

Linked Designs of the MEPS Medical Provider and Organization Surveys

Marc Zodet¹, Sadeq Chowdhury¹, Steven Machlin¹, and Joel Cohen¹

¹Division of Statistical Research and Methods, Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality, 5600 Fishers Lane, Rockville, MD 20857

Abstract

The Medical Expenditure Panel Survey (MEPS) is an ongoing household survey that yields national estimates of various health care metrics; including health care use, expenditures, and insurance coverage. The MEPS also collects information from a sample of health care service providers reported by the household. The medical provider data are an invaluable complement to the household reported data. Often more detailed and accurate, the provider data serve as the gold standard for MEPS expenditure estimates and are the source for MEPS expenditure imputations. Because of increased demand for data on organizational characteristics of providers and/or health care practices, the Robert Wood Johnson Foundation has sponsored a Medical Organization Survey (MOS) which will collect this type of data from a subset of the MEPS medical provider sample. This paper presents the underlying design considerations of the MOS (e.g., identification of the sampling frame, sample selection, construction of analytic weights, linking of person-level and organization level characteristics) and discusses the analytic potential of the data.

Key Words: Medical Expenditure Panel Survey (MEPS), survey design, weighting, establishment survey

1. Introduction

There is increased interest in how recent changes to the U.S. health care delivery system may be impacting the relationship between patients and provider organizations, specifically in terms of access to care, utilization of services, and overall healthcare spending. With Medicare, some state insurance programs, and commercial insurers demonstrating strong support for alternative payment models, such as the Accountable Care Organization (ACO), providers have been motivated to restructure their organizations and financial partnerships around more consolidated provider networks (e.g., group practices, networks of practices, hospitals, hospital-physician joint ventures, and other health care groups). The resulting impact of this restructuring is yet to be determined (i.e., whether or not these models will ultimately improve efficiency and quality of care). To effectively evaluate the impact these changes may have on access to, use of, and expenses for health care services, information on both providers and health care recipients is required (Lake, 2012). However, there is currently no nationally representative database linking provider characteristics to the characteristics of their patients. The Robert Wood Johnson Foundation has provided the Agency for Healthcare Research and Quality (AHRQ) the funding to conduct a new survey of health care

providers called the Medical Organization Survey (MOS) that is designed to fill this data void.

The Medical Expenditure Panel Survey (MEPS) collects data on health care utilization, expenditures, sources of payment, insurance coverage, and health care quality measures. The survey, conducted annually since 1996 by AHRQ, is designed to produce national and regional estimates for the U.S. civilian noninstitutionalized population (Ezzati-Rice, 2008). MEPS collects data from both household respondents (Household Component – HC) and from a sample of their health care providers (Medical Provider Component – MPC). AHRQ is undertaking the MOS by incorporating it into MEPS' established and ongoing MPC data collection activities. The ultimate objective is to create a database that will not only contain information on individuals' characteristics and health care use and expenditures, but also information on the providers they use.

The purpose of this paper is to describe how this newly initiated medical organization survey is being integrated into the existing MEPS-MPC.

2. Background

2.1 MEPS-HC

The MEPS-HC collects data from individual households and their members. These households are a nationally representative subsample of households that participated in the prior year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The annual sample size of the MEPS is approximately 14,000 households. The data collected can be analyzed at the person-, event-, and family-levels. Details regarding the MEPS sample design and the construction of analytic weights can be found in Cohen (2000), Ezzati-Rice (2008), and Machlin (2010).

Data from the household are collected for a period of two years and includes information on health care use, expenditures, sources of payment, health insurance coverage, health status, demographic and socio-economic characteristics, employment, access to care, and satisfaction with health care by means of a series of personal interviews. The interviewer also requests the names and contact information of any medical providers seen by members of the household. Household members seen by a medical provider are asked to provide signed permission forms authorizing AHRQ to contact each identified provider. Providers for whom a signed authorization form is obtained are eligible for the MEPS-MPC survey. In addition, for each individual family member it is ascertained whether there is a particular doctor's office, clinic, health center, or other place that the individual usually goes to if he/she is sick or needs advice about his/her health. That is, it is determined whether or not they have a usual source of care (USC).

2.2 MEPS-MPC

The MEPS-MPC is a targeted sample of the medical providers (physicians, hospitals, home health agencies, and pharmacies) who provided medical care to MEPS-HC respondents during the survey reference period. Again, only providers for whom a signed authorization form is obtained from the household are eligible for the MEPS-MPC survey. AHRQ attempts to contact all hospitals (including any separately billing doctors), home health care agencies, institutions and pharmacies. Among office-based medical doctors a sample is selected for follow-back.

The primary purpose of the MEPS-MPC is to compensate for household nonresponse on expenditure items. In this capacity, it only provides supplemental information to the MEPS-HC and is intended solely for editing and imputation purposes. It is not designed as a stand-alone survey to yield national estimates.

Data in the MEPS-MPC are collected at the event/encounter level and include dates of visits/services, diagnosis and procedure codes (ICD-9, CPT-4), and charges and payments.

2.3 MEPS-MOS

The purpose of the MEPS-MOS is to collect supplemental information on usual source of care (USC) practice characteristics for MEPS sample persons who saw their USC during the year. These data will be appended to the MEPS person-level data and included on MEPS public use files, subject to confidentiality review. MOS data determined to be confidential will be made available to researchers in the AHRQ and Census Research Data Centers.

The release of data that are nationally representative of persons with a usual source of care that they visited at least once during the year should enhance the scope of person-level analyses. It is anticipated that these data will support studies of the association between practice characteristics and consumer access, service use, expenditures, and quality of care.

MEPS staff have designed a 23-item questionnaire to investigate various organizational aspects of medical practices including ownership, financial incentives, provider and patient mix, access to services, coordination and quality of care, and establishment use of electronic health records. Further, they have convened a panel of technical experts to help develop and assess a set of proposed research questions.

3. Methods

3.1 MEPS-MPC / MOS Sample Selection

The target population for the MEPS-MOS is all persons with an office-based medical doctor identified as their USC who saw that USC during the year. The target population excludes persons whose USC was in a hospital, institution, or home health agency.

The MEPS-MOS is being fielded as an extension of the MEPS-MPC. Hence, USC providers will be identified and selected to receive the MOS questionnaire as part of the sample selection process already established for the MEPS-MPC. All usual source of care providers selected to be administered the MOS are also in the sample contacted for the MPC survey. Sampled non-USC providers will only be administered the MPC survey.

The MPC-MOS sampling frame is derived from a file compiled during the processing of the MEPS-HC data and is provided to AHRQ by the household survey contractor. This file is a person-provider pair level file and contains records for all MPC event types (e.g., hospital, office-based visits, institutions, home health agencies). This discussion focuses only on the selection of office-based event types because the MEPS-MOS is focused on the patient-USC relationship in physician office-based settings.

3.1.1 General Sampling Strategy

The objective of the MEPS Provider Component is to select person provider pairs where the MPC data are likely to be more complete, accurate, and informative than the data collected in the MEPS-HC. To help achieve this, two subpopulations are identified where household reported expenditure data are likely to be partially or fully missing, highly variable, or characterized by particularly high cost medical events: 1) the Medicaid population and 2) a group of people labeled as ‘targeted.’ Medicaid recipients frequently have missing expenditures data because State agencies pay medical providers directly for the care provided, so recipients are not aware of the true payment amounts. The ‘targeted’ group is identified as a group with more variable and likely higher expenditures. This group is comprised of persons who were in an institution for at least part of the year, who died during the year, or who had at least one of the following events: hospital stay, home health visit, surgery in an office setting, surgery in a hospital outpatient setting. These subpopulation classifications are used to create four sampling strata: (1) Medicaid—Targeted [M-T], (2) Medicaid—Non-Targeted [M-nT], (3) Non-Medicaid—Targeted [nM-T], and (4) Non-Medicaid—Non-Targeted [nM-nT].

