

Precision of Survey Estimates Derived from the Medical Expenditure Panel Survey (MEPS)

William Yu, Agency for Healthcare Research and Quality
540 Gaither Road, Suite 5054, Rockville, MD 20850-6649

KEY WORDS: Design Effect, MEPS, Health Care Expenditures

Introduction

The sample design of the Medical Expenditure Panel Survey Household Component (MEPS-HC) is characterized by a complex multi-stage area probability design that includes disproportionate sampling of specified policy relevant population groups. As a consequence of departures from simple random sampling assumptions, the variances of survey estimates derived from the MEPS will generally exhibit design effects that are substantially greater than unity. A summary of design effect variations for 1996-1998 MEPS estimates has been previously reported (Yu, W., 2002). Based on data from the 1999 and 2000 MEPS-HC, this paper will evaluate and contrast the design effects achieved for national estimates of health care utilization, expenditures, and sources of payment; the level of design effect variation in related survey estimates; and design effect variation by alternative population subgroups and by different geographic regions of the nation. This analysis will also include an evaluation of design effects achieved for health care estimates associated with individuals that incur high levels of medical expenditures. The results will help improve the sample design specifications for the selection of future new sample panels of households for the annual MEPS-HC.

MEPS Household Component

MEPS is conducted annually to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. It is co-sponsored by the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS).

The core survey for MEPS is the Household Component (HC). The MEPS-HC collects data through an overlapping panel design. In this design, data are collected through a series of five rounds of

interviews over a period of 2 ½ years. Interviews are conducted with one member of each family, who reports on the health care experiences of the entire family. Two calendar years of medical expenditure and utilization data are collected in each household and captured using computer-assisted personal interviews. This series of data collection rounds is launched again each subsequent year on a new sample of households to provide overlapping samples of survey data that provide continuous and current estimates of health care expenditures (Cohen J., 1997).

The sampling frame for the MEPS-HC is drawn from respondents to the previous year's National Health Interview Survey (NHIS), conducted by NCHS. NHIS provides a nationally representative sample of the U.S. civilian noninstitutionalized population, with over sampling of Hispanics and blacks.

Source and Definition of Data

This study is based on two years of MEPS-HC data: HC-038 (1999) and HC-050 (2000). Expenditures in MEPS are defined as the sum of direct payments for health care provided during the year, including out-of-pocket payments and payments by private insurance, Medicare, Medicaid, and other sources. Payments for over the counter drugs, alternative care services, and phone contacts with medical providers are not included in MEPS total expenditure estimates. Indirect payments unrelated to specific medical events such as Medicaid Disproportionate Share and Medicare Direct Medical Education subsidies also are not included (Cohen JW, Machlin SR, Zuvekas SH, *et al.*, 2000).

The use and expenditure data included in this paper were derived from the MEPS-HC and Medical Provider Components (MPC). MPC data were collected for some office-based visits to physicians (or medical providers supervised by physicians), hospital-based events (e.g. inpatient stays, emergency room visits, and outpatient department visits), and prescribed medicines. HC data were collected for physician visits, dental and vision services, other medical equipment and services, and home health care not provided by an agency. Data on expenditures for care provided by home health agencies were collected only in the MPC. MPC data were used if complete; otherwise HC data were used if complete. Missing data for events where HC data were not complete and MPC data were not

The views expressed in this paper are those of the author and no official endorsement by the Department of Health and Human Services or the Agency for Healthcare Research and Quality is intended or should be inferred. The author wishes to thank Trena Ezzati-Rice and Doris C. Lefkowitz for their helpful reviews of the paper.

Collected or not complete were derived through an imputation process (Cohen S. and Carlson B., 1994).

Only persons with positive person weight are included in this study. The sample sizes for the target populations were 23,565 and 23,839 for calendar years 1999 and 2000, respectively. Table 1a and 1b contain utilization, expenditure, and sources of payment variables selected from the 1999-2000 MEPS consolidated data files for this study.

