factual	No	Yes	2	1
Yes/No	Factual	Yes	No	6	0
Yes/No	Factual	No	No	2	0
Categorical	Factual	No	No	1	0
				53	25

* Significance at .05 level

Table 2: TELEPHONE VS. MAIL MODE EFFECTS

<u>Variable</u>	<u>Telephone Estimate</u>	<u>Mail Estimate</u>
Percent having ever had problem with incontinence	88.8%	84.7%
Percent having problem with incontinence currently	67.6%	77.6%*
<hr/>		
Percent claiming surgery reduced ability to have erections	92.9%	93.7%
Percent having had <u>no</u> full erections any time since surgery	86.5%	93.4%*
<hr/>		
Percent having had treatment for strictures	19.7%	20.5%
Percent having had treatment to help sexual function	14.6%	15.9%
Percent having surgery to help stop incontinence	7.5%	6.8%
Percent having recurrence of cancer	26.3%	24.8%
Percent having erections before surgery	92.4%	90.4%
<hr/>		
Percent feeling <u>no</u> physical discomfort currently	67.8%	53.0%*
Percent having <u>no</u> limitations on activities currently	66.8%	52.0%*
Percent having <u>no</u> problem with frequent urination currently	52.6%	39.7%*
Percent feeling calm <u>all</u> of the time in past month	23.2%	14.0%*

Table 3: MASSACHUSETTS SAMPLE RESULTS

<u>VARIABLE</u>	<u>Mail Estimate</u>	<u>Face-to-Face Estimate</u>
Number Of Interviews	173	173
Response Rate	86.9%	89.2%
<hr/>		
<u>Questions in Self-Administered Booklet:</u>		
Incontinence Any Time Since Surgery	84.7%	82.7%
Incontinence Currently	78.5%	73.9%
Received Treatment For Strictures	21.6%	14.8%
Percent not Able To Have Full Erections After Surgery	95.0%	88.7%
Percent With Reduced Ability To Have Erections	93.2%	95.1%
<hr/>		
<u>Questions not in Self-Administered Booklet:</u>		
Percent Reporting Very Good Or Excellent Health Now	85.8%	85.4%
Percent Having Recurrence Of Cancer	20.9%	25.9%
Percent Feeling <u>No</u> Physical Discomfort currently	55.5%	62.3%
Percent Having <u>No</u> Limitations Currently	66.9%	73.5%
Percent Having <u>No</u> Problem With Frequent Urination Currently	24.2%	38.3%*
Percent Feeling Calm <u>All</u> The Time In The Past Month	15.8%	13.6%

*significant at .05 level

340 (ASA.tab)