

An Assessment of the Impact of Two Distinct Survey Design Modifications on Health Care Utilization Estimates in the Medical Expenditure Panel Survey

Cohen, Steven

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

540 Gaither Road

Rockville, Maryland 20850, USA

E-mail: scohen@ahrq.gov

1. Background: There are several factors that contribute to the variation in estimates of the health care utilization related measures across respective national surveys. These factors include survey content and questionnaire design, survey design and methods, sample design, and post-survey data processing and estimation techniques. Survey design issues such as mode of administration, number of interviews, length of recall period, and resultant response rates will also affect the accuracy and precision of associated survey estimates of health care utilization. Alternative methodologies employed for editing the survey data, imputation procedures and adjustments for survey nonresponse likewise can impact the final survey estimates that are generated. Another critical dimension that distinguishes estimates both within and across surveys is the duration of the time period or the interval that the survey estimates cover. Gaining a better understanding of how the respective surveys either converge or diverge in terms of these various components is critical to helping policymakers make the best use of resulting statistics.

The Medical Expenditure Panel Survey (MEPS) was designed to provide annual estimates at the national level of the health care utilization, expenditures, sources of payment and health insurance coverage for the U.S. civilian non-institutionalized population. The MEPS consists of a family of interrelated surveys, which include a Household Component (HC) and a Medical Provider Component (MPC). In addition to collecting data to yield annual estimates for a variety of measures related to health care use and expenditures, MEPS provides estimates of measures related to health status, demographic characteristics, employment and access to health care. Estimates can be provided for individuals, families and population subgroups of interest. The data collected in this ongoing longitudinal study also permit studies of the determinants of the use of services and expenditures, and changes in the provision of health care in relation to social and demographic factors such as employment or income; the health status and satisfaction with health care of individuals and families; and the health needs of specific population groups such as the elderly, children and minorities. The set of households selected annually for the MEPS HC is a subsample of those participating in the prior year's National Health Interview Survey (NHIS), an ongoing annual household survey of approximately 35,000 households (~85,000 individuals) conducted by the National Center for Health Statistics, Centers for Disease Control and Prevention. Initiated in 1996, the current MEPS annual survey consists of approximately 13,500 families and 32,000 individuals, and reflects an oversample of Hispanics, blacks and Asians. Data from two panels are combined to produce estimates for each calendar year.

2. Comparisons of MEPS Panel Specific Estimates of Health Care Utilization Estimates: In 2007, the survey experienced two dominant survey design modifications: (1) a new sample design attributable to the sample redesign of the NHIS in 2006, and (2) an upgrade to the CAPI platform for the survey instrument, moving from a DOS to a Windows based environment. These design changes were implemented starting with the MEPS Panel 12. Alternatively, the MEPS Panel 11, which was initiated in the previous year, carried forward the prior NHIS sample design and DOS based CAPI instrument when obtaining calendar year 2007 data (Year 2 data for panel 11). Consequently, the comparison of the national estimates of health care utilization for specific services derived separately for these two distinct MEPS panels providing calendar year data for 2007 provides a unique opportunity to assess the impact of the *joint effects* of these two recent

MEPS design modifications relative to the original MEPS design. This is partially achieved as a component of the survey quality control procedures, through which the national survey estimates derived from each of the MEPS panels are compared, to determine whether the national estimates are consistent across the two independent samples.