Table 1a - Utilization and Expenditure Variables by Health Services Category: MEPS

Health Service Category	Utilization Expenditure	
	Variables	Variables
All Health Services	-	TOTEXPy
<u>Office Based Visits</u>		
Total Office Based Visits (Physician + Non-physician + Unknown)	OBTOTVyy	OBVEXPy
<u>Hospital Outpatient Visits</u>		
Total Outpatient Visits (Physician + Non-physician + Unknown)	OPTOTVyy	OPEXPyy ⁽¹⁾
Total Emergency Room Visits	ERTOTyy	EREXPy ⁽¹⁾
Total Inpatient Stays (Including Zero Night Stays)	IPDISyy	IPEXPyy ⁽¹⁾
Total Dental Visits	DVTOTyy	DVTEXPy
Total Home Health Care	HHTOTDyy	HHEXPyy ⁽¹⁾
<u>Other</u>		
Prescription Medicines	RXTOTyy	RXEXPy
Vision Aids	-	VISEXPy

Table 1b - Sources of Payment Variables: MEPS

Sources of Payment	Variables
Total Amount Paid by Self/Family	TOTSLFyy
Total Amount Paid by Medicare	TOTMCRyy
Total Amount Paid by Medicaid	TOTMCDyy
Total Amount Paid by Private Insurance	TOTPRVyy
Total Amount Paid by Others	TOTOTHyy

yy-applicable year of data (e.g., 99 or 00)

⁽¹⁾ Sum of facility and "separately billing doctor"(SBD) expense

Source: Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality

Variables used to form population, geographic, and economic subgroups in this analysis are gender (male, female), age (<=17, 18-44, 45-64, 65+), race/ethnicity (Hispanic, black/non-Hispanic, others), Census region (Northeast, Midwest, South, West), family income as percent of poverty line (0%-199%, 200%-400%, 400%+), and insurance coverage (public/private, uninsured) .

Design Effect in the MEPS-HC

Given the complex nature of the MEPS-HC survey design, the assumptions of independence and equal selection probabilities are not satisfied. The resulting impact on variance estimation may be described as follows (Cohen, S., 1982):

$$\sigma^2_{\text{complex}} = \sigma^2_{\text{SRS}} [1 + \rho (\bar{n} - 1)]$$

where

$\sigma^2_{\text{complex}}$ is the true variance of a statistic given the complex survey design,

σ^2_{SRS} is the variance estimate obtained for the statistic under simple random sampling assumptions,

ρ is the intra cluster correlation coefficient, and

\bar{n} is the average cluster size.

The design effect consequently is expressed as:

$$\text{Design Effect} = (\sigma^2_{\text{complex}} / \sigma^2_{\text{SRS}}) = [1 + \rho (\bar{n} - 1)]$$

For a given variable, the design effect deviates from unity when the effects of clustering are dominant in a survey design and/or the average cluster size is moderate to large. Variances of all estimated parameters presented in this paper were derived using SUDAAN (Shah, Barnwell, Bieler, *et al.*, 1996) with the Taylor series linearization method to account for the complex survey design.

Evaluation of Design Effect Variation

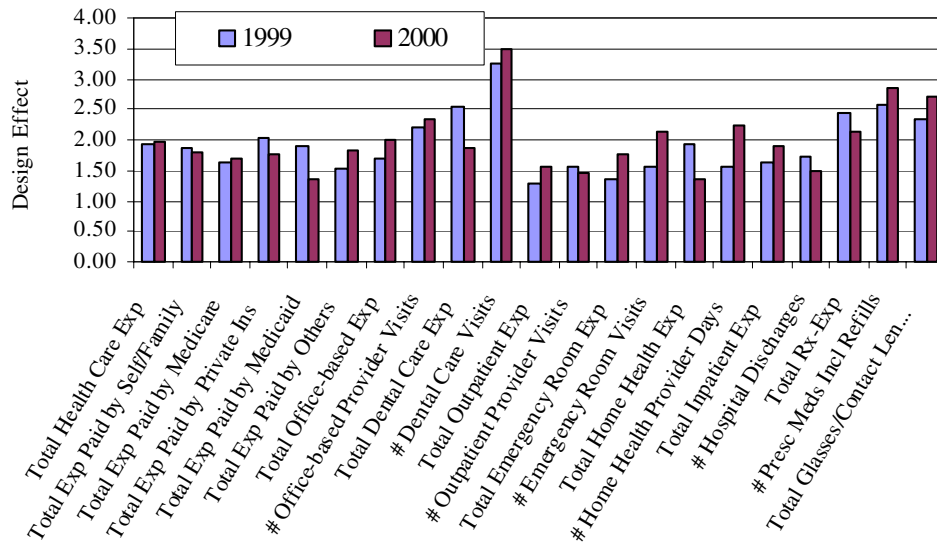
For each of the two years covered in this study, design effects were determined for each of the 21 selected survey statistics for the nation overall as well as for alternative population subgroups and different geographic regions of the nation. Figure 1 compares the level of design effects achieved for the selected set of national mean health care use, expenditure, and sources of payment estimates in 1999 and 2000.

