

**ASKING QUESTIONS ABOUT DISEASES:
HOW QUESTIONNAIRE DESIGN CHANGES IMPACT THE CHRONIC DISEASE DATA OF
THE MEDICARE CURRENT BENEFICIARY SURVEY**

Sophia Chan
Westat, Rockville, MD 20850

KEY WORDS: CAPI, PAPI, Health Survey, Long-term Care Residents, Minimum Data Set, Reference Period

Background

The Medicare Current Beneficiary Survey (MCBS) is a continuous study of a nationally representative sample of aged and disabled Medicare beneficiaries sponsored by the Centers for Medicare & Medicaid Services (CMS). It provides comprehensive data on health and functional status, health care expenditures, and health insurance for Medicare beneficiaries. Data on sample persons may be collected in the Community Interview, the Facility Interview, or both, depending on whether they reside at home or in a long-term care setting.

The MCBS Facility Interview collects information about institutionalized patients from proxy respondents, most of whom are nurses or other primary care givers. To enhance data quality, proxy respondents are encouraged by the interviewers to refer to the sample person's medical records when answering questions about health status, including chronic diseases. One of the key components of every resident's medical record in a nursing home is the Minimum Data Set (MDS), a set of federally-mandated clinical assessments administered to nursing home residents. Information on the sample person's chronic diseases is recorded in the Disease Diagnosis section in the MDS, in the subsection of Diseases. Forty-three categories of chronic diseases, and the options of "other" and "none of above," were listed in this subsection. The Disease Diagnosis section includes an instruction for nursing home staff to "check only those diseases that have a relationship to current ADL status, cognitive status, mood and behavior status, medical treatments, nursing monitoring, or risk of death." Inactive diagnoses are not recorded.

The Facility Questionnaires underwent significant questionnaire design changes in 1997. In particular, the Health Status section of the Facility Instruments was redesigned to build around the use of the MDS in data collection. To mirror MDS format and wording, the reference period, question format, and wording of the

chronic disease questions, among others in the Health Status section, were revised. At the same time, the mode of administration was changed from paper-and-pencil interview (PAPI) to computer-assisted personal interview (CAPI). As a result, the distributions of some of the items in the Health Status section after the changes were very different from those before the changes. The chronic disease items were especially affected by these changes. This paper discusses the background and nature of the questionnaire design changes to the chronic disease questions, the impact of these changes on the data, and the implications of these changes.

Questionnaire Design Changes

The changes to the chronic disease questions include reference period, question format, types of diseases, and reference to medical records in the question wording.

Reference Period. In the PAPI version (1996 and before), the reference period was not specified in the questions. For instance, the series of items about chronic diseases asked if "a doctor has ever told (SP) that (he/she) had any of these conditions." Thus, this series covered not only the period during which the sample person resides in the current facility, but also his or her entire life. These items assumed that once a disease was diagnosed, then it would always be present after the initial diagnosis. This series was developed in a community or household context. Thus, a problem with administering these questions in a long-term care setting is the assumption that the proxy respondent has knowledge about the medical history of the sample person's entire life, before and after he or she was admitted to the facility. In reality, the staff at long-term care facilities may have only a limited perspective of it.

Changes were made to the reference period of the chronic disease items to mirror that of the MDS in 1997. Because the MDS Disease Diagnosis section instructs long-term care staff to check only those diseases that have a relationship to the sample person's current health status, the chronic disease item in the CAPI version has been reworded to emphasize the current aspect of the diseases of interest. The

reference period of the chronic disease items, as well as that of the Health Status items, is explicitly stated. The CAPI version of the chronic disease questions reads, “What active diseases were checked on {SP’s} MDS assessment}?”. Also, CAPI determines the date of the MDS full assessment forms completed closest to the day of interview and displays on the screen the version and section in the MDS where the proxy respondents can locate the answer to these questions.

Question Format. In the PAPI version, the chronic disease items were a series of questions with three response options: yes, no, and don’t know. In the CAPI version, the format of the chronic disease items mirrors that of the Disease subsection of the MDS. Data about chronic diseases are captured by a select-all-that-apply item. The same 45 options in the Diseases subsection in the MDS are listed in the item as response categories. Interviewers enter all the diseases checked in the MDS into the computer. Unlike the disease item in the MDS, the response categories in the CAPI chronic disease item are listed alphabetically to facilitate search during data collection.

Types of Diseases. Because the new design mirrors the wording of the MDS, the types of chronic diseases asked about in the CAPI version are slightly different from those in the PAPI version:

- The PAPI version contained one item on diabetes, while the CAPI chronic disease item includes two diabetes-related response categories: diabetes mellitus and diabetes retinopathy. Similarly, in the PAPI version, Alzheimer’s disease and dementia were asked about in the same item; in the CAPI version, “Alzheimer’s disease” and “dementia, other than Alzheimer’s” are two separate response categories at the chronic disease item.
- In the PAPI version, emphysema, asthma, and chronic obstructive pulmonary disease (COPD) were asked about in the same item; in the CAPI version, “asthma” and “emphysema/COPD” are two different response categories in the chronic disease question.
- The PAPI version asked two separate items about arthritis: rheumatoid arthritis and “arthritis, other than rheumatoid arthritis.” The CAPI version of the chronic disease item contains “arthritis” as the only response category related to the disease. Similarly, the PAPI version asked about skin cancer and

“any other kind of cancer, malignancy, or tumor” separately. In the CAPI chronic disease item, “cancer” is the only response category that pertains to the disease.

