A RECORDS - SURVEY COMPARISON OF ELIGIBILITY AND HEALTH CARE UTILIZATION MEASURES FOR MEDICAID BENEFICIARIES: ADULT AND CHILD REPORTS

Donna Eisenhower and Angela Schmitt, Mathematica Policy Research, Inc.
Donna Eisenhower, MPR , Box 2393, Princeton, NJ 08540

Key Words: Record check, Self-proxy report

Background

This study is part of a larger effort to assess the feasibility of conducting a national survey of Medicaid beneficiaries. The study has been conducted by Georgetown University's Center for Health Policy Studies and Mathematica Policy Research under a grant from the Physician Payment Review Commission. A national survey is being considered to produce state-specific estimates and generate enough information for monitoring access-to-care. Many sampling and other survey design issues need to be evaluated to determine the feasibility of conducting such a survey.[1] This paper reports on an effort to assess the accuracy of survey estimates related to Medicaid eligibility and health care utilization for adults reporting for themselves and for adults reporting for one randomly selected child. The accuracy of the survey estimates will be assessed by comparing the survey data to Medicaid records data.

The survey data are the result of a pilot survey of Medicaid beneficiaries that was conducted in one northeastern state between August 1 and October 25, 1993. The survey resulted in 358 completed adult self-reports and 320 adult reports for one randomly chosen child. Medicaid records data were abstracted for these same individuals after waiting an additional three months to assure that any lagging medical claims were processed.

To conduct a national survey requires that eligibility information be accurately reported by survey respondents or collected from records. If the accuracy of the Medicaid eligibility determination is in question, all measures related to access or utilization will be meaningless if such measures are evaluated in terms of Medicaid participation. This matter is complicated in Medicaid populations where a portion of the participants have inactive or active eligibility (go on or off Medicaid) during the course of a year. Therefore, accessing the ability of respondents to report eligibility accurately is critical to evaluating the feasibility of using survey data or the related need of abstracting such data from records.

Secondly, the assessment of the accuracy of utilization measures provides a measure of accuracy in general and also provides a basis of assessing the feasibility of collecting accurate measures from a survey. Recall bias in collection of health care utilization measures is well documented in the literature.[2] The record-survey comparison will shed light on the need for the use of bounding or memory aids,[3] the appropriate recall period, the need for statistical adjustments of the estimates, or the need to use records data for certain measures.

The measures for which survey responses will be compared to records data for both the self-report of the adult and the adult report for the child include:

- Eligibility for Medicaid by month over the past year

- Number of doctor's visits for either last month or last three months[4]

- Date of admission for overnight hospitals stays

- Number of nights hospitalized

Results

Eligibility. Several measures of eligibility were assessed as part of the full study and are available as part of the final report. The most important eligibility question asked of survey respondents regarding eligibility was, "Now thinking back over the last twelve months beginning in July 1993 and ending in August 1992, please tell me to the best of your ability if your (then the child's) eligibility for Medicaid was active for every month during that year?" As Table 1 indicates, 91.8 percent of survey responses for adults matched the records data. Of these, 71.5 percent were active for all twelve months and matched and 20.3 percent matched but were not active for all twelve months. Of the 8.2 percent of responses which were mismatches, 4.8 percent

mismatching was because the survey said they were active when the record data showed they were not, and 3.4 percent mismatching was because the record data said they were active and the survey said they were not. At 89.0 percent, the number of survey responses of the adult reports for the child was very similar to those for adult self-reports.

While this is a modest degree of mismatch, it should be improved if survey data alone is to be used to produce national estimates. Therefore, two subsequent analyses were conducted to examine the reasons for mismatch. The first examined if recall bias explained the mismatches. If recall bias alone accounted for the mismatches, more mismatches should occur further back in time. While there was a pattern in that direction, it was not consistent for either the adult or child report.

The analyses explored the concept of "transition month" as a reason for the mismatch. The transition month is the last month either with or without Medicaid and the ensuring month where there is a change of status. An analysis of this phenomenon showed a mismatch rate for adults of 40.3 percent in transition months compared to a 10.3 percent mismatch rate in non-transition months.[5] Similarly, the mismatch rate for the adult report for a child was 33.3 percent in transition months compared to 10 percent in non-transition months. Clearly, the greater percentage of mismatches occur during a transition month when a Medicaid beneficiary may not even be aware that their official status has changed. Regarding eligibility status, lack of awareness rather than recall bias is more of an explanation. This actually suggests that the survey data may be adequate if the data users account for the transition month phenomenon when looking at other survey measures.

Utilization

Doctor Visits. The first utilization question to be compared to records data read, "During the last month (or past three months), how many times did you see a medical doctor for an in-person visit?"[6] As mentioned, one-half the sample responded to a one-month recall period and one-half to a three-month recall period. The findings presented in Table 2 for adults indicate that for exact matches, the one-month recall is significantly more accurate with 51 percent of the survey responses matching records data compared to 26.9 percent for the three month-recall period.

When a match is defined reasonably as ±1 visit, 78.4 percent match for the one-month recall for adults while 58.2 percent match for the three-month. Of adults reporting, 83 percent reported 0, 1, or 2 visits within the last month (35 percent reported 1 visit). Within the last three months, of adults reporting 75 percent reported 0, 1, or 2 visits (42 percent reported 1 visit).[7]

Clearly, a one-month recall period for doctor visits for adults is preferred to a three-month recall period. In fact, if accurate estimates greater than ±1 visits are required, a two-week recall period may be necessary.[8] As Table 2 also indicates, when adult respondents err, more (80 percent) overestimate the number of visits at a one-month recall period while slightly more (55 percent) underestimate at a three-month recall period. This is evidence of some telescoping at a one-month period and some recall decay at three months. This also suggests that accuracy would be improved by the use of bounding techniques for either a one-month or two-week recall period.

As for most measures, there is at least slightly greater accuracy for adults reporting for a child than for adults reporting for themselves. As Table 2 indicates, exact matches for children were 57.6 percent for the one-month recall period and 50.6 percent for the three-month recall period. The most noticeable improvement for children is at the three-month period with 50.6 percent compared to 26.9 percent for adults. However, even 50.6 percent is less than the desired accuracy for survey estimates. When a match is defined as ±1 visit for children, the matches for children increase dramatically to 93.2 percent for the one-month recall period and to 78.5 percent for the three-month.

If a match is defined as ±1 visit for children, the one-month recall period provides a good survey estimate of the number of doctor visits. An exact match could probably be improved further by using a two-week recall period. Error for the one-month recall period is usually an overestimate of one visit suggesting telescoping. This problem can also be improved by the use of bounding procedures. Error for the three-month period is slightly more often (53.8 percent), an underestimate suggesting some recall decay.

Hospitalizations. The second utilization question asked, pertains to overnight hospitalizations and reads "Since July 1, 1992 a year ago, were you a patient in a hospital OVERNIGHT?" Hospitalizations are large events subject to telescoping in survey reporting. As Table 3 indicates, only 63.3 percent of the adult hospitalizations reports

for yes/no responses matched records data. Most of this error is due to overestimating, suggesting the effect of telescoping in hospitalizations from a previous period or to less likely record error (some reports may be missing due to unexpected lags in billing). For those reporting "yes",they then answer the question, "How many different times did you stay in any hospital overnight or longer since July 1, 1992 a year ago?" As Table 3 indicates, 84.1 percent have exact matches on the actual number of hospital stays. Of these, all errors were either missing one (8.7 percent) or two (7.2 percent) overnight stays. Therefore, getting more accurate estimates for adult hospitalizations requires increasing accuracy at the "yes" or "no" question. Perhaps asking the number of hospitalizations directly over the last year without asking a yes/no question first would in itself improve response accuracy.

