

Completeness of Household Reported Medical Expenditure Data for Physician Office Visits in the Medical Expenditure Panel Survey

Steven R. Machlin, Center for Financing, Access, and Cost Trends,
Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850
and Diana Wobus, Westat, Research Blvd., Rockville, MD 20850

KEY WORDS: MEPS, medical expenses, physician visits, completeness

Introduction

The Medical Expenditure Panel Survey (MEPS) is a complex national probability survey of the civilian noninstitutionalized population, and has been conducted on an annual basis since 1996 by the Agency for Healthcare Research and Quality (AHRQ). One of the primary purposes of the survey is to collect data that can be used to analyze national medical expenditures (i.e., the amount paid for health care services).

Unfortunately, it is difficult to obtain complete information on medical expenditures from household survey respondents because the type of information being collected is often not straightforward and requires extensive record keeping over time, especially for households with members that frequently use the health care system. Further, in a significant number of instances, respondents are simply not aware of either the total amount billed or how much the provider is paid for the services that were received. Classic examples are individuals enrolled in the Medicaid program, where financial transactions occur primarily between the provider and the state Medicaid agency, and enrollees of Health Maintenance Organizations (HMO) or other managed care plans who only may be aware of paying some predetermined co-payment that is not necessarily related to the total amount the provider receives (Cohen et. al., 1997).

As a consequence of these factors, there is a substantial amount of item nonresponse on medical expenses in the MEPS household survey component (HC). To compensate for these missing data and to improve accuracy, data on expenses for sample persons are also collected from a sample of their health care providers in the Medical Provider Component (MPC) of MEPS (Machlin and Taylor, 2000). Office-based care to physician offices constitutes the largest category of provider visits in the MEPS, comprising more than 1.2 billion visits in 2003. While generally more complete and accurate than HC data, it is not practical to collect

MPC data on medical expenditures from all office-based physician offices that provided care to sample persons because of survey budget constraints and lack of full cooperation among respondents and providers. Consequently, while the MPC is the preferred source of data, it is necessary to also use HC data and imputation techniques to complete MEPS expenditure data for office-based physicians (Machlin and Dougherty, 2004).

This paper examines the extent to which expenditures reported in the 2003 MEPS-HC for visits to physicians' offices are complete. Variations in completeness according to selected characteristics of the household, respondent, interview, and person receiving care are also examined. A complementary analysis examines the extent to which household reported data on expenditures for physician visits are accurate (Kashihara and Wobus, 2006). Taken together, these papers help inform the extent to which it is possible and appropriate to rely on household data in the MEPS-HC for estimates of office-based physician expenses.

MEPS Sample Design

The sample of households for the MEPS-HC is a subsample of households that responded to the prior year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics (National Center for Health Statistics, 2002). The MEPS sample is drawn from approximately half of the PSUs selected for the NHIS. For example, the 1996 MEPS-HC sample was selected from households that responded to the 1995 NHIS (Cohen S., 1997). This selection was comprised of 195 Primary Sampling Units (PSUs) and 1,675 sample segments (second-stage sampling units). Over-sampling of households with Hispanics and blacks carries over from the NHIS to the MEPS sample design.

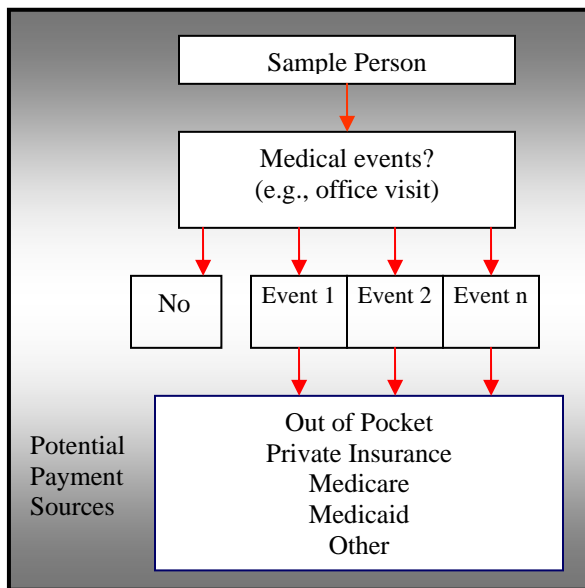
The sample design of the Medical Expenditure Panel Survey is an overlapping panel design, with data collected for each new MEPS panel covering a two-year period (Cohen J., 1997). As a result of the overlapping panel design, MEPS annual data for 1997 and beyond are constructed based on data collected from two consecutive panels.

MEPS Household Expenditure Data Collection

Primary data collection in the MEPS-HC employs computer-assisted personal interviewing (CAPI). The HC questionnaire is designed to collect use and expenditure data for two consecutive years through a series of five interviews. In general, annual health care utilization and expenses for sample persons are derived from information collected in 3 of the 5 interviews (Cohen J., 1997).

Figure 1 provides a summary of the data collection process for medical events (utilization) and expenditures associated with these events in MEPS. For each person in a sample household, the core instrument first asks the respondent to identify all health care visits that occurred since the previous interview (since January 1 for the first interview), and then asks about expenditures for each health care event reported (see definition of expenditures in the section below). While data on all types of medical events are collected (e.g., office-based medical provider visits, hospital emergency room visits, hospital outpatient visits, hospital inpatient stays, dental visits, home health, prescribed medicines, and other medical expenses), this analysis is based exclusively on office-based medical provider visits.

Figure 1. Illustration of collection of medical event and source of payment data in MEPS



Expenditures for each event are collected according to the following 10 sources of payment categories: Out of pocket, Medicare, Medicaid, Private insurance, Veteran’s Administration, TRICARE, Other Federal sources, Other State and Local Sources, Workers’ Compensation, and Other unclassified sources.

Payments for a particular medical event are typically made by one or a combination of sources.¹ Total expenses for a given event are obtained by summing across all payment sources.

Completeness Defined

This study is based on office-based physician visits (events) *reported* in the survey. Thus completeness of expenditures is conditional on the events that were reported, and our analysis does not examine the extent to which expenditures may be incomplete due to underreporting of events. A reported event is defined as “complete” if expenditures are reported for *all* of the *potential* sources of payment (i.e. none are missing). For example, if a person was covered by private insurance and had a routine visit that required a co-payment, then the expenses for that visit would be complete if the person reported *both* the amount they paid out of pocket to the provider *and* the amount the insurer paid to the provider (note that \$0 for any particular source can be a valid response). One common pattern is for respondents to report the amount paid out-of-pocket for the visit but then to not know or report the amount paid by one or more third-party sources.

While our definition of completeness for an event requires that all potential sources of payment be reported, it should be noted that completeness in reporting is substantially higher for out of pocket payments than for other sources. While 94 percent of the office-based events in our analysis were complete on out-of-pocket payment, only about one-quarter of the events for persons with private health insurance were complete on the amount paid by private insurance.

Data and Methods

Our analysis is based on 120,113 office-based physician events reported in the 2003 MEPS-HC.² The overall completeness rate on expenditure information for these events was about 22 percent. We examined variations in completeness across all sources according to selected characteristics of the respondent (age, sex, health status, race/ethnicity, self vs. other family member), household (region, MSA status, and family poverty status), interview (round number, visits in round, and reporting aids used), and sample person (type of insurance coverage).³

¹ Total payments for a small proportion of events each year are considered to be \$0, which occurs when it is reported that no payments were or will be made by any source.

² Events that are part of a flat fee bundle in which a lump sum is paid for multiple visits are excluded.

³ See appendix table for variable definitions used in analysis.

We used a multivariate logistic regression model to determine which characteristics were significantly associated with the likelihood of an event being completely reported in the MEPS-HC, holding the other factors constant. The dependent variable was coded as 1 if all potential payment sources for the event were completely reported and 0 otherwise. Bivariate completeness rates for the independent variables as well as the results of the logistic model (expressed as odds ratios) are shown in Table A. All estimates are weighted to take into account the complex survey design, nonresponse, and post-stratification and standard errors were conducted using a Taylor Series approach. Only differences that are statistically significant at the .05 level or better are discussed below.

Results (Appendix Table A)

Among the respondent characteristics examined, there were significant but moderate associations between health status, race/ethnicity, and education level of the respondent and completeness (holding other characteristics in the model constant). More specifically the odds of completeness were 34 percent higher for college graduate respondents than those with less than 12 years of education, 31 percent higher for those in excellent/good versus fair/poor health, and about 50 percent higher for white non-Hispanics relative to the minority race/ethnic group categories. Age, sex, and respondent type were not significantly associated with completeness.

Of the three household characteristics included in the model, only MSA was statistically significant. Expenses for events in non-metropolitan statistical areas were slightly more likely to be complete than those in other areas (OR=1.17).

All of the interview characteristic variables included in the model were statistically significant. Moderate associations were found for the interview round and events per person in round variables. More specifically, expenses for events reported in the last round (i.e., round 5) were more likely to be complete than those reported in the first round, and expenses for events for persons with only one event reported in the round were slightly more likely to be complete than those for persons with multiple events reported for the round.

