MEDICAL EVENTS: ENHANCING SELF-REPORT QUALITY IN A LONGITUDINAL CAPI DESIGN

Brad Edwards, Westat
Brad Edwards, Westat, 1650 Research Blvd., Rockville, MD 20850

Key Words: Medicare, Memory Recall, Record Matching, Survey Methods

INTRODUCTION

In medical expenditure surveys, underreporting utilization of health care services is a fundamental methodological issue. Primary concern has focused on memory decay. Neter and Waksberg (1964) observed that the longer the recall period, the more likely medical events were forgotten or reported inaccurately. A number of researchers have noted that some events are more easily forgotten than others. The approximate dates and other data regarding an inpatient hospital stay may be remembered for years, but similar details for a routine doctor visit might be remembered for only several months. Some methodologists have suggested that event frequencies beyond a certain threshold may no longer be retained in episodic memory; rather, a generic or form-type memory may be created that is a summary or estimate of many discrete events (Mathiowetz and Duncan, 1988; Burton and Blair, 1991).

Most large-scale medical expenditure studies follow the general model established by the first National Medical Care and Expenditure Survey (NMCES) in 1977. Events are defined as the provision of specific services in specific settings on specific dates. Survey respondents are asked to report all medical events received within a certain time period (Kasper, 1980; Edwards and Berlin, 1989; Cohen and others, 1996). The time period is usually several months long, going back to the date of the previous interview. A longer time period is thought to produce inaccurate reports of certain types of events (Cohen and Burt, 1985). The survey reports are aggregated and weighted to produce national estimates of medical events by service type. The Medicare Current Beneficiary Survey (MCBS) was the first to incorporate this model in a computer-assisted personal interviewing (CAPI) design (Adler, 1994).

Medical expenditure surveys have used a variety of features to aid respondent recall. For example, a calendar or diary may help respondents keep track of medical appointments and can be brought into subsequent interviews as a data source. A summary of events reported in the last interview may prompt recall of events that occurred since the last interview. Prescription medicine bottles, checkbook registers, and statements from insurers have also been useful tools to aid recall of utilization detail.

Some studies use alternative sources to enhance the survey report, to help correct for underreporting, or to validate self-reports. For the MCBS fee-for-service population, survey-reported events are matched to Medicare claims. This process adds utilization and cost data, beyond what the survey respondent can report about the event. It is also a mechanism for bringing into the MCBS database claims for events that were not reported by the survey respondent, but for which a provider filed a claim. The assumption is that virtually all of these "claims-only" events are valid and represent events that should have matched a survey report, if respondents were able to report with 100 percent accuracy and if project staff were able to match all survey reports and claims (Eppig and Chulis, 1997). One of the most striking features of the MCBS data is the degree to which the claims bolster the survey reports.

During the past decade, CAPI has provided designers with additional tools for aiding recall. This paper examines five CAPI design features on MCBS that can help increase reporting of medical events: interrupting (using a menu that allows the interviewer to review data or return to a previous questionnaire section); automatic jumping (adding a medical event the respondent failed to report in proper sequence, in the Utilization Section); making a comment (indicating some event that should be added by home office staff); dependent interviewing (correcting utilization reports from the previous interview); and dynamic rostering (selecting events from context-specific, user-created lists to link cost data from paper statements to events). MCBS field staff have used these features on more than 240,000 interviews in the past 7 years.

DATA SOURCES

Data on the use of CAPI features are collected as variables in the CAPI database that Westat delivers to Health Care Financing Administration (HCFA). The most recent MCBS Cost and Use Public Use File contains events for the calendar year 1995 and indicates the event type and source (survey report only, survey report matched to HCFA administrative claims, and claim-only events). For this study a subset of the 1995 public use file was created, containing only those

survey respondents who lived in the community the entire reporting year, who were not enrolled in any managed care plan during the year, and who reported data for each of the 1995 rounds of interviewing. There are 7,806 individuals who meet these criteria. The examination was confined to inpatient hospital stays, hospital outpatient department visits, and office-based medical provider visits. These services can be considered more traditional event types than prescribed medicines, home health services, and other medical expenses: Each inpatient, outpatient, and medical provider event has a clear date or period of service and can be attributed to an identifiable medical provider.

