

Deriving a Person-Level Weight for Analyzing MEPS Supplemental Data from a Linked Medical Organization Survey

Marc Zodet¹, Sadeq Chowdhury¹,

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 includes a medical provider component (MEPS-MPC) that is designed to collect information from the health care service providers reported by the household. To address the increased demand for data on organizational characteristics of providers and/or health care practices, the Robert Wood Johnson Foundation sponsored a pilot Medical Organization Survey (MOS). This survey is an extension of the MEPS-MPC and collects this type of data from a subset of MEPS sample medical providers. The MOS pilot was first fielded in 2016 and the data were made available through the release of a public use data file (PUF) in spring 2017. This paper discusses the construction of the analytic weights included on the MOS PUF.

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

1. Introduction

Discourse on health care reform continues to dominate the American political landscape. With any discussion on changing the health care delivery system there is particular interest in how changes may impact the relationship between patients and provider organizations, specifically in terms of access to care, utilization of services, and overall healthcare spending. It is yet to be determined whether or not the trend toward more consolidated provider networks (e.g., group practices, networks of practices, hospitals, hospital-physician joint ventures, and other health care groups) will 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). The Robert Wood Johnson Foundation provided the Agency for Healthcare Research and Quality (AHRQ) with funding to conduct a new survey of health care providers called the Medical Organization Survey (MOS). This new survey yields the first nationally representative database linking provider characteristics to the characteristics of their patients.

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 has incorporated the MOS into MEPS’ established and ongoing MPC data collection activities. The resulting MOS database not only contains 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 the development of MOS analytic weights that can be used to produce nationally representative estimates based on a sample of individuals in MEPS who received care from their office-based usual source of care (USC) during the year.

2. Background

2.1 MEPS-HC

The MEPS-HC collects data from sample households on their members. These households are a nationally representative subsample of respondents to 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 covering two years and includes information on health care use, expenditures, payment sources, health insurance coverage, health status, demographic and socio-economic characteristics, employment, access to care, and satisfaction with health care through a series of computer-assisted 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.

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 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 are available as a MEPS public use file (PUF) and can be merged onto the MEPS person-level 2015 full-year consolidated file.

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 enhances the scope of person-level analyses by enabling studies of the association between practice characteristics and consumer access, service use, expenditures, and quality of care.

The MOS survey instrument is a 23-item questionnaire designed 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. MEPS staff worked in conjunction with 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 reported as a hospital, institution, or home health agency.

The MEPS-MOS is fielded as an extension of the MEPS-MPC. Hence, USC providers are identified and selected to receive the MOS questionnaire as part of the sample selection process 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.

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 only collects data on the patient-USC relationship in physician office-based settings.

3.1.1 General Sampling Strategy

The MEPS-HC sampling strategy selects person-provider pairs where the MEPS-HC data are more likely to be partially/fully missing, highly variable, or characterized by particularly high cost. The MPC sampling process has historically utilized a differential sampling procedure performed across four strata. The process selects with certainty

Medicaid pairs also characterized as more likely to have more variable and higher expenditures (e.g., Medicaid beneficiaries with one or more of the following: hospital stay, home health visit, surgery in office setting, surgery in hospital outpatient setting, or who died during the year or were in an institution for part of the year). The optimal sample size for each of the remaining strata is then determined using Neyman Allocation. 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 was used to select the MPC-MOS cases, incorporating the USC relationship among pairs. To maximize the sample for the MOS project, all USC pairs were also selected with certainty and are administered both the MPC survey and the MOS questionnaire.

Implementation of the sampling strategy is outside the scope of this report. For a more detailed discussion see Zodet (2016).

3.1.2 Sample Selection Results

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 across the 3 waves were released for fielding; 9,972 of these for the MEPS-MOS. Sample totals for the 2015 MEPS-MPC/MOS are provided in Table 1.

Table 1. Frame size and actual sample counts for 2015

	Frame (N)	Number of sampled pairs		
		MPC & MOS	MPC Only	Total
Total	37,464	9,972	8,028	18,000

During the fielding process the contractor identified any sampled person-provider USC pairs that were not eligible for the MOS for any of the following reasons: 1) the usual source of care setting is something other than an office-based physician practice, 2) the pair is a duplicate from a prior wave or 3) the person was not eligible to receive an analytic weight for the 2015 full year file (see section 3.2 below). Exclusion of these ineligible cases resulted in the sample size for the 2015 MOS being reduced from 9,972 to 9,252.

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 a MOS analytic weight.

3.2.1 General Weighting Strategy

An analytic weight was assigned to all MOS sample persons for which an MOS response was obtained. The MEPS-HC full-year person weight was used as the base weight from which to develop the MOS weight. A two-step non-response adjustment was used to

ensure that the MOS yields nationally representative estimates for those receiving care from their USC.

The first step used 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*:

Step 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

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

The second step used the same weighting-class method to adjust the resulting weight from Step 1 for *practice survey non-response*:

Step 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

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

Weighting adjustment classes are defined by characteristics across which there are differential propensities for a given outcome (e.g., propensity to sign an authorization form, propensity to respond). Each step utilized both person- and provider-level information in the forming of adjustment classes. Person-level characteristics were extracted from the MEPS-HC. Practice-level characteristics were obtained from the SK&A Physician database (<http://www.skainfo.com/databases#physicians>) using a hierarchical matching algorithm. The matching protocol for this algorithm was as follows:

1. Exact matching on provider level National Provider Identifier (NPI)
2. Exact phone matching on records unmatched after 1.
3. Bi-gram name and address matching on records unmatched after 1. and 2.

Any situation in 1. and 2. where a MEPS record exact matched to two or more SK&A records (duplicate matches), the bigram matching algorithm used name

and address to adjudicate the best match among the duplicate records.

The match rate among persons with signed authorization forms was significantly higher (85%) than among those without signed authorization (51%).

Variables highly correlated with a particular adjustment (i.e., signing status, non-response) or that might be associated with variance in key MOS items being collected were identified as initial candidate variables for defining weighting adjustment classes (Table 2).

Table 2. Candidate variables by source

MEPS-HC	SK&A
Age	Number of doctors at the site
Race/Ethnicity	Site specialty
Education	Number of unique doctors in the medical group
Interview language	Health system ownership
Perceived physical health	Accepts Medicaid
Perceived mental health	Use of electronic medical record (EMR) software

For the Step 1 weighting adjustment (i.e., non-signing of authorization form) candidate variables were evaluated separately for each of four strata: 1) children with SK&A data, 2) children without SK&A data, 3) adults with SK&A data, and 4) adults without SK&A data. Stratification by child/adult was implemented because some of the candidate variables were more readily defined specific to children and adults. Multiple logistic regression models were used to evaluate the associations between the signature status and the candidate variables. The same evaluation process was carried out for the Step 2 weighting adjustment though with only two strata. Logistic regression models were used to assess the relationship of the candidate variables to MOS survey response among those pairs with and without SK&A data.

In selecting the final set of variables to cross and the appropriate levels for each variable when forming adjustment cells we balanced statistical significance (from logistic models) and achieving minimum cell size considerations. Tables 3 and 4 present the variables and value levels that were used for the final weight adjustment classes for Steps 1 and 2 respectively.

3.2.2 Applied Weighting Strategy

The purpose of the MOS is to describe the practice characteristics of usual source of care providers for persons in the U.S. civilian, noninstitutionalized population who identified as having an office-based physician as a usual source of care and who visited them during the year. In the 2015 MEPS data, 10,766 persons were identified as meeting these criteria. Of these 10,776 observations, 1,514 (14%) were not eligible for fielding due to lack of a signed permission form. Of the 9,252 fielded cases, 2,091 (23%) were lost due to provider nonresponse. The weighting adjustment strategy presented above was used to create a preliminary analytic weight which summed to 101,159,262. The analytic sample size, weight and target population estimates presented here are preliminary and will be

updated in a future file that will contain revised weights reflecting adjustments for poverty status as well as additional provider links to sample persons.

Table 3. Final weight adjustment classes Step 1

Children with SK&A	Children without SK&A	Adults with SK&A	Adults without SK&A
Age <i>0-11 years</i> <i>12-17 years</i>	Age <i>0-11 years</i> <i>12-17 years</i>	Age <i>18-64 years</i> <i>65+ years</i>	Age <i>18-64 years</i> <i>65+ years</i>
Race/Ethnicity <i>Hispanic</i> <i>White</i> <i>Black</i> <i>Asian/Other</i>	Race/Ethnicity <i>Hispanic</i> <i>White</i> <i>Black</i> <i>Asian/Other</i>	Race/Ethnicity <i>Hispanic</i> <i>White</i> <i>Black</i> <i>Asian/Other</i>	Race/Ethnicity <i>Hispanic</i> <i>White</i> <i>Black</i> <i>Asian/Other</i>
Use of EMR software <i>Yes</i> <i>No/Not specified</i>	Region <i>Northeast</i> <i>Midwest</i> <i>South</i> <i>West</i>	Accepts Medicaid <i>Yes</i> <i>No</i>	Region <i>Northeast</i> <i>Midwest</i> <i>South</i> <i>West</i>
Number of unique doctors in medical group <i>Non-group</i> <i>/ Not specified</i> <i><10</i> <i>10-24</i> <i>25-49</i> <i>50-149</i> <i>150-199</i> <i>200+</i>		Number of doctors at the site <i><10</i> <i>10+</i>	
Health system owns site <i>Yes</i> <i>No/Not specified</i>		Site specialty <i>Family practice</i> <i>/ Pediatrics</i> <i>Internal medicine</i> <i>Multispecialty</i> <i>Other</i>	

*Data were stratified by child/adult because some of the sociodemographic variables evaluated as possible adjustment cell candidates were specific to children/adults.

Table 4. Final weight adjustment classes Step 2

With SK&A		Without SK&A
Age	Site specialty	Age
<i>0-17 years</i>	<i>Family practice</i>	<i>0-17 years</i>
<i>18+ years</i>	<i>/ Pediatrics</i>	<i>18+ years</i>
	<i>Internal medicine</i>	
Region	<i>Multispecialty</i>	Race/Ethnicity
<i>Northeast</i>	<i>Other</i>	<i>Hispanic</i>
<i>Midwest / South</i>		<i>White, NH</i>
<i>West</i>		<i>Black, NH</i>
		<i>Asian/Other, NH</i>
Health system owns site	Number of doctors at the site	Region
<i>Yes</i>	<i><10</i>	<i>Northeast</i>
<i>No/Not specified</i>	<i>10+</i>	<i>Midwest / South</i>
		<i>West</i>
Use of EMR software		
<i>Yes</i>		
<i>No/Not specified</i>		
Number of unique doctors in medical group		
<i>Non-group</i>		
<i>/ Not specified</i>		
<i><10</i>		
<i>10-24</i>		
<i>25-49</i>		
<i>50-149</i>		
<i>150-199</i>		
<i>200+</i>		

*Data were not stratified by child/adult, but still retained age which place children and adults in different adjustment classes.

A summary of the complete adjustment procedure is presented in Table 5. Application of the weighting class adjustment among the 9,252 with signed authorization forms using the class cells shown in Table 3 adjusted the base weight (PERWT15P) for the observed 14% non-signing rate. The resulting Step 1 weight then summed to the desired target population of just over 101 million. The Step 1 weight adjustment factor ranged between 1.0 and 2.3 with an average of 1.175.

Building off of the Step 1 weight, the adjustment cell classes shown in Table 4 were used to perform the Step 2 adjustment among the pairs with a responding provider. This second step adjusted the Step 1 weight among responding pairs from approximately 79 million back up to the target of 101,159,262. The adjustment factor for Step 2 ranged from 1 to 2.4 with an average of 1.278. The Step 2 weight then underwent a final raking procedure.

Table 5. Summary of the Weight Adjustment Procedure

Step 1 (adjust base weight for lack of permission to contact provider)

Patient-USCP pairs identified in the MEPS-HC:

1,514	Not eligible, No AF	(14%)		
<u>9,252</u>	Signed AF	(86%)	$\Sigma_{\text{PERWT15P}} =$	85,403,004
10,766	Total		$\Sigma_{\text{PERWT15P}} =$	101,159,262

After Step 1 weight adjustment . . .

	N	Min	Max	Sum	Mean
Adjustment factor	9,252	1	2.3	10,871	1.175
Step 1 weight	9,252	774	97,350	101,159,262	10,934

Step 2 (adjust base weight for provider non-response)

Patient-USCP pairs fielded for the MOS (adjusted frame):

2,091	Non-responders	(23%)		
<u>7,161</u>	Responders	(77%)	$\Sigma_{\text{Step1 weight}} =$	79,075,104
9,252	Total		$\Sigma_{\text{Step1 weight}} =$	101,159,262

After Step 1 weight adjustment . . .

	N	Min	Max	Sum	Mean
Adjustment factor	7,161	1	2.4	9,148	1.278
Step 1 weight	7,161	1,236	127,172	101,159,262	14,126

3.2.3 Raking

Because there is no external source to provide a benchmark for this study population, the raking procedure utilized internal control totals based on the MEPS-HC. The marginal totals for the raking procedure were selected because they are often of analytic interest to MEPS data users and are also used for the raking procedures associated with the full-year weight development. The raking dimensions are presented in Table 6.

Table 6. Marginal Dimensions for final raking procedure

Age (0-17, 18-64, 65+ years) x	Region (Northeast, Midwest, South, and West)
Age (0-17, 18-44, 45-64, 65+ years) x	Race/Ethnicity [Hispanic, White (NH), Black (NH), Asian/Other (NH)]
Age (0-9, 10-17, 18-44, 45-64, 65+ years) x	Sex (Male, Female)
Age (<18, 18+ years) x	Race/Ethnicity [Hispanic, White (NH), Black (NH), Asian/Other (NH)]

The preliminary MOS weight, MOSWT15P, is non-zero (range: 1,217-120,064) for all 7,161 persons (person-provider pairs) included on the 2015 MOS public use file.

4. Summary

The MEPS-MOS pilot study is a supplemental survey built upon the existing MEPS-MPC. The prime objective of the MOS project is 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. Using common survey weighting adjustment methods and both internal and external data sources, an analytic weight was developed that accounts for both household respondent non-authorization and provider non-response. The effectiveness of the design and estimation strategy will be evaluated for future implementations. An initial preliminary weight was released with the 2015 MOS PUF in the spring of 2017. A final poverty adjusted weight is currently in development and is scheduled to be released mid-November 2017. This file will contain revised weights reflecting adjustments for poverty status as well as additional provider links to sample persons.

Acknowledgements

Funding for the MEPS-MOS was 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

Zodet, M., Chowdhury, S., Machlin, S., and Cohen, J. 2016. Linked designs of the MEPS Medical Provider and Organization Surveys. In JSM Proceedings, Survey Research Methods Section. Alexandria, VA: American Statistical Association. 1914-1921.