The reporting aids variable had an extremely large association with completeness. Respondents who used explanation of benefits (EOB) forms from insurers and those who used bills from providers (but no EOBs) were dramatically more likely to report complete expenses than those who relied on memory only (ORs of 15.5 and 9.2, respectively). Events where expenses were reported

using checkbooks and/or information on prescribed medicine bottles (but not EOBs or provider bills) were also more likely to be complete than those reported based on memory only (OR=1.65).

There was also a dramatic association between type of health insurance coverage and completeness. Among events for persons under 65 years of age, expenses for those covered by public only insurance were substantially less likely (OR=0.14) and those for the uninsured were substantially more likely (OR=3.2) to be complete than those covered by private insurance. Among persons age 65 and over, expenses for events covered by Medicare and supplemental public insurance were much less likely to be complete than those covered by Medicare only or Medicare and supplemental private insurance. In addition, events covered by fee-for-service payment arrangements were nearly twice as likely to be complete (OR=1.82) than those covered by HMOs or plans with gatekeepers.

Summary/ Discussion

We found that respondent's age and sex as well as family income (reflected in poverty status variable) and geographic region were not significant predictors of completeness. Surprisingly, we also found that expenses for events where a respondent reported for themselves were not significantly more likely to be complete on expenditures than those reported by a family member.

While moderate associations were found for several respondent (race/ethnicity, health status, and education), household (MSA, region), and interview (number of visits, round) characteristics, by far the most dramatic effects were for reporting aids used by the respondent and insurance coverage of the person with the visit. It is not surprising that events in which the respondent used EOBs from insurers or bills from providers were substantially more likely to be complete, since these aids provide the detailed expenditure information requested in the interview. Further, use of these aids may be an indicator of a more conscientious and thorough type of respondent.

It is also not surprising that type of insurance coverage of the sample person had a strong association with completeness of expenditure reports. Persons under 65 years of age with public insurance only (mainly Medicaid) were substantially less likely to have complete expenditures reported because payments to providers for Medicaid recipients are typically made directly by State agencies, so the recipient has no information regarding the amount paid by Medicaid. Conversely, it is likely events for uninsured persons were more likely to be complete because more expenses

are paid out of pocket and the respondent is less likely to have payment amounts from other sources to report. It is also not surprising that persons in HMOs or plans with a “gatekeeper” were substantially less likely to have complete expenditure reports. This result is due to the fact that members of these types of plans are frequently only aware of the required co-payment, and the amount paid by the insurer can only be obtained from the MPC or completed through imputation (Machlin and Dougherty).

Based on our results, the most hopeful strategy for obtaining more complete expenditure data from respondents is increased emphasis on the importance of using EOBs and provider bills by respondents. For example, 71 percent of the events where respondents used EOBs and 63 percent of those where the respondent used provider bills were complete, compared with only a 22 percent overall completeness rate. However, only 5 percent of events were reported by respondents who relied on EOBs and 10 percent of events were reported based on provider bills. While increased use of these aids could increase completeness rates, some respondents will never have access to these types of records due to uncontrollable factors such as the type of insurance coverage for the event or timing of the interview.

In summary, the U.S. health care system has a wide range of complex public and private financing arrangements, so it is difficult and unrealistic to collect complete information on health care expenses in a household survey. Therefore, wherever possible the MEPS survey relies on expenditure data collected in the MPC. Unfortunately, it is not cost effective to do a full MPC for all providers who care for MEPS sample persons. The other options to complete expenditure data include using data reported in the MEPS-HC or employing imputation methods. This paper on completeness of HC reported expenses, in conjunction with an analysis of the accuracy of HC reported expenses (Kashihara and Wobus, 2006), helps inform the extent to which it is possible and appropriate to rely on MEPS-HC data for estimates of expenditures for office-based physician visits.

References

Cohen J. Design and Methods of the Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. *MEPS Methodology Report No. 1*. AHCPR Pub. No. 97-0026.

Cohen J., Monheit A., Cohen S., et. al. (1997). “The Medical Expenditure Panel Survey: A National Health

Information Resource”. *Inquiry* 33: 373-389 (Winter 1996/97).

Cohen S. Sample Design of the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997. *MEPS Methodology Report No. 2*. AHCPR Pub. No. 97-0027.

Kashihara D. and Wobus D. (2006). Accuracy of Household Reported Medical Expenditure Data for Physician Office Visits in the Medical Expenditure Panel Survey. *2006 ASA Proceedings of the Section on Survey Research Methods (CD-ROM)*.

Machlin, S. and Dougherty D. (2004). Overview of Methodology for Imputing Missing Expenditure Data in the Medical Expenditure Panel Survey. *2004 ASA Proceedings of the Section on Survey Research Methods (CD-ROM)*.

Machlin S. and Taylor A. (2000), Design, Methods, and Field Results of the 1996 Medical Expenditure Panel Survey Medical Provider Component, Rockville, MD: Agency for Healthcare Research and Quality. *MEPS Methodology Report No.9*. AHRQ Pub. No.00-0028.