For children, the accuracy in the yes/no response to the first question on whether hospitalized is almost equal to that for adults at 63.2 percent. Once again, most err in the direction of saying "yes" when the records suggest the correct answer is "no". The percentage of matches for the number of hospitalizations reported for children is 83.3 percent: again almost equally accurate to that of adults. When there is error in reporting the number of hospitalizations for children, most are underestimates off by one less overnight stay than the records indicate. Although this would be unusual, there may be an unexpected lag in the billing or reporting in the records accounting for this difference. Hospitalizations are large, infrequent events and for these the accuracy is more equivalent for both adult self-reports and adults reports for a child. For both adults and children, the clear majority (88% for children, 87% for adults) reported only 1 overnight stay in the hospital over the previous year.[9]

Questions during the hospitalizations were also asked as follows: "Since July 1, 1992 a year ago, on what date did you enter the hospital the last time? The time before that?" A high 94.8 percent of the month of hospitalization matched for adults. The match for the full date which included the day was 70.7 percent. The correct match for the number of nights of hospital stays was 56.9 percent for exact match and 86.2 percent if a match is defined as ±1 night. The survey question for this reads, "For the stay beginning (READ DATE), how many nights were you in the hospital?" The respondents interpretation or perception of days verses nights in the hospital may be the source of error for this question. If they were in the hospital for a substantial part of a day, they might want to report this as an overnight's stay. In other words, the respondent may want to get credit for staying in the hospital and want to actually report "days"in the hospital.

There was an approximately equal percentage of matches on month of hospitalizations for children as adults at 93.3 percent. There was a much better match on the full date of hospitalization for children at 90 percent. The exact match on number of nights stay for children was 53.3 percent (slightly less than that for adults) or 83.3 percent, if the ±1 night criterion for match is used.

Conclusion

The survey literature supports the premise that eligibility and health utilization measures will be subject to recall bias in respondent reporting in **any** population, including a poorer, less educated one. In fact, the literature suggests the recall error increases proportionately with the length of the recall period for such measures. The findings from this record check study support these premises for utilization measures. However, error in respondent reporting for eligibility status is better explained by the lack of awareness of their status during a one to two month transition period rather than to recall bias. Certain measures of eligibility and utilization data are more critical than others to a national survey of Medicaid beneficiaries and accuracy in these measures should be emphasized.

Whether the respondent was actively on Medicaid over the entire 12-month period is critical information. The match of survey to records data is much higher than the noncritical for this critical information at 91.8 percent for adult report and 89 percent for the child report. However, this information is so critical that it needs to be even closer to a 100 percent match to dismiss the problem. A subsequent analysis indicated that mismatch in survey response error was largely due to a transition month phenomenon which resulted in lack of awareness of the respondent as to their Medicaid status. Recall bias was shown to be less of a problem for respondent reports of eligibility for both the adult and child. Therefore, accuracy in respondent reports of eligibility may be sufficient to support the purposes of the national survey.

Depending on the emphasis on the national survey, precise utilization measures may be less critical than this eligibility information. The one-month recall period for doctor's visits was a better match (more accurate) for both adult and child

reports, but substantially better for adult reports. If the match is defined as $\pm$ 1 visit, the match is improved dramatically, especially for child reports. The three-month recall period for doctor visits is of poor accuracy for adult reports and not adequate but better for the child reports. The one month recall period introduced telescoping in of visits, the three-month period was subject to recall decay. These findings suggest that accuracy in reporting doctor's visits may be optimum at a 2-week recall period, especially if used with bounding procedures to reduce the effect of telescoping.

The reporting for hospitalizations is good at 83-84% marker for the number of hospitalizations occurring during the past year. However, hospitalizations are rare events (the majority of adults and children only reported 1 hospitalization) in this young population and should be reported more accurately. The greatest source of error was in the initial question which asked whether or not there was a hospitalization during the previous 12 months. Only 63 percent of the adult and child reports matched the record. Furthermore, the accuracy was highest at around 93-94% for the reporting for month of occurrence. These findings indicate that asking the more specific information on hospitalizations first will improve the accuracy of reporting. Certainly, skipping the question on whether hospitalized in favor of asking the number of hospitalization should reduce error in respondent reports. Finally, the accuracy of reporting of number of nights stayed was not good and might be improved by emphasizing "nights" not "days" in the hospital.

Accuracy of utilization measures need to be improved by either shortening the recall period; using memory aids and bounding procedures (a costly alternative); or by supplementing survey data with records data. It is important to keep in mind that while these measures are subject to recall bias, there are other critical measures such as access to care or satisfaction measures which will be asked on a national survey. Because these measures do not involve precise frequency of event reporting, they are not likely to be subject to the recall problem to this extent. Recall bias is more problematic in behavioral frequency questions.