Reporting round and source, plus some additional methods data about the interview were obtained from the survey files Westat delivered to HCFA. Event report frequencies were reviewed by feature and reporting period, in the context of results from matching self-reported fee-for-service events to Medicare claims.

Some other aspects of report quality were also examined: whether a calendar was used in the interview as an aid to memory and whether statements were used at all. These variables were created by aggregating interviewer reports across the five rounds that contributed data to the 1995 year. If the interviewer reported the respondent used a calendar in more than one round, the calendar variable was coded "Yes." (Both use of the MCBS project calendar and use of a personal calendar were counted as calendar use.) Similarly, if statements were used in more than one interview, the "used statements" variable was coded "Yes." Statements can mean Explanation of Medicare Benefit forms or documentation from private insurers.

RESULTS

For 1995, there were nearly 150,000 inpatient, medical provider, and outpatient events in the database (Table A).

Table A. 1995 events by source

	In-patient	Medical Provider	Out-patient	Total
Survey only	200 (7.3%)	26,982 (23.8%)	6,352 (18.9%)	33,534 (22.4%)
Survey + Claims	2,049 (75.3%)	54,439 (48.1%)	15,317 (45.6%)	71,805 (48.0%)
Claims only	472 (17.4%)	31,841 (28.1%)	11,934 (35.5%)	44,247 (29.6%)
Total	2,721 (100%)	113,262 (100%)	33,603 (100%)	149,586 (100%)

More than three-quarters of the reported events were medical provider visits. Nearly 30 percent of the 150,000 were "claim-only" events; this is close to the 32 percent Eppig and Chulis (1997) found across all event types in the 1992 MCBS data. However, the proportion of claim-only events varies considerably by type: 17 percent of the inpatient, 28 percent of the medical provider, and 36 percent of the outpatient visits were found only in the claims. If one considers only the events reported in the survey, 68 percent of the reported inpatient, medical provider, and outpatient events matched to administrative claims. The match rate was much higher for inpatient stays (91%) than for medical provider events (67%) or outpatient events (59%). Both the match rates and the claim-only rates support the theory than inpatient stays are more easily remembered and can be reported accurately for a longer period of time, than other event types. Outpatient events can be difficult to remember; repeat visits are common, and episodes of care may be associated with several visits to the outpatient department, when what is remembered most clearly is the episode, rather than each individual visit. ("Repeat" visits are visits to the same provider for the same condition; episodic memories of these events are thought to blur more readily than memories of other visit types.)

There are a number of reasons why a reported event might not match a claim. Medicare does not cover all events, some events have total payments equal to zero, some beneficiaries are not covered by both Part A and Part B, some events occurred in Veterans Administration facilities (which do not generate Medicare claims), etc. (Eppig and Chulis, 1997). However, match rates can be useful measures of report quality, especially when confined to event types that are similar.

By the same token, the claim-only events are not exactly comparable to the survey reported events. They may include some detailed breakouts of services, with each provider service counted as an event in its own right, when the survey respondents might have considered the various services as a single event. Nonetheless, poor reporters would be expected to have a higher proportion of claim-only events.

Neither the frequency of reported events nor the match rate deteriorates noticeably by panel or round. (Data not shown.) This is contrary to expectations; one might suppose that once respondents have experienced the full scope of the data collection request over the course of the first three interviews, some might seek to minimize the cognitive and time burden by suppressing events or details about the events. This effect might be most pronounced for event types that are hardest to

remember over a 4-month period, such as prescribed medicine purchases, but at least for inpatient stays, outpatient visits, and doctor visits, no relationship between event frequency and time in sample was observed.