Zuvekas SH and Cohen JW. A Guide to Comparing Health Care Expenditures in the 1996 MEPS to the 1987 NMES. *Inquiry*. Spring 2002;39:76-86.

The views expressed in this paper are those of the authors and no official endorsement by the Department of Health and Human Services or the Agency for Healthcare Research and Quality are intended or should be inferred. The authors wish to thank David Kashihara for his contributions to the presentation and Trena Ezzati-Rice, Joel Cohen and Steven Cohen for their helpful reviews of the paper.

Table A. Completeness rates and odds ratios for characteristics included in model

Characteristics in model		Number of Events (unweighted)	Bivariate % Complete	Multivariate Beta Coeff.	Odds Ratio¹
Total		120,113	21.7		
Respondent Characteristics					
Age	<25	7,244	17.6	-0.05	0.95
	25-64	85,636	22.7	0.00	(Ref)
	65+	27,233	19.5	-0.08	0.93
Sex	Male	27,565	21.2	-0.03	0.98
	Female	91,769	21.9	0.00	(Ref)
Self rated health	Fair/poor	32,602	14.7	0.00	(Ref)
	Good-Excellent	87,511	23.7	0.27	1.31**
Race/Ethnicity	White non Hispanic	75,798	24.4	0.00	(Ref)
	Black non Hispanic	15,363	12.1	-0.45	0.64**
	Asian/multiple races non Hispanic	6,041	16.4	-0.35	0.70**
	Hispanic	22,132	12.1	-0.45	0.64**
Education level	Below HS (<12 yrs)	45,383	17.3	0.00	(Ref)
	HS graduate	54,319	21.1	0.05	1.05
	College graduate	20,411	28.9	0.29	1.34**
Respondent type ²	Self	69,227	21.8	0.04	1.04
	Family	50,181	21.7	0.00	(Ref)
Household Characteristics					
Poverty Status	Less than 125% FPL	30,696	12.1	0.00	(Ref)
	125-199% FPL	18,262	18.3	0.15	1.16
	200-399% FPL	33,909	22.0	0.18	1.19
	At least 400% of FPL	36,467	26.6	0.18	1.20
Region	Northeast	20,135	17.2	0.00	(Ref)
	Midwest	25,186	25.4	0.22	1.24
	South	45,290	23.5	0.25	1.28**
	West	28,722	19.4	0.12	1.13
Metropolitan Statistical Area	MSA	92,953	21.0	0.00	(Ref)
	NonMSA	27,160	24.6	0.15	1.17*
Interview Characteristics					
Round indicator	Round 1	17,091	19.2	0.00	(Ref)
	Round 2	26,989	20.4	0.05	1.05
	Round 3	28,409	21.4	0.00	1.00
	Round 4	28,162	22.1	0.13	1.14
	Round 5	19,462	25.5	0.29	1.33**
Events per round	1 event	18,189	24.1	0.00	(Ref)
	2 events	16,354	22.9	-0.12	0.89**
	3-5 events	32,788	21.7	-0.23	0.79**
	6+ events	52,782	20.5	-0.24	0.79**
Reporting aids ³	Explanation of Benefits (EOB)	4,691	71.0	2.74	15.50**
	Bill from provider	9,714	63.1	2.22	9.24**
	Calendar	28,033	13.0	-0.14	0.87
	Checkbook/Medicine bottle	1,855	30.2	0.50	1.65**
	Memory only	75,820	14.6	0.00	(Ref)
Insurance Coverage					
Type ⁴	Any private (under 65)	55,651	24.8	0.00	(Ref)
	Public only (under 65)	26,212	2.7	-1.96	0.14**
	Uninsured (under 65)	8,984	46.5	1.16	3.20**
	Medicare only (65 and over)	8,530	18.3	-0.48	0.62*
	Medicare/any private (65 and over)	15,950	21.4	-0.53	0.59**
	Medicare/public only (65 and over)	4,747	2.5	-2.41	0.09**
HMO/Gatekeeper ⁵	HMO/Gatekeeper	53,366	14.3	0.00	(Ref)
	Other	66,747	27.7	0.60	1.82**

1. * p<.05 ** p<.01

2. Excludes 705 events reported by proxy.

3. Although some respondents used multiple reporting aids, events were classified hierarchically along the continuum from EOB to memory only.

4. Excludes 39 events for persons age 65 and over reported as having no insurance coverage.

5. HMO/Gatekeeper includes events where it was reported the person was covered by an HMO or plan with a gatekeeper. Other group includes events for all other persons, including those reported as covered by traditional fee for service plans and uninsured persons.