The overall design effects are summarized further in Table 2 below with the lowest and highest design effects noted in { }. For example, in 2000, the design effects ranged from 1.36 for the estimated total home health expenditures to 3.49 for the estimated number of dental care visits with an overall average of 1.99.

Table 2 - Summary of Design Effect Variation: MEPS 1999-2000

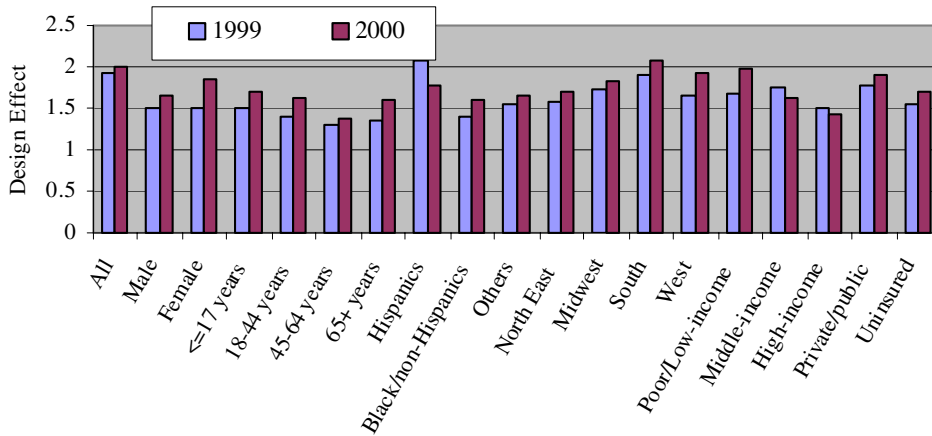
Year	Lowest Estimate {variable}	Highest Estimate {variable}	Mean (Std)
1999	1.30 {total outpatient expenditures}	3.27{# dental care visits}	1.93 (0.49)
2000	1.36 {total home health expenditures}	3.49 {# dental care visits}	1.99 (0.52)

Figure 1 - Design Effects for Mean Use, Expenditure, and Source of Payment Estimates (1999-2000 MEPS)



Source: Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality

Figure 2 - Average Design Effects for Population and Geographic Groups (1999 - 2000 MEPS)



Source: Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality

Figure 2 and Table 3 present comparisons of average design effects from the selected health care use, expenditure, and sources of payment measures across all of the alternative population groups and by different geographic regions of the nation from 1999 to 2000. There are no significant differences in the average design effects for the selected estimates by population subgroups or geographic regions between 1999 and 2000.

Table 3 - Summary of Design Effects by Population Groups: MEPS 1999-2000

Population Subgroups	Mean(Std)*	
	1999	2000
All	1.93 (0.49)	1.99 (0.52)
<u>Gender</u>		
Male	1.53 (0.22)	1.64 (0.33)
Female	1.76 (0.33)	1.85 (0.35)
<u>Age Group</u>		
<=17 years	1.75 (0.48)	1.69 (0.53)
18-44 years	1.56 (0.27)	1.62 (0.30)
45-64 years	1.46 (0.31)	1.38 (0.38)
65+ years	1.51 (0.33)	1.59 (0.33)
<u>Race/Ethnicity</u>		
Hispanics	2.08 (0.50)	1.77 (0.57)
Black/non-Hispanics	1.26 (0.23)	1.60 (0.48)
Others	1.64 (0.38)	1.64 (0.39)
<u>Census Region</u>		
North East	2.05 (0.59)	1.71 (0.51)
Midwest	1.54 (0.53)	1.83 (0.58)
South	1.80 (0.51)	2.08 (0.80)
West	2.24 (0.99)	1.92 (1.02)
<u>Poverty Level</u>		
100% - 200%	1.67 (0.30)	1.97 (0.52)
200% - 400%	1.75 (0.43)	1.63 (0.40)
400%+	1.49 (0.39)	1.42 (0.35)
<u>Insurance Coverage</u>		
Private/public	1.77 (0.45)	1.90 (0.48)
Uninsured	1.54 (0.38)	1.69 (0.38)

*Mean and standard deviation of design effects for the 21 selected estimates listed in Table 1a and 1b.