Turning to the source of the 150,000 events, about 60 percent of the events were reported in the Utilization Section, the traditional way of asking about utilization. The proportion reported in response to these initial questions was greatest for inpatient (77%), followed by medical provider (62%), then outpatient (51%). This suggests the utilization questions work best for salient events recorded in episodic memory and is consistent with the literature on recall.

About 30 percent of the events were only found in the claims. MCBS was designed to give the administrative claims a central position in the survey database; this approach was conceived to minimize respondent burden and recognized that HCFA already had a firm grip on what Medicare paid for covered services. Still, it is surprising to see that so many events are apparently not reported by the survey respondent.

Interrupting. The "Interrupt" feature (a menu that allows the interviewer to go back to a previous section of the interview and review data or make changes) accounted for very few additional events in the database. Just over 1 percent of the 150,000 events were entered through Interrupt, about one in every four interviews. It is worth noting that interviewer use of Interrupt was somewhat constrained in the CAPI application. During the Charge Series (which follows the Utilization Section), Interrupt is disabled because another, more structured mechanism is available in the program for adding events in that section of the interview. A discussion of this other mechanism, "automatic jumping," follows.

The Interrupt menu was designed specifically for MCBS, but the capability was in some respects similar to what is sometimes known as "Jump back" in other systems (including Westat's Cheshire system and Blaise). Jump back is a system feature that allows the interviewer to zoom rapidly back to a previous item without going screen by screen. The MCBS data available for this study do not permit an examination of global use of the Interrupt feature, and it is not known how often interviewers used Interrupt to review events that had been entered previously or to correct spelling, dates, and other details about providers and services. But the relatively low rate of Interrupt usage for *adding* events is consistent with reports on the frequency of Jumpback use in several CAPI applications, at Westat and elsewhere (Sperry and others, 1998; Couper, 1997; Wojcik and Baker, 1992).

Automatic Jumping. After the Utilization Section is completed, the interviewer abstracts data from Medicare and private insurance statements. If an event is discovered that does not match any event already entered in the Utilization Series, the interviewer can add the event. The CAPI program serves up the appropriate questions to collect details about the event, just as if the event had been reported in its proper sequence. In other words, there is an "automatic jump" back to the Utilization Section for the additional event. Once details have been collected about the event (date, provider name, condition, etc.), the program returns the interviewer to the screen in the Charge Series where the event was discovered.

Overall, about 8 percent of the events in the 1995 file were reported in this way. Of course, most of these events matched to claims, since they were discovered in the course of reviewing claim documentation. Without this CAPI design feature, which allowed the interviewer to add an event out of sequence, it is likely that a substantial number of events would have been missed (i.e., been unreported by the survey respondent). Since most of the events entered through this automatic jumping process were Medicare covered, they would have been present in the claims. But the CAPI feature allows the interviewer to prompt the respondent directly for details about the event (including out-of-pocket payments and reimbursements by private insurance sources), information that would have been missing from the claim. Oksenberg and others (1996) have noted the importance of flexible structures for collecting data about medical events and expenditures. It seems that this feature in the MCBS provides interviewers with an effective tool for entering events reported outside their "normal" sequence.

The proportion of events entered in the statement series was highest for outpatient (10%) and lowest for inpatient (4%). The relatively low use of this feature for reporting hospital inpatient stays supports the notion that hospital stays are easily remembered as distinct episodes in memory for a considerable time, and thus one would expect a higher proportion of them to be reported initially in the Utilization Section, compared to other event types.

Making Comments. Comments are another way for interviewers to add events, albeit not directly. When the interviewer encounters an uncodable response or some other problem that cannot be easily resolved on the spot, the interviewer enters a comment in an open text field. The comment is stamped with the

point of entry, and there is no limit on comment length. Data preparation staff in the home office read each comment and determine whether some corrective database action is warranted. Although Comments are a standard feature in most computer-assisted survey information collection (CASIC) systems, reviewing comments is not a commonplace practice. However, on a longitudinal CAPI study that features extensive dependent interviewing, there is a compelling need to review comments. Anything that might affect the data that travel out with the case when it is fielded for the next round of interviewing needs attention. Failure to review comments could undermine confidence in the quality of the overall effort, for both the interviewer and the respondent.