Reference to Medical Records. Another important difference between the PAPI and CAPI versions is that the CAPI version has been reworded to reinforce the use of the MDS, or other medical records if the MDS is not available, during data collection. In the PAPI version, the chronic disease items did not contain any reference to the MDS or any medical records in the question wording. In the CAPI version, on the other hand, the words “MDS assessment” are included in the question wording, which reads, “What active diseases are checked on {SP}’s MDS assessment?” The advantage of mentioning the MDS assessment in the question wording is that by reading the question to the proxy respondents, the interviewers are also telling them to refer to the MDS or other medical records when answering the questions about chronic diseases.

Analysis

As a result of the design changes discussed here, the distributions of several chronic disease items in the CAPI version may be very different from those in the PAPI version. Research on the impact of design changes on health and nutrition survey data has demonstrated that even minor differences in the structure and wording of the questionnaires resulted in major differences in estimates of various health indicators (Kuskowska-Wolk et al., 1992; Picavet and van den Bos, 1996). This paper compares the distributions of the chronic disease variables before and after the implementations of the questionnaire design changes. Data collected in 1996 and 1997 were selected for the analysis. The data of these consecutive years were used to minimize the possibility that any change in health status was a function of time. The 1996 data were collected using PAPI; the 1997 data were collected by CAPI.

To enhance the comparability of the chronic disease data, only full-year facility residents were included in the analysis. Part-year facility residents were excluded from this study because their health status tends to be different from that of full-year facility residents. The 1996 full-year facility sample contained 799 cases; the 1997 full-year facility sample included 827 cases.

To compare the changes in the distributions of chronic disease variables before and after the questionnaire design changes, only the chronic diseases asked about in both the 1996 (PAPI) and 1997 (CAPI) versions were analyzed. These chronic diseases include Alzheimer's disease, arthritis, asthma, cancer, COPD, dementia, diabetes, emphysema, hip fracture, hypertension, osteoporosis, and stroke. As discussed earlier, some of these chronic diseases were asked individually in 1996/PAPI, and with other diseases in 1997/CAPI, or vice versa. To make the results more reader-friendly, the diseases that were asked separately in one version and together in another version were combined into one summary category in the tables. This recoding scheme was also applied to the items of cancer, skin cancer, Alzheimer's disease, dementia, emphysema, asthma, and COPD.

Because supplemental sample cases are introduced every year to the MCBS sample to replace cases who were deceased or retired from the sample, it is necessary to reduce the possibility that any changes in the distributions of the chronic disease variables could be artifacts introduced by incoming sample cases. To do so, the analysis of the chronic disease variables was restricted to the overlapping sample, that is, full-year facility residents who were in the MCBS sample in both 1996 and 1997. The overlapping sample consists of 462 cases.

The 1996 and 1997 full-year facility samples shared similar demographic characteristics. About half of the sample persons were the oldest old (i.e., at least 85 years old [51% in 1996; 47% in 1997]). At the same time, almost one-fifth was under 65. Most of the sample persons were female (69% in 1996; 68% in 1997). The majority of the sample (87%) was White non-Hispanic. The majority of the sample persons had an annual income under \$10,000 (67% in 1996; 66% in 1997). The demographic profile of the cases who were in the Facility sample in both 1996 and 1997 (i.e., the overlapping sample) was basically similar to those of the full-year samples, except that the overlapping sample consisted of a higher proportion of cases under 65 and those at least 85.

Results

Between 1996 and 1997, there was a decrease in the percentage of cases with various types of chronic diseases. The decrease ranged from 2 to 21 percent (Table 1). In particular, the

variable of "hip fracture" saw a 21 percent drop, and stroke 17 percent.

In terms of cancer, there was a 16 percent decrease in 1997. In 1996, when skin cancer and "any other type of cancer, malignancy, or tumor" are asked separately, 3 percent of the sample persons were reported to have skin cancer, while 17 percent had other types of cancer. Altogether, 20 percent suffered from either skin cancer or another type of cancer. In 1997, when "cancer" was the only response category in the chronic disease item, only 4 percent were reported to have cancer.

In terms of arthritis, there was a 12 percent decrease in 1997. In 1996, rheumatoid arthritis and "other types of arthritis" were asked separately. Five percent of the sampled persons were reported to have rheumatoid arthritis, 29 percent suffered from other types of arthritis, and 2 percent had both. In sum, 32 percent of the sampled persons suffered from rheumatoid arthritis, arthritis other than rheumatoid arthritis, or both. In 1997, when "arthritis" was the only response category in the chronic disease question that pertains to the condition, 20 percent were reported to have arthritis.

