THE USE OF A MULTIPLE ENTRY DIARY IN A PANEL STUDY ON HEALTH CARE EXPENDITURE

Robert A. Wright, Richard H. Beisel, Julia D. Oliver, Michele C. Gerzowski
National Center for Health Statistics

A major problem in the production of reliable national statistics on the health status of the country is that most of the data gathering techniques rely on recall periods of up to a year. Using a recall period of this length may cause substantial reporting errors to occur.

Review of Related Literature

The use of diaries to obtain health care information has been shown to be an effective way to help reduce this problem in collecting health care information.

Prior to 1970, the diary as a means of securing data had been investigated primarily in the fields of consumer expenditures and food consumption.[1,2] Both journal and ledger diaries have been used in consumer expenditure studies. In journal diaries, the entries are made in time sequence, while in ledger diaries they are made by categories of events such as visits to the doctor, absences from work and purchases of medicines. For consumer expenditure studies, the work of Sudman and Ferber (1971)[3], has shown that ledger diaries obtain more complete information and a higher level of reporting. However, in health expenditure studies a form of diary, "an attractive household calendar," was used in the Canadian Sickness Survey[4] as early as 1952 and a diary was used in the family health survey conducted by the University of California School of Public Health in 1954.[5] In these instances the diary served as a reminder for reporting at the interview. In 1954, Allen et al. conducted a study on the characteristics and quality of data obtained by diary keeping. It was found that those maintaining diaries reported higher rates of illness for a calendar month than did those who were interviewed for whatever form of rate was used as an expression of morbidity levels. But rates for the groups maintaining diaries were close to rates for interviewed groups when medically attended illness was considered. Differences between rates for disabling illness were less than differences between rates for nondisabling illness. The difference between diary rates and interview rates was about the same for males, females, preschool children, school children and adults under 65. For older persons, a group of special interest in health surveys, diary rates were much closer to interview rates than for other groups.

Some of the most systematic work on the use of diaries for medical care has been by Sudman et al., at the Survey research Laboratory of the University of Illinois. They conducted an experimental study from October 1973 through March 1974 to investigate the use of diaries to obtain health care information. The study focused on two methodological issues: (1) what should be the format (i.e., ledger or journal) for a medical diary?, and (2) should households be compensated for keeping records?

Respondents were chosen from the membership in two prepaid health maintenance organizations (HMO's). An attempt was made to validate all medical events which resulted in a visit to an HMO physician or to a clinic or hospital for outpatient or overnight care. The results of this experiment confirm the results of early studies which found that households that use a diary report a higher number of medical events than do households who recall information on a personal interview. There seemed to be no meaningful differences between ledger and journal diaries.

Medical Economic Research Study Methodology

A multi-entry format was designed for the diary used in the Medical Economic Research Study. This study was designed to test various methods of collecting the medical care expenditures for families and individuals. Because average family out-of-pocket expenditures for medical care does not adequately serve the needs of policy makers for data of this type the research study had two components. The first component consisted of a longitudinal survey of a panel of households. The second component consisted of a survey of all providers and third party payors mentioned by the persons in the household survey.

The research study tested several methods of household data collection, i.e., monthly versus bimonthly interviewing, telephone versus personal interviewing and receiving permission to do the provider survey early in the survey period versus receiving such permission at the last interview. The research study was conducted in a rural area and an urban area and included areas where a high proportion of persons belonged to an HMO.

To assist household respondents in keeping track of health care information during the survey period a diary was designed and distributed to all participating households. The diary incorporated together features used separately by other longitudinal surveys: a calendar portion for recording health care events chronologically; a ledger portion where details of health care events could be recorded in different sections, one section for each type of event; and a pocket for household members to collect bills, receipts, and other information which would aid the household respondent in recalling events. A diary using three distinct methods of entry was used for a number of reasons. First, respondents vary in their ability to understand and use something as complex as a ledger. For reasonably educated respondents with a strong motivation, the ledger provides a place on which to record almost all the information needed about a particular event. For persons who use a calendar,

even if they don't, won't or can't use a ledger, the calendar provides a place for recording some details. At a minimum the fact an event took place can be indicated and the date is thus known. As bills come in, the pocket provides a convenient receptable for all respondents to reduce searching for them at interview time.

At the final interview at the household, the diary was picked up by the interviewer. The household was not informed at the initiation of the study that the diary would be collected. Thus, the fact that the diary was to be collected caused no pressure on the respondents to utilize the diary.

The coding of the diary was done by the staff of the Utilization and Expenditure Statistics Branch in the Division of Health Interview Statistics.

Findings and Analysis

The analysis is primarily concerned with how diary use is related to certain demographic variables: urban versus rural residence; age of head of household; education of head of household; and income.

In addition, two variables concerning the actual interviewing technique were used: personal versus telephone interviews; and monthly versus bimonthly interviews.

Before beginning the analysis, those households who reported no medical events during the survey were dropped from consideration. They were not included, because they could not possibly use the diary if they had nothing to record. There were 32 such households. Their elimination dropped the population size to 650.

At the beginning of the study, neither the interviewers nor the respondents were told that the diaries would be picked up at the end. The respondents were told, in fact, that the diaries were to be used as a memory aid for subsequent interviews and were theirs to keep. When the interviewers returned to collect the diaries they picked up 66 percent of them. The remaining 34 percent were either lost, thrown away, or not collected for some other reason (see Table 1).

TABLE 1. Diary Status at the End of the Study

Returned diary		428 (65.8%)
1. used diary	311 (47.8%)	
2. did not use diary	107 (16.5%)	
Did not return diary		222 (34.2%)
1. lost diary	126 (19.4%)	
2. threw diary away	36 (5.5%)	
3. all other reasons for not returning diary	60 (9.2%)	
TOTAL		650 (100.0%)

For the 222 households in Table 1 who did not return the diary, there is no way of knowing if the diaries were used or not before they were lost or thrown away. We, therefore, do not know the overall proportion of diary usage in the population or the demographic characteristics of all the users. There is only information on those 428 households who returned the diary. We also do not know if the group who returned the diary was somehow different from the group who did not.

The demographic characteristics of the two groups are known, however. Table 2 shows these characteristics for the two groups. Table 3 shows the demographic characteristics of the households who used the diary and of those who did not, within the population of those who returned the diary.

TABLE 2. Selected Demographic Characteristics of Those Who Returned the Diary and Those Who Did Not

	Diary returned	Diary not returned
Percent of total population	66%	34%
Number	428	222
Median age of head of household	46 (yrs.)	43 (yrs.)
Median income	$13,900	$12,900
Median education of head of household	12 (yrs.)	12 (yrs.)
Percent urban	48%	64%
Percent rural	52%	36%
Percent personal interviews	48%	47%
Percent telephone interviews	52%	53%
Percent monthly interviews	45%	53%
Percent bimonthly interviews	55%	47%
Percent diary use	70%	???

TABLE 3. Selected Demographic Characteristics of Those Who Used the Diary and Those Who Did Not Use the Diary for All Those Who Returned the Diary

	Diary used	Diary not used
Number	311	117
Percent of population who returned the diary	73%	27%
Median age of heat of household	44 (yrs.)	46 (yrs.)
Median income	$15,100	$10,800
Median education of head of household	12 (yrs.)	12 (yrs.)
Percent urban	41%	62%
Percent rural	59%	38%
Percent personal interviews	48%	50%
Percent telephone interviews	52%	50%
Percent monthly interviews	47%	41%
Percent bimonthly interviews	53%	59%

The 428 households who returned their diaries were the ones used in the analysis. Usage versus no usage was used as a dependent variable in a stepwise regression analysis. The independent variables used were the demographic and interview variables mentioned earlier. For the basic question of diary usage, three regressions were used: diary use versus no diary use; calendar use versus no calendar use; and ledger use versus no ledger use.

In all 3 regressions, income came out as the most significant independent variable with rural residence second (see Table 4). What was surprising was that education was not included as a significant variable for either diary or ledger usage. It was only included for calendar usage, but was number 6 our of 6.

TABLE 4. Independent Variables Kept in Regression Equation by Order

Diary Use	Calendar Use	Ledger Use
1. Income	Income	Income
2. Rural residence	Rural residence	Rural residence
3. Telephone interviewing	Age	Telephone interviewing
4. Age	Monthly interviewing	Age
5. Monthly interviewing	Telephone interviewing	Monthly interviewing
6. —	Education	—

The second phase of analysis dealt with how well the diary was used, what items were recorded most and least often, and what, if any, time trends were observed. Certain items, such as hospital stays, listed on the Summary were matched with those same items recorded on either the calendar or ledger. Correlation coefficients were then determined for certain events as they were recorded on the Summary and on either the calendar or ledger. A problem to note about this approach is that those items listed in the Summary were those which the respondent reported to the interviewer in the normal course of the interview. They were not necessarily all the events which acutally occurred in the household. What the correlations actually represent, then, are comparisons between those items given orally to the interviewer, which were then transcribed to the Summary, and those items which the respondent himself wrote in the calendar or ledger. The correlations then represent the correlations between two methods of recall.

Table 5 shows the correlations for types of events as recorded in the Summary and in the calendar. Similarly, Table 6 shows correlations for types of events as recorded in the Summary and in the ledger. The calendar protion of the diary had only a blank square surrounding the date in which to record an event. The ledger was more formalized. There were specific spaces for specific items; such as doctor and dental visits, prescriptions, and hospitalizations; and check boxes for those visits which included x-rays, lab tests, and visits to the emergency room. There was also a space for "other medical expenses" which included non-medicinal items prescribed or recommended by a doctor, i.e., corrective shoes, crutches, syringes for diabetics. The ledger also provided special space to record not only the encounter, but the date, provider's name, prescription numbers and conditions.