In the 1995 data, about 20 percent of the cases contained comments. About one-third of the cases with comments required some updating in the MCBS database, but most of these updates did not require an event to be added. The Comments feature was used to add just over 1,000 events, or less than 1 percent of all 1995 events in the Public Use File. Most of these were medical provider or outpatient events. Again, this is consistent with the sense that hospital inpatient events are rarely forgotten, especially in a 4-month recall period.

Dependent Interviewing and the Summary Review. The MCBS design employs dependent interviewing to an unusual degree. In the subsequent interview, a printed summary of events is presented to the respondent, "in the form of an annotated calendar, with icons representing doctor visits, hospital stays," and so on (Edwards and others, 1992). The purpose of the summary review is primarily to bound recall (to avoid double counting of events) and to aid memory (by reminding the respondent of providers seen and reported in the previous interview). Interviewers and respondents were not encouraged to review the summary in detail, but if the respondent noticed that an event was missed or entered incorrectly, the interviewer was trained to add or delete it. This process has been called "passive dependent interviewing."

During the summary review of utilization in the following round, some 341 events were discovered and added, less than .3 percent of all reported events. Given the design and objectives of the summary, this result is not surprising. Outpatient visits were added more frequently than the other 2 event types, perhaps reflecting the difficulty many respondents have in distinguishing this event type from others and the relatively high frequency of repeat visits to outpatient departments. The extremely infrequent use of the summary to add events in 1995 is generally consistent with the level reported in a first look at the MCBS data: Edwards and others (1992) reported .7 percent of all reported events were added in the summary, and about 1 percent of events reported previously were deleted.

So overall, the Utilization Section accounts for the majority of events. Self-reports in utilization combined with the claim-only events represent about 90 percent of all events in the 1995 Cost and Use File. However, there is substantial variation by event type, and the four CAPI features—especially the automatic jumping from charge data entry—are the source of the remaining 10 percent.

Dynamic Rostering, Calendars, and Statements. The innovative use of statements in MCBS could be seen as another CAPI design feature. It works by taking full advantage of multilevel, dynamic rostering, a capability of advanced CAPI systems. Multilevel rostering allows the interviewer to enter event detail efficiently and with much greater accuracy than flat, rectangular designs (such as those found in a paper questionnaire). For instance, entry of the first event initiates two lists, one for providers and one for dates of visits to that provider. A second event to the same provider can be entered by selecting the provider name from a case-specific list. This is simpler, quicker, and less error prone than typing the provider's name again. In the Charge Series, the event can be linked with charge and payment data by selecting the provider and the date, and can be linked with insurance coverage information by selecting third-party payers from another case-specific roster.

Before CAPI, calendars were perhaps the most common memory aid in medical expenditure surveys. About 51 percent of the MCBS respondents used a calendar in at least two interviews for reporting 1995 data. Figure A shows events per person by source (survey only, survey and claim, claim only) and by calendar and statement usage. The mean number of unmatched events reported in the survey did not vary by calendar and statement use, except for those who used both statements and calendars—they reported 3.9 events in 1995, compared to 4.4 to 4.6 for the other groups. Those who used neither a calendar nor statements had the highest proportion of claim-only events (43%); the other group with no statements was also higher than average, at 38 percent. The two groups that used statements have the highest proportion of events added through the statement series.