ENDNOTES

1. See "Assessment of Access to Care Pilot Survey of Medicaid Beneficiaries" by Eisenhower et al., Physician Payment Review Commission, November 24, 1993.

2. A review of the literature is available as part of the final report.

3. Bounding involves presenting some information gathered from one period for a respondent to that respondent for a subsequent survey. Memory aids include the use of checklists, flashcards, maps, pictures, and calendars.

4. The survey sample was randomly assigned to either a one or three month recall period.

5. Complete tables are available as part of the final report or upon request.

6. The probe read "Include all types of doctors, such as dermatologists, psychiatrists and ophthalmologists, as well as general practitioners and osteopaths. Do not count times while an overnight patient in a hospital.

7. This data is not presented in a table but reported from a separate report.

8. A two-week recall period was not tested as part of this study.

9. Again, this data is not reported in a table but elsewhere in a separate report.

TABLE 1

MEDICAID ELIGIBILITY FOR ALL MONTHS--AUGUST 1992 TO JULY 1993

Adults	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Match Not Active[1]	59	20.3	59	20.3
Match Active[2]	208	71.5	267	91.8
Mismatch Survey Active	14	4.8	281	96.6
Mismatch Record Active	10	3.4	291	100.0

Frequency Missing = 5

Children	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Match Not Active	69	22.3	69	22.3
Match Active	207	66.8	276	89.0
Mismatch Survey Active	20	6.5	296	95.5
Mismatch Record Active	14	4.5	310	100.0

Frequency Missing = 5

[1] This is a match where the respondent said they were not actively on Medicaid for all 12 months.

[2] This is a match where the respondent said they were actively on Medicaid for all 12 months.

TABLE 2

SUMMARY STATISTICS ON MISMATCH/MATCH FOR ONE MONTH AND THREE
MONTH RECALL OF DOCTOR'S VISITS - ADULT AND CHILD

I. ADULT SELF-REPORT

A. One Month Recall - Adult	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Mismatch	25	49.0	25	49.0
Match	26	51.0	51	100.0
Magnitude -				
Minus	4	16.0	4	16.0
Plus	21	84.0	25	100.0
B. Three Month Recall - Adult				
Mismatch	49	73.1	49	73.1
Match	18	26.9	67	100.0
Magnitude -				
Minus	27	55.1	27	55.1
Plus	22	44.9	49	100.0

II. ADULT REPORT FOR THE CHILD

A. One Month Recall - Child	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Mismatch	25	42.4	25	42.4
Match	34	57.6	59	100.0
Magnitude -				
Minus	6	24.0	6	24.0
Plus	19	96.0	25	100.0
B. Three Month Recall - Adult				
Mismatch	39	49.4	39	49.4
Match	40	50.6	79	100.0
Magnitude -				
Minus	21	53.8	21	30.8
Plus	18	46.2		100.0

TABLE 3

A. HOSPITALIZATION: ANY OVERNIGHT STAYS DURING THE PAST YEAR (YES/NO RESPONSE)

Adults	Frequency	Percent	Cumulative Frequency	Cumulative Percent
I/P CLAIM NO SURVEY	12	11.0	12	11.0
NO I/P CLAIM MATCH	28	25.7	40	36.7
I/P CLAIM MATCH	69	63.3	109	100.0

Children	Frequency	Percent	Cumulative Frequency	Cumulative Percent
I/P CLAIM NO SURVEY	6	10.5	6	10.5
NO I/P CLAIM MATCH	15	26.3	21	36.8
I/P CLAIM MATCH	36	63.2	57	100.0

B. HOSPITALIZATIONS: OF THOSE SAYING "YES," DEGREE OF MATCH OF OVERNIGHT STAYS

Adults	Frequency	Percent	Cumulative Frequency	Cumulative Percent
2 Less than record	5	7.2	5	7.2
1 Less than record	6	8.7	11	15.9
EQUAL	58	84.1	69	100.0

Children	Frequency	Percent	Cumulative Frequency	Cumulative Percent
1 Less than record	5	13.9	5	13.9
EQUAL	30	83.3	35	97.2
5 More than record	1	2.8	36	100.0