Differential sampling is performed across each of the four strata. Historically, the MPC sampling process has selected all Medicaid—Targeted pairs with certainty; with the belief that this group benefits the most from MPC data collection. The optimal sample size for each of the remaining strata has then been determined using Neyman Allocation. The Neyman method allocates the remaining total sample (i.e., after certainty selection) among the remaining strata in proportion to stratum size and stratum variance (i.e., the variance of average expenditures per person). The purpose of the Neyman Allocation is to maximize survey precision given a fixed sample size; it assumes the costs of sampling within each stratum are equal. The same strategy is now being used to select the MPC-MOS cases, incorporating the USC relationship among pairs. Table 1 shows the basic counts of records eligible to be sampled.

Table 1. Basic MPC-MOS sampling frame counts, stratum by pair USC status

Sampling Strata	USC pair		Totals
	No	Yes	
M-T	n_{11}	n_{12}	n_{1+}
M-nT	n_{21}	n_{22}	n_{2+}
nM-T	n_{31}	n_{32}	n_{3+}
nM-nT	n_{41}	n_{42}	n_{4+}
Totals	n_{+1}	n_{+2}	n_{++}

Note: Certainty selections are highlighted in green.

Following the previous years’ protocols all Medicaid—Targeted pairs (i.e., M-T [n_{1+}]) are selected with certainty to be administered the MPC survey. To maximize the sample for the MOS project, all USC pairs (i.e., USC pair=Yes [n_{+2}]) are also selected with certainty and will be administered both the MPC survey and the MOS questionnaire.

Because the USC pairs are now also being selected with certainty they need to be accounted for when optimizing the sample sizes for each remaining stratum (i.e., M-nT, nM-T, nM-nT). Table 2 shows the basic optimal sample sizes after accounting for the certainty selections of the Medicaid—Targeted and USC pairs. The Neyman allocation can become an iterative task. For example, if the sample size in a non-certainty stratum based on Neyman allocation is less than the count of USC pairs in that stratum the sample size requirement is already met and no further sample is selected from non-USC pairs in

that stratum. Any such strata become certainty, and the remaining sample size is then reallocated by reapplying Neyman allocation across the remaining strata.

Table 2. Sample sizes to be selected based on Neyman Allocation

Sampling Strata	Non-USC	USC	Observed Marginal Totals	Neyman Allocation Sample Size	Remaining Sample to be Selected from the Non-USC
M-T	n_{11}	n_{12}	n_{1+}	n_{1+} (fixed)	0
M-nT	n_{21}	n_{22}	n_{2+}	$n_{2+[NA]}$	$n_{2+[NA]} - n_{22}$
nM-T	n_{31}	n_{32}	n_{3+}	$n_{2+[NA]}$	$n_{3+[NA]} - n_{32}$
nM-nT	n_{41}	n_{42}	n_{4+}	$n_{2+[NA]}$	$n_{4+[NA]} - n_{42}$
Totals	n_{+1}	n_{+2}	n_{++}	Targeted sample size	

Note: Final sample sizes to be selected are highlighted in green.

3.1.2 Applied Sampling Strategy

As with past MEPS-MPC samples the 2015 MEPS-MPC/MOS sample was selected and released for fielding in three waves. The general sampling strategy described above was applied for each of the three sample selections. For the 2015 MEPS-MPC a total of 18,000 pairs were released for fielding; 9,972 for the MEPS-MOS. It's anticipated that the 9,972 pairs for the MOS will equate to approximately 6,000 unique medical practices. Sample totals for the 2015 MEPS-MPC/MOS are provided in Table 3.