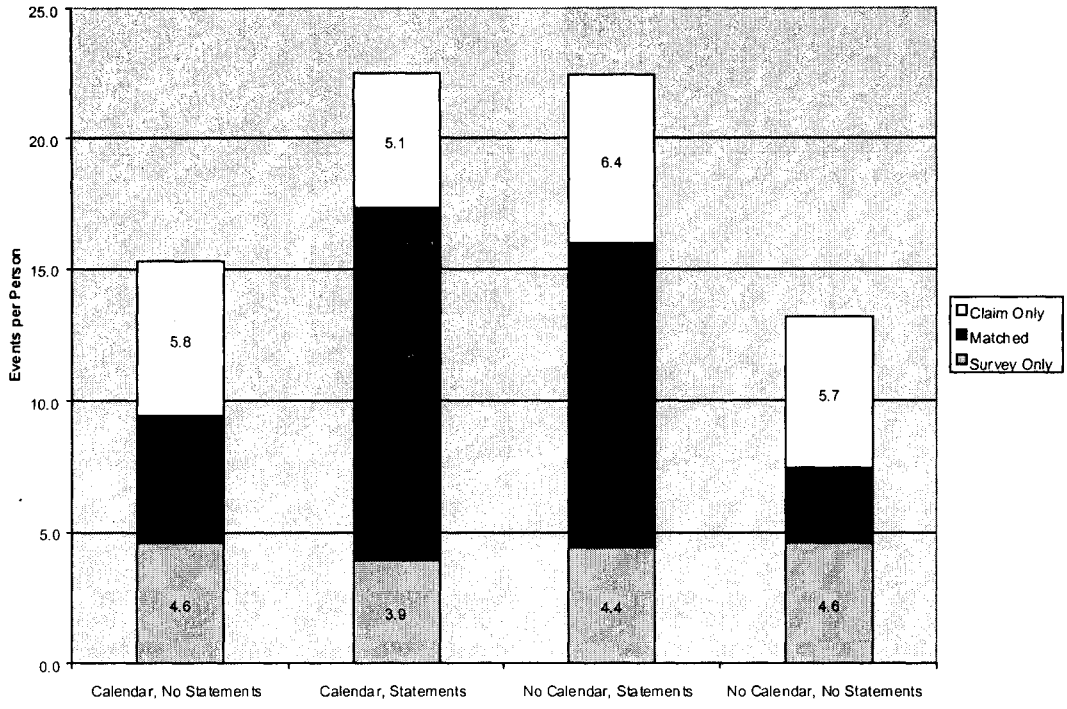

Figure A. Calendars and statements

There is striking variation among these groups in terms of the mean number of matched events. Those who used no calendars nor statements had a mean of only 2.9 matched events. Calendar users without statements had a mean of 4.9. But those with statements reported 13.5 matched events (if a calendar was also used), or 11.6 matched events (if no calendar was used). The "best reporters"—those with both calendars and statements—reported a total of 17.4 events per person, compared to 16.0 for those with statements but no calendars, 9.5 events for those with calendars but no statements, and 7.5 for those with no statements and no calendars.

DISCUSSION

The MCBS data for 1995 provide an opportunity to reflect on the relative contributions of various design features to health care utilization data quality and on the cost-effectiveness of the effort required to create and maintain the features. It is clear that reliance on self-report alone would be insufficient to create estimates of utilization. Even data from the some of the best respondents does not match administrative data 100 percent of the time. The MCBS database includes virtually no imputed utilization for respondents, but it relies heavily upon administrative claims. The "claim-only" events add considerably to the accuracy of the MCBS estimates. The survey design used the administrative claims as a foundation for the self-reports. Although this required a great deal of custom planning, testing, and programming in 1990 to put in place, the investment has served the project well.

The CAPI features discussed here represented another investment early in this decade. The design decision to decouple the collection of utilization data from the collection of charge and payment data was difficult. It risked making it harder for respondents to report all the data on simple events (e.g., a visit to the doctor for a broken toe, with an office charge of $50; the respondent paid $10, and Medicare paid the rest). But for a complex series of events, the ability to easily link several events with one charge, and several payments with one event, was essential. These many-to-many relationships were best served by a design with multilevel rosters and a flexible scheme for adding events and charge data, without regard to a predetermined sequence of how events should be reported. The 1995 MCBS data confirm the value of this approach.

The Comments feature entailed no design costs, but incurs considerable maintenance costs, in the sense that a number of data management staff are required to review comments in time to field the cases in the next round. If the sole purpose of Comments was to add events, it would be hard to justify this ongoing cost. But when viewed as a critical form of communication between the interviewer and the home office, the data are enriched in many ways by keeping up with the

comments review. Adding a few events in the process is a useful by-product.

Two of the other design features discussed here (Interrupt and the Summary Review) were implemented with little or no empirical evidence for their utility in a CAPI study (though summaries played a central role in the early NMCES studies, conducted with paper questionnaires). They required a great deal of effort to develop and maintain, and they add only a limited number of events. Although both were designed primarily to accomplish other purposes, the "event addition" capability that both of them carry accounted for a large part of their design cost. In any future MCBS redesign, a case could be made for carefully reviewing the objectives and usability of these features and refining their functions.

The analysis of the use of calendars and statements generally confirms the sense that memory aids and CAPI features play important roles in improving report quality. But the analysis also suggests that the ability to incorporate events from a variety of sources (memory, documentation) in flexible ways, which can only be supported in a CAPI environment, has a more powerful effect on utilization report quality than the use of a calendar or diary.

REFERENCES

Adler, G.S. (1994). A profile of the Medicare Current Beneficiary Survey. *Health Care Financing Review,* 16(4).

Burton, S., and Blair, E. (1991). Task conditions, response formulation processes, and response accuracy for behavioral frequency questions in surveys. *Public Opinion Quarterly,* 55.

Cohen, S.B., and Burt, V.L. (1985). Data collection frequency effect in the National Medical Care Expenditure Survey. *Journal of Economic and Social Measurement,* 13.

Cohen, J.W., Monheit, A.C., Beauregard, K.M., Cohen, S.B., Lefkowitz, D.C., Potter, D.E.B., Sommers, J., Taylor, A.K., and Arnett, R.H. III (1996, Winter). The Medical Expenditure Panel Survey: A national health information resource. *Inquiry,* 33.

Couper, M. (1997, December). *Interviewer CAPI Behavior on the National Health Interview Survey.* Presented at the Washington Statistical Society.

Edwards, W.S., and Berlin, M. (1989). Questionnaires and data collection methods for the Household Survey and the Survey of American Indians and Alaska Natives (DHHS Publication No. (PHS) 89-3450). *National Medical Expenditure Survey Methods 2,* National Center for Health Services Research and Health Care Technology Assessment. Rockville, MD: Public Health Service.

Edwards, W.S., Sperry, S., and Edwards, B. (1992). Using CAPI in a longitudinal survey: A report from the Medicare Current Beneficiary Survey. *Proceedings of Statistics Canada Symposium 92: Design and Analysis of Longitudinal Surveys,* Statistics Canada.

Eppig, F.J., and Chulis, G.S. (1997). Matching MCBS and Medicare data: The best of both worlds. *Health Care Financing Review,* 18(3).

Kasper, J.D. (1980, September). Perspectives on health care: United States, 1980. National Medical Care Utilization and Expenditure Survey. *Series B, Descriptive Report No. 14.* Office of Research and Demonstrations, Health Care Financing Administration. Washington, DC.

Mathiowetz, N.A., and Duncan, G.J. (1988). Out of work, out of mind: Response errors in retrospective reports of unemployment. *Journal of Business and Economic Statistics,* 6(2).

Neter, J. and Waksberg, J. (1964). A study of response errors in expenditure data from household surveys. *Journal of the American Statistical Association.*

Oksenberg, L., Cannell, C., and Blixt, S. (1996). Analysis of interviewer and respondent behavior in the household survey. *AHCPR Publication No. 96-N016,* National Medical Expenditure Survey Methods 7, Rockville, MD: Agency for Health Care Policy and Research, U.S. Public Health Service.

Sperry, S., Edwards, B., Dulaney, R., and Potter, D.E.B. (1998, in press) Evaluating interviewer use of CAPI navigation features. In *Proceedings of the International Conference on Computer Assisted Survey Information Collection* (working title). New York: John Wiley and Sons.

Wojcik, M., and Baker, R. (1992). Interviewer and respondent acceptance of CAPI. *Bureau of the Census 1992 Annual Research Conference Proceedings.*