TABLE 5. Ranking of Correlations Between Visits Recorded in the Summary and Visits Recorded in the Calendar

Type of Visit	Correlation
Dental, Feb. - March	.758
Doctor, Feb. - March	.673
Hospital Stays, Feb. - July	.663
Dental, April - May	.654
Dental, June - July	.627
Doctor, April - May	.580
Doctor, June - July	.517
Medicine, April - May	.495
Medicine, Feb. - March	.449
Emergency Room Visits, Feb. - July	.401
X-rays, Feb. - July	.399
Medicine, June - July	.370
Lab Tests, Feb. - July	.340
Other Medical Expenses, Feb. - July	-.461

TABLE 6. Ranking of Correlations Between Visits Recorded in the Summary and Visits Recorded in the Journal

Type of Visit	Correlation
Medicine, April - May	.868
Emergency Room Visits, Feb. - July	.815
Medicine, Feb. - March	.705
Hospital Stays, Feb. - July	.659
Doctor, Feb. - March	.635
Dental, Feb. - March	.632
Dental, April - May	.542
Medicine, June - July	.497
Doctor, April - May	.473
Dental, June - July	.448
X-Rays, Feb. - July	.420
Doctor, June - July	.394
Lab Tests, Feb. - July	.393
Other Medical Expenses, Feb. - July	-.548

TABLE 7. Correlations Between Visits Recorded on the Summary and Visits Recorded in the Diary, by Type of Diary Use

Type of Visit	(1) Correlation between visits recorded in the Summary and visits recorded in the calendar	(2) Correlation between visits recorded in the Summary and visits recorded in the journal	Col. (1) minus(-) Col. (2)
Dental, Feb. - Mar.	.758	.632	+
Dental, April - May	.654	.542	+
Dental, June - July	.627	.448	+
Doctor Feb. - Mar.	.673	.635	+
Doctor, April - May	.580	.473	+
Doctor, June - July	.517	.394	+
Medicine, Feb. - Mar.	.449	.705	-
Medicine, April - May	.495	.868	-
Medicine, June - July	.370	.497	-
Hospital Stays	.663	.659	+
Emergency Room Visits	.401	.815	-
X-Rays	.399	.420	-
Lab Tests	.340	.393	-
Other Med. Exp.	-.461	-.548	+

Table 7 compares correlations for specific encounters between those visits recorded on the Summary with those recorded in the calendar and those visits recorded on the Summary with those recorded in the ledger. The results were mixed with some events, emergency room visits, prescribed medicine, having higher correlation with encounters on the ledger, some other events, dental and doctor visits having higher correlation with encounters on the calendar.

Another aspect of the analysis was to see if there was any trend as the study progressed for changing habits of recording items in either the calendar or ledger. This could be determined for doctor visits, dental visits and prescribed medicines as they were coded by two-month time periods; February - March, April - May, June - July. Tables 8 and 9 show the results.

TABLE 8. Correlations Between Visits Recorded on the Summary and Visits Recorded on the Calendar by Time Period

	Feb.-Mar.	Apr.-May	June-July
Dental Visits	.758	.654	.627
Doctor Visits	.673	.580	.517
Prescribed Medicines	.449	.495	.370

TABLE 9. Correlations Between Visits Recorded on the Summary and Visits Recorded on the Ledger by Time Period

	Feb.-Mar.	Apr.-May	June-July
Dental Visits	.632	.542	.448
Doctor Visits	.635	.473	.394
Prescribed Medicines	.705	.868	.497

Both tables show a decrease in correlation by time period, indicating less usage of the diary as the study progressed with increases in April-May for prescribed medicines.

Conclusions

In retrospect, this data set could have furnished more information on diary usage if we had known more about those households who did not return the diary. If they had been asked, "Before you lost (threw away) your diary, did you use it at all? If so, what sections?", there would have been a better grasp of the differences between diary users and non-users. Also, looking at other demographic variables, particularly race, would have added interest. However, we did find out how households used the diary as a memory aid, what items they used if for and how they used it over time.

References

1 Brady, Dorothy S. and Williams, Fath M. "Advances in the Techniques of Measuring and Estimating Consumer Expenditures." Journal of Farm Economics, 27, No. 2 (May 1945), 315-344.

2 Murray, J.E.C., Blake, D. Dickens, and Moser, A.M. "Collection Methods in Dietary Surveys." Southern Cooperative Series, Bulletin 23. Clemson, S.C.: South Carolina Agricultural Experiment Station, April 1952.

3 Sudman, Seymour and Feber, Robert. "Experiments in Obtaining Consumer Expenditures by Diary Methods." Journal of the American Statistical Association, 66, No. 336 (December 1971), 725-736.

4 Peart, A.F.W. "Canada's Sickness Survey: Review of Methods." Canadian Journal of Public Health, 43, No. 10 (October 1952), 401-414.

5 Allen, George I., et al. "Interviewing Versus Diary Keeping in Eliciting Information in a Morbidity Survey." American Journal of Public Health, July 1954, pp. 919-927.