Source: Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality

There is a notable downward trend for the values of average design effect by ascending age groups 0-17, 18-44, and 45-64 in both years. For race/ethnicity, the black/non-Hispanic group has the lowest and second lowest average design effects for 1999 and 2000, respectively. For insurance coverage, the uninsured group has lower average design effect than private/public for both years. There is no notable pattern of difference in the average design effects for

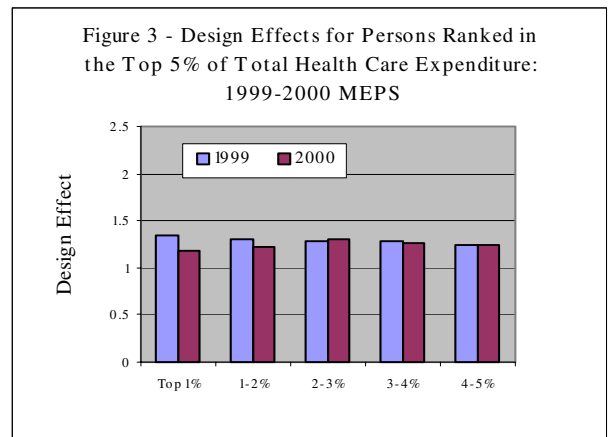
people classified in each of the census regions or income group.

Table 4 shows that the top 5% of the population (ranked by health care expenditure) expends approximately 52% of the total health care expenses by the civilian noninstitutionalized population. Figure 3 presents a comparison of design effects achieved for health care estimates associated with individuals that incur high levels of medical expenditures. For persons ranked in the top 5% by health care expenditures, the average design effects for the selected survey estimates are 1.40 and 1.33 for 1999 and 2000, respectively. This average design effect does not vary significantly by 1% increments within the top 5% group.

Table 4 - Top 5% Healthcare Expenditures Distribution: 1999-2000 MEPS

Percent of Population Ranked by Health Care Expenditures	1999 Health Care Expenditures (%)	2000 Health Care Expenditures (%)
Top 1%	25.4	24.5
Top 2%	36.2	34.6
Top 3%	42.9	41.6
Top 4%	48.1	46.9
Top 5%	52.3	51.5

Source: Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality



Source: Center for Financing, Access, and Cost Trends, Agency for Healthcare Research and Quality

Summary

A part of the overall precision requirements for the 1999 and 2000 MEPS-HC was the achievement of an average design effect of ~2.0 for survey estimates of policy relevant population subgroups (e.g., households with Hispanics and blacks, persons with family incomes less than 200% of poverty line, persons 65 years or older). The study findings confirmed that this requirement was generally satisfied with respect to these subgroups and for selected measures of health care utilization, expenditures, and sources of payment for the two years of data studied.

For both 1999 and 2000, “# of dental care visits” has the highest design effect estimate. One possible explanation is that the ultimate cluster units in the MEPS HC sample design are the household or family. It is expected that a strong positive correlation exists between individuals in the same household with respect to their number of dental care visits.

There are no significant differences in the average design effects for the selected survey estimates by population subgroups and geographic regions between 1999 and 2000. For individuals with the highest total health care expenditures (ranked in the top 5% of the population), the average design effect for the selected survey estimates is about 1.3. This design effect does not vary significantly by 1% increments within the top 5% group.

In summary, design effects are useful for estimating the sample size needed for a survey and for estimating the precision of estimates achieved in the survey.

References

- Yu WW. (2002), “Design Effect Variation for Estimates Derived from the MEPS(1996-1998).” 2002 Proceedings of the American Statistical Association, Section on Survey Research Methods [CD-ROM], Alexandria VA: American Statistical Association.
- Cohen JW., “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 JW, Machlin SR, Zuvekas SH, *et al.*, “Health care expenses in the United States, 1996.” Rockville (MD): Agency for Healthcare Research and Quality; 2000. MEPS Research Findings 12. AHRQ Pub. No. 01-0009.
- Cohen S. and Carlson B., “A comparison of household and medical provider reported expenditures in the 1987 NMES.” *Journal of Official Statistics* 1994; 10(1):3-29.
- Cohen SB., “Comparison of Design Effect and Relative Variance Curve Strategy for Variance Estimation from Complex Survey Data,” presented at the annual meetings of the American Public Health Association, 1982.
- Shah, B.V., B.G. Barnwell, G.S. Bieler, K.E. Boyle, R.E. Folsom, L. Lavange, S.C. Wheelless, and R. Williams, Technical Manual: Statistical Methods and Algorithms Used in SUDAAN Release 7.0, Research Triangle Park, NC: Research Triangle Institute, 1996.