

Cognitive Testing of a Questionnaire for a Dual-mode Survey of Medicare Beneficiaries

Mary Grace Kovar, Consultant, Lisa Lee and Mary Hess, National Opinion Research Center, and Susan Arday, Centers for Medicare and Medicaid Services

Mary Grace Kovar, Suite 500, 1350 Connecticut Ave NW, Washington DC 20036

As part of its mission to improve the quality of care provided to its beneficiaries, the Centers for Medicare and Medicaid Services (CMS) is carefully developing a survey designed to permit the beneficiaries to report on their own health status and care. One might think that the extensive administrative records maintained by CMS would be sufficient, but the claims data have limited utility. They do not measure people's behavior, their perceptions of their own health status, or reasons why they do or do not seek care. There are no claims data for the beneficiaries who are in managed care plans; managed care plans are based on a fixed amount per enrollee regardless of utilization, therefore no individual claims are filed. Finally, there are no claims for services that are received outside the Medicare system or for services that are not covered by Medicare.

Since the CMS is making a transition to a data-driven approach, the data are needed to help CMS's 53 state Peer Review Organizations (PROs) monitor the quality of care received and to target beneficiary populations under the Health Care Quality Improvement Program.

To meet this need CMS is developing the Medicare Quality Monitoring Surveillance Survey (MQMSS), a survey of beneficiaries age 65 and older.

The characteristics of the system are that it is an annual cross-sectional survey with a focus on the community-dwelling elderly that will obtain reliable data for each state. It will provide data on clinical preventive services, outpatient services, fee-for-service and managed care, and care paid for by Medicare and by other sources. Finally it will provide survey data to be linked with claims data

The first plan was that the MQMSS could be a telephone survey. The possibility of a telephone survey was tested in Macon GA on beneficiaries living as of July 1, 1998. Of the 19,425 beneficiaries who were living in the community and whose addresses were not "in care of" someone else, 11,845 records (61 percent) were

matched to telephone numbers. Match rates were lower for black than for white people and lower for people age 85 and older than for younger beneficiaries (Arday et al, 1999). That pretest was completed as a telephone survey but a coverage rate of 61 percent combined with a CASRO response rate of 50 percent is not sufficient for reliable data.

The three lessons learned from that pretest were that there needs to be a mail survey for beneficiaries who can not be contacted by telephone; the survey must incorporate proxy respondents (some people are just too old and sick to respond for themselves); and that interviewing older people takes longer than anticipated so the questionnaire must be short and must be cognitively tested.

Accordingly, CMS let a contract for the second stage of testing. This project consisted of developing a questionnaire that met CMS's needs, cognitively testing it in person, revising the questionnaire, and cognitively testing it by telephone and by mail.

The basis for the questionnaire was the one used for the Behavioral Risk Factor Surveillance System, a national state-owned telephone survey that is co-ordinated by the Centers for Disease Control (CDC). Sections that were obviously irrelevant to older persons were deleted. Ms. Arday searched the CMS files for CMS goals and issues, and for guidance on what Medicare will pay for. These were the main issues that governed the content of the questionnaire. For example, Medicare will pay for a mammogram for women and a flu shot for everyone every year and a pneumonia vaccination once in a life time. The questions had to fit these regulations. The question on the mammogram had to be asked only of women but that question and the one about flu shot had to be phrased "within the year." The question on the pneumonia vaccine could be "ever", or "in your life."

The content was also governed by that of other surveys. CMS wanted questions that had been tested and questions that were used on other surveys so that the MQMSS data

could be validated against data from household-based surveys.

The purpose of the first phase of cognitive testing was to ascertain whether respondents could understand the questions, recall the information needed to answer questions, and select the appropriate answer category. To address those issues we needed a diverse group of respondents with some of the most prevalent chronic conditions such as diabetes and congestive heart disease. Thirty respondents were recruited from urban and suburban senior centers. Half were African-American and half were white; the age range was 65 - 93 years; the majority were female. A trained cognitive interviewer interviewed each respondent for about 45 minutes (including asking for informed consent and explaining the purpose of the study). The interviews were taped and transcribed and the interviewers were debriefed after the interviews had been completed.

The cognitive interviewing we are referring to here is the variety of cognitive interviewing that is applied by survey researchers to study the general cognitive processes that are assumed to be used by survey respondents as they answer questions (Willis, DeMaio, and Harris-Kojetin, 1998). They are intensive interviews requiring respondents to verbalize their thought processes. They are used to identify cognitive problems with survey questions and to help develop solutions to those problems:

There are four basic think-aloud techniques used in cognitive interviews. They include:

Concurrent protocols

Respondents think aloud while answering the survey questions.

Retrospective protocols

Respondents answer the survey questions first, then are asked to think back about how they arrived at the answers.

Follow-up probes

Specific questions are asked of respondents to learn more about answers they gave to the survey items.

Paraphrasing

Respondents demonstrate their understanding of a survey question by repeating it in their own words.

The second, third, and fourth methods were used most often in this study. Two examples showing both the original question and the probe questions are at the end of this paper.

Some of the findings have been reported before (Jobe,

Keller, and Smith, 1996). Many terms, both medical and common, were not well understood. Respondents had trouble with carotid artery stenosis, urinary incontinence, and obese. They over-reported, which also may mean that they did not understand. When asked what impairments or health problems limited their activities, they tended to report any problem or impairment *regardless* of whether that impairment affected their activities. When asked the number of prescribed medicines that they took, some included non-prescription drugs such as aspirin.

They did not understand some of the questions. When asked about primary and secondary insurance, respondents thought they were asking for the same information. That too has been reported before (Center for Survey Research, 1999). When asked about the last visit to a doctor for a routine checkup, they tended to report their last visit whether routine or not and some women did not know whether an annual gynecological examination counted as a routine checkup. (It is likely that some did and others did not - depending on how extensive the examination was).

Although the research group had learned most of these things before, the administrative people and physicians at CMS had not. Thus, the development of the questionnaire to be tested and the first phase of cognitive testing served to increase the education of the research team and to introduce the CMS team to how much older people really do understand if the concepts and terms are used by members of a discipline but are unfamiliar to outsiders.

The questions were sharpened and focused better. The unfamiliar language was dropped if a more common term could be found - and explained if one couldn't. The questionnaire was ready for phase two of the cognitive testing.

Because this was to be a mail and telephone survey we wanted to test both modes of data collection under simulated conditions; to establish completion times; to cognitively test questions that had been revised after the first phase; and to explore the effects of language barriers and respondent characteristics on success in completing the survey. Forty respondents were recruited in the same way as before. They were assigned at random to the mail or telephone version. Again diversity was sought and was achieved. The age range was 65-87 years; there were 17 males and 23 females; there were five Asian/Pacific Islanders, seven African-Americans, and 25 white Americans. There were three Hispanic Americans.

Because the potential respondents had agreed to take part in the study when they were recruited, unit response rates

to a real field test can not be determined. Nevertheless, the item response for the mail questionnaire was 96 percent in contrast with 99 percent for the telephone questionnaire.

As anticipated, respondents completing the mail questionnaire made more skip errors than interviewers asking respondents on the telephone questionnaire. (The questionnaires were all paper and pencil. No CATI was used.). Overall, about eleven percent of the questions had incorrect skip patterns on the mail questionnaire and two percent had incorrect skip patterns on the telephone questionnaires.

There were, even in this very small sample of respondents to a mail questionnaire, clear differences in the number of errors according to the respondent's age and the respondent's education. The mean number of errors for people ages 65-74 years was seven percent; the mean number for people age 85 or older was thirteen percent. Similarly, the mean number of errors for respondents with college degrees was six percent; for those with less than a high school degree was twelve percent.

There was an additional problem with the responses to the mail questionnaire that had not been anticipated. All of the respondents said at the time they were recruited that they could complete a questionnaire in English. The post-completion interview revealed that many of them had required, or at least received, the help of a translator. Others, who were native English speakers, had received help for other reasons. Often it was not clear who the real respondent to the mail questionnaire was. That's an important issue because there is evidence that self-respondents and proxy-respondents will give different answers to the same question.

Conclusions

The process of developing the questionnaire and having it reviewed by people who actually ran the programs was beneficial to both groups. Although we had read the regulations carefully, we made some errors and did not word the questions to catch the exact meaning. Although they thought they were speaking a language understood by patients, they learned that often they were not.

The first phase of cognitive questionnaire testing was also beneficial. We knew much of what we learned there, but we were reminded that medical terms are virtually unknown to many people and that what we think is clear is capable of many interpretations. We were reminded that people know very little about their medical insurance and interpret the questions on

Activities of Daily Living in ways we never anticipated (Keller et al, 1993).

The second phase of the cognitive testing was, we think, unique. Telephone and personal interviews have been tested (Herzog, Rodgers, and Kulka, 1983; Herzog and Rodgers, 1988). Mail questionnaires have been tested (Turner et al, 1997), but as far as we know this is the first time telephone and mail questionnaires have been tested in a randomized experiment. Our results certainly suggest that the telephone mode is superior, but the testing was limited to interpreting and administering the questionnaire - it was not a field test.

A field test shows a different story. Working from a Medicare sample, it is unlikely that we could reach more than half the beneficiaries by phone (the CASRO rate for the Macon study was 50 percent). The Medicare records do not contain telephone numbers and a large proportion of older people do not list their telephones, live with children or other people with telephones in their names, keep telephones in the name of a deceased spouse, or do not have a telephone. Even if the telephone mode is superior in eliciting response and minimizing skip patterns, that will not compensate for a 61 percent coverage rate. An alternate must be used.