Table 3. Frame size and actual sample counts for 2015

Sampling Strata	Frame (N)	Number of sampled pairs		
		MPC & MOS	MPC Only	Total
M-T	2,650	592	2,058	2,650
M-nT	9,017	3,399	0	3,399
nM-T	6,573	996	3,213	4,209
nM-nT	19,224	4,985	2,757	7,742
Total	37,464	9,972	8,028	18,000

3.2 MEPS-MOS Weight Development

Unlike the MPC, the design of the MOS is intended to yield nationally representative estimates. The intent is to produce estimates of USC practice characteristics for a nationally representative sample of individuals who received care from their USC during the year. This requires the construction of an MOS analytic weight.

3.2.1 General Weighting Strategy

The objective is to assign an analytic weight to all MOS sample persons for which an MOS response is obtained. The MEPS-HC full-year person weight will serve as the base weight from which to develop the MOS weight. To ensure that the MOS yields nationally representative estimates for those receiving care from their USC, a two staged non-response adjustment will be used.

The first stage uses a weighting-class method (Lohr, 1999) to adjust the MOS base weight (i.e., MEPS-HC full-year person weight) for *lack of permission to contact the provider*:

Stage 1: Adjust for lack of permission to contact provider

$$W_{MOS0} = W_0 \times \hat{\phi}_C$$

W_0 = MEPS-HC full-year person weight

$\hat{\phi}_C$ = adjustment factor for respective weight adjustment classes

$$= \frac{\sum W_{0[s]} + \sum W_{0[ns]}}{\sum W_{0[s]}} \quad s \text{ and } ns \text{ are signers and non-signers}$$

Person-level characteristics of MEPS respondents (e.g., demographics, health status, language, etc.) across which there are differential propensities to sign the authorization form are used to specify the weight adjustment classes.

The second stage uses the same weighting-class method to adjust the resulting weight from Stage 1 for *practice survey non-response*:

Stage 2: Adjust for practice survey non-response

$$W_{MOS1} = W_{MOS0} \times \hat{\phi}_C$$

W_{MOS0} = adjusted weight from Stage 1

$\hat{\phi}_C$ = adjustment factor for respective weight adjustment classes

$$= \frac{\sum W_{MOS0[r]} + \sum W_{MOS0[nr]}}{\sum W_{MOS0[r]}} \quad r \text{ and } nr \text{ are responding and non-responding practices}$$

The Stage 2 weight adjustment classes will be determined using practice-level characteristics. This is somewhat more challenging than identifying the weighting-classes in Stage 1 because practice-level characteristics are not as readily available; hence the need for an MOS-type of survey. MEPS staff are exploring alternative sources of provider/practice level data to be used for such purposes. To the extent that practice level data is not available, person-level characteristics will be assessed for the Stage 2 adjustment.

Following the Stage 2 adjustment a raking procedure will be performed to align the MOS sample to MEPS population control totals.

3.2.2 Applied Weighting Strategy

To test this weighting strategy, we used a 2014 MOS test file to implement Stage 1 of the weight development strategy. Cross-tabulations and logistic regression models were used to evaluate the MOS eligible respondents' propensity to sign the authorization form. Among adults, differential non-signing rates were observed across age, race, and perceived health status: 18-64 year olds (21%) vs. 65+ years (17%); Asian/Native

Hawaiian/Pacific Islanders (30%) vs. Others (19%); Fair/Poor (17%) vs. Good/Very Good/Excellent (21%). Among children, differential non-signing rates were observed across age and race: 0-11 year olds (17%) vs. 12-14 year olds (23%) vs. 15-17 year olds (37%); Asian/Native Hawaiian/Pacific Islanders (29%) vs. Others (20%). The Stage 1 weight adjustment was applied separately for adults and children using weight adjustment classes defined by the respective variables above. Table 4 shows the Stage 1 adjustment calculations for adults:

Table 4. Stage 1 weighting-class adjustment for adults aged 18 or more years

Age	Race/Ethnicity		Perceived Health	
			Fair/Poor	G/VG/Excellent
18-64 yrs.	Asian/NH/PI	Total	71	296
		Signed	52	194
		Not-signed	19	102
		$\Sigma W_{0[s]}$	338,312	1,442,442
		$\Sigma W_{0[ns]}$	140,518	763,226
		$\hat{\phi}_C$	1.42	1.53
	Other	Total	1,394	2,829
		Signed	1,173	2,205
		Not-signed	221	624
		$\Sigma W_{0[s]}$	10,036,883	25,595,229
		$\Sigma W_{0[ns]}$	1,987,114	7,430,833
		$\hat{\phi}_C$	1.20	1.29
65+ yrs.	Asian/NH/PI	Total	61	87
		Signed	42	70
		Not-signed	19	17
		$\Sigma W_{0[s]}$	239,890	533,914
		$\Sigma W_{0[ns]}$	75,576	111,304
		$\hat{\phi}_C$	1.32	1.21
	Other	Total	570	864
		Signed	468	735
		Not-signed	102	129
		$\Sigma W_{0[s]}$	4,578,341	9,815,829
		$\Sigma W_{0[ns]}$	1,227,766	1,858,222
		$\hat{\phi}_C$	1.27	1.19

This Stage 1 approach will be implemented using the fielded 2015 MOS sample data once key components become available in 2017.

MEPS staff are considering two potential sources of provider characteristic data for the Stage 2 adjustment: the National Provider Identifier (NPI) [<https://npiregistry.cms.hhs.gov>] database and the SK&A Physician database (<http://www.skainfo.com/databases#physicians>). Assessing the match potential of these data sources with the MEPS and the quality of information they contain, particularly for weighting-class construction, is a work in progress. If the match rates are low and/or the provider characteristic data are non-informative, person-level characteristics will be used for the provider non-response adjustment.

4. Summary

The MEPS-MOS pilot study has been initiated to capture information on medical providers, link that information to patient-level information of health care use and expenditures, and make this information widely available through the release of public use data files. This effort is expected to enhance patient-level analyses and aid in evaluating the relationship between provider organization characteristics and individuals' health care use and expenditures. The effectiveness of the design and estimation strategy will be evaluated for future implementations. The MOS survey data will be available in Spring 2017. The full year 2015 use and expenditure person level public use file is scheduled for release in the Fall of 2017.

Acknowledgements

Funding for the MEPS-MOS is provided by the Robert Wood Johnson Foundation. The results and conclusions expressed here are those of the authors, and do not necessarily represent the views of the Department of Health and Human Services, the Agency for Healthcare Research and Quality, or the Robert Wood Johnson Foundation.

References

Cohen, S.B., *Methodology Report #11: Sample Design of the 1997 Medical Expenditure Panel Survey Household Component*. November 2000. Agency for Healthcare Research and Quality, Rockville, MD.

https://meps.ahrq.gov/data_files/publications/mr11/mr11.pdf

Ezzati-Rice, T.M., Rohde, F., Greenblatt, J., *Sample Design of the Medical Expenditure Panel Survey Household Component, 1998-2007*. Methodology Report No. 22. March 2008. Agency for Healthcare Research and Quality, Rockville, MD.

http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr22/mr22.pdf

Lake, T., Marincic, J., Valenzano, J., and Aggarwal, R. 2012. Feasibility of Collecting Data on Physicians and Their Practices: Environmental Scan Report. Mathematica Policy Research, Washington, DC.

Lohr, S., 1999. *Sampling: Design and Analysis*. Pacific Grove, CA: Brooks/Cole.

Machlin S.R., Chowdhury S.R., Ezzati-Rice T., DiGaetano R., Goksel H., Wun L.-M., Yu W., Kashihara D. Estimation Procedures for the Medical Expenditure Panel Survey Household Component. Methodology Report #24. September 2010. Agency for Healthcare Research and Quality, Rockville, MD.

http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr24/mr24.pdf