Other chronic disease variables showed smaller decrease in percentage between 1996 and 1997. For example, there was a 4 percent decrease in the percentage of cases with emphysema, asthma, or COPD. In 1996, when emphysema, asthma, and COPD were asked about in the same item, 13 percent of the sample persons were reported to suffer from at least one of these pulmonary diseases. In 1997, emphysema and COPD were given as one response category, while asthma was used as a separate response category. Eight percent of the cases had either emphysema or COPD, while 2 percent were reported to have asthma. In general, 9 percent of the sampled persons were afflicted by at least one of these three types of pulmonary diseases in 1997.

The percentage of cases suffering from Alzheimer's disease or other types of dementia had a 3 percent drop in 1997. In 1996, "Alzheimer's disease" and "dementia" were asked in the same item. Forty-five percent of the sample persons were reported to have either Alzheimer's disease or dementia. In 1997, "Alzheimer's disease" and "dementia" became separate response categories in the chronic disease item. Fifteen percent were reported to suffer from Alzheimer's disease, 29 percent dementia. In total, 42 percent of the sampled

persons were reported to have either Alzheimer's disease or dementia in 1997.

Table 1. Percentages of MCBS Full-Year Facility Cases with Selected Chronic diseases/Conditions

Chronic diseases	1996 (N=799)	1997 (N=827)	Diff.
Alzheimer's disease / Dementia	45%	42%	-3%
Arthritis	32	20	-12
Cancer	20	4	-16
Diabetes	17	15	-2
Emphysema / Asthma / COPD	13	9	-4
Hip Fracture	25	4	-21
Hypertension	40	33	-7
Osteoporosis	15	11	-4
Stroke	32	15	-17

The pattern of change in the distribution of the chronic disease variables among the general sample also holds up among the overlapping sample (Table 2). In particular, the size of change between 1996 and 1997 was almost identical for most of the chronic disease variables.

Table 2. Percentages of Full-Year Facility Cases in the MCBS Sample in Both 1996 and 1997 (N=462), with Selected Chronic Diseases/Conditions

Chronic diseases	1996	1997	Diff.
Alzheimer's disease / Dementia	44%	42%	-2%
Arthritis	32	21	-11
Cancer	17	3	-14
Diabetes	18	15	-3
Emphysema / Asthma / COPD	12	9	-3
Hip Fracture	23	3	-20
Hypertension	39	31	-8
Osteoporosis	15	13	-2
Stroke	30	15	-15

Another interesting finding is that some chronic diseases show a larger effect than others. For instance, hip fracture shows a change of 21 percent, while Alzheimer's disease shows a change of only 3 percent, and diabetes only 2 percent. A plausible explanation is that some diseases or conditions, like certain types of cancer or a hip fracture, are recoverable. Once these diseases or conditions are resolved, there is no immediate cause for concern, thus they would not be counted as "active" diseases anymore. Other diseases, like diabetes, Alzheimer's disease/dementia, or osteoporosis, have no cure. People suffering from these diseases or conditions often have to undergo special treatments or diets on an ongoing basis. Thus, these diseases will always be regarded as active diseases. This might explain why some diseases show more percentage changes than the others.

Implications

In the PAPI version of the MCBS Facility Questionnaire, the chronic disease questions used the wording of "Has a doctor ever told {SP}..." This wording has been used by

many national health surveys during the past 30 years and is still being used in the MCBS Community Interview. In 1997, the chronic disease questions were revised to mirror the MDS in order to facilitate data collection and enhance data quality. Often, questionnaire designers have a tendency to stick to old items, even when they realize that some of these items may not be working very well. It takes a lot of courage on the part of a questionnaire designer to make changes to questions in midstream of the survey to improve data quality. Making such changes might impact longitudinal data, but what's lacking in continuity will be compensated by the timeliness and overall quality of the data.

For future research, questionnaire designers should continue to improve the wording of the chronic disease questions and collect data more closely related to the current health status and functioning of the sample persons.

Because the questionnaire design changes discussed in this paper were implemented at the same time as the mode conversion from PAPI to CAPI, it is not possible to analyze the main effects of each of these changes or their interaction effects on the chronic disease data. This is a limitation of this study. Further research using experimental design would be needed to help researchers understand the dynamics of various types of questionnaire design changes on health data.

The author would like to express her gratitude to Brad Edwards and Tom Harper for their helpful comments. Any shortcomings of this article belong to the author.

Bibliography

Kuskowska-Wolk, A., S. Holte, E.M. Ohlander, A. Bruce, L. Holmberg, H.O. Adami, and R. Bergstrom. (Dec. 1992). "Effects of Different Designs and Extension of a Food Frequency Questionnaire on Response Rate, Completeness of Data and Food Frequency Responses." *International Journal of Epidemiology*. 21(6), 1144-1150.

Picavet, H.S., and G.A. van den Bos. (Feb. 1996). "Comparing Survey Data on Functional Disability: The Impact of some Methodological Differences." *Journal of Epidemiology and Community Health*. 50(1), 86-93.