Aside from a major pretest, which is planned, we would recommend one more round of cognitive testing. This questionnaire was developed in the traditional way; through interviewers asking questions of respondents. The format of the mail questionnaire was not tested before we used it. It is likely that we could improve the layout of the questions to avoid some of the problems we uncovered. The American Council for the Blind has a flyer on presentations for people with low visual acuity and following those directions might help. Further, we discovered during the re-interviews that some respondents received help with the questionnaire, which means that on the mail questionnaire we really don't know who answered the questions or why they received help. Questions on who the respondent was need to be tested and added.

There is a need for multiple methods for obtaining information from older people. In addition to the lack of telephone numbers on the Medicare files, many older people, especially older men, are hard of hearing and can not hear well over the telephone. But many older people also have trouble with vision and can not read well. The audience is important, especially for health studies, but our means of obtaining information has not kept pace with the need. The number of cognitive studies on older people's ability to respond to questionnaires remains small; we located only five prior studies. If there are more, we'd like to hear about them.

References

Arday, S, P Weiss, M Bedinger, P Minor, D Arday. (1999). Methodology for Conducting a List-based Telesurvey of Elderly Community-dwelling Medicare Beneficiaries. Paper presented but not published.

Center for Survey Research, University of Massachusetts Boston. June, 1999. *Survey Instrument for Medicare Beneficiaries Whose Health Plans Have Terminated: Cognitive Interview Results.*

Herzog, RA, WL Rodgers, RA Kulka. (1983) Interviewing Older Adults: A Comparison of Telephone and Face-to-Face Modalities. *POQ* 47(3): 405-418.

Herzog, RA and WL Rodgers (1988) Interviewing Older Adults: Mode Comparison Using Data from a Face-to-Face Survey and a Telephone Resurvey. *POQ* 52(1): 84-89.

Jobe, J, DM Keller, AF Smith. (1996) Cognitive Techniques in Interviewing Older People. In: *Answering Questions: methodology for determining cognitive communication processes in survey research.* Eds: N Scharz and S. Sudman. Jossey-Bass Publishers, San Francisco.

Keller, DM., MG Kovar, JB Jobe, & LG Branch. (1993) Problems eliciting elders' reports of functional status. *Journal of Aging and Health.* 5(3): 306-318.

Turner, CF, TK Smith, LK Fitterman, T Reilly, K Pate, MB Witt, AM McBean, JT Lessler, and BH Forsyth (1997). "The Quality of Health Data Obtained in a New Survey of Elderly Americans: A Validation Study of the Proposed Medicare Beneficiary Health Status Registry (MBHSR)." *Journal of Gerontology: Social Sciences*, Vol. 52B, No. 1, pp. S49-S58.

Willis, GB, TJ DeMaio, and B Harris-Kovetin (1998) Is the Bandwagon Headed to the Methodological Promised Land? Evaluating the Validity of Cognitive Interviewing Techniques. In: *Cognition and Survey Research*, Eds: Sirken, MG et al. John Wiley & Sons, Inc., New York.

QUESTIONS:

“Are you limited in any way in any activities because of any impairment or health problem?”

- Yes
- No
- Don't Know/Not sure
- Refused

“What is the major impairment or health problem that limits your activities?”

Check all that apply

- Arthritis/rheumatism
- Back or neck problem
- Fractures, bone/joint injury
- Walking problem, etc.
- Depression/anxiety/emotional problem
- Urinary incontinence
- Fecal incontinence
- Hyperlipidemia
- Other impairment/problem
- Don't know/Not sure
- Refused

PROBES:

“Let's review the list of major impairments and health problems. I want to make sure that all the terms on the list are clear. Are there any items on this list that you could not redefine in your own words?”

“You indicated that your activities are limited by (impairment/problem). I'd like to hear more about how (impairment/problem) limits your activities.”

“Are there any other health problems from this list that limit your activities?”

QUESTION:

“How many different kinds of medicines do you take regularly that were prescribed by a doctor?”

- None
- 1 or 2 different kinds
- Between 3 and 5 different kinds
- Between 6 and 8 different kinds
- More than 8 different kinds
- Don't know/Not sure
- Refused

PROBES:

“I would like to hear more about the medicines that you take. What is the name of each medication, what is it for, and when do you take it?”

(Interviewer observes if the enumeration of medicines is in agreement with the answer given to the question. Interviewer probes discrepancies as needed.)

“This question asks about prescription medicines that you take ‘regularly.’” What does ‘regularly’ mean to you? For example, what if you took a medicine:

- every day?
- a few times a week?
- less than once a week?”

“Do you take any non-prescription medicines? If yes, did you include these in your count or did you leave them out?”