In each of the comparisons in this study, the panel that represents the subset of MEPS sample respondents that are participating in their 2nd year of the survey, is characterized by a response rate that is between 3-5% lower than the panel based on the year one respondents. Several analyses of specific types of MEPS estimates have been conducted across panels to assess the impact of survey attrition. Measures that have been the focus of these assessments include health insurance coverage status, medical expenditures, the concentration of health care expenditures, the prevalence of specific chronic diseases (e.g. heart disease, mental disorders, hypertension, osteoarthritis and diabetes) and the annual mean health care expenditures for individuals with chronic diseases. Study findings have revealed no systematic evidence of nonresponse bias attributable to survey attrition affecting the national estimates of these criterion measures derived from the older panel, when contrasted with the new panel estimates. These findings were confirmatory in terms of the soundness of the nonresponse adjustments employed to adjust for potential nonresponse bias attributable to survey attrition associated with the second year of the survey. Viewed in this light, the comparisons of use estimates in this study across panels serve to assess the impact of the survey design changes implemented in 2007 on resultant MEPS survey estimates of a variety of health care utilization estimates. Conditioned on the same set of primary sampling units (e.g. counties, groups of contiguous counties), and sample sub-areas (segments) that define the samples each year, the NHIS/MEPS sample of households are independently selected. The standard errors of the survey estimates derived from each of the panels have been adjusted for the impact of clustering due to the multistage survey design, and the test statistics used to test for equivalence in estimates have also been adjusted to control for the shared set of geographic areas that define the samples across the panels.

The first set analyses focus on national health care utilization estimates derived from the MEPS for the following health care services: ambulatory visits (office-based visits and outpatient facility visits); in-patient stays, ER visits, dental visits and prescribed medicine purchases. These utilization estimates are produced for the overall population, and further subset by age classification (0-17, 18-64, 65+) for a given calendar year. The table includes comparable overall and panel specific estimates of these health care utilization measures for the years 2002-2006, to allow for comparisons of the stability of these estimates over time, for periods when both panels of the survey shared the same design features, versus the new design in 2007. The MEPS estimation weight used to derive the full year estimates was the estimation weight released on the full year health care expenditures public use file. This person-level estimation weight was post-stratified and raked to national estimates of the population control totals obtained from the Current Population Survey for the given year on the following measures: Census region (Northeast, Midwest, South, and West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, Black but non-Hispanic, Asian, and other); sex, age and poverty status. The panel specific MEPS utilization estimates are based on panel specific weights which were individually raked to national control totals. In addition, the MEPS estimation weights for producing national estimates from the combined samples across panels includes further post-stratification adjustments and raking adjustments to the pooled panel specific estimation weights.

2.1 Utilization estimates expressed as population totals: When comparing the calendar year health care utilization estimates of the total number of office based visits; outpatient visits; in-patient stays, ER visits, and dental visits for the overall population across panels for years 2002-2006, no statistically significant differentials in estimates were detected at the .05 level of

significance. Comparable results were obtained when examining the alignment in the panel specific estimates for the population under age 18, in addition to those for ages 18-64, and those age 65 and older. These findings are indicative of a strong degree of convergence in estimates across panels prior to the recent MEPS redesign.

When focusing attention on the comparison of these annual utilization measures for 2007, a significant differential in the overall population estimates of the total number of office based visits was detected across MEPS panels. This differential in population estimates of the annual number of office based visits also held for individuals ages 18-64, again with a lower office based use totals associated with the MEPS panel that employed the redesign. Other differentials in the annual utilization estimates expressed as totals across MEPS panels were noted in the 2007 for estimates of inpatient stays (ages 0-17 and 65+) and emergency room visits (0-17 years), though here, the estimates obtained from the panel with the survey redesign were higher. For the comparisons of the annual estimates of the total number prescribed medicine purchases across panels for years 2002-2006, some significant differentials were occasionally detected, either for specific age groups or for the overall U.S. civilian population. These differentials, when detected, were consistently characterized by higher totals derived from the more current panel. Alternatively, no significant differentials in these national estimates of total number of prescribed medicine purchases were detected in 2007, the year of the MEPS sample re-design.

2.2 Utilization estimates expressed as population means, conditional on some use: Conditioned on the population with at least one visit (or prescribed medicine purchase) in a given year, the next set of comparisons focused on the mean number of health care utilization estimates for the same measures across panels for years 2002-2006. Across a set of 120 comparisons by panel, 13 significant differentials were noted, with each of the utilization measures exhibiting at least one cross panel differential. Regarding the direction of the differences in panel estimates, when detected, higher mean estimates of prescribed medicine purchases and lower mean estimates of ambulatory office visits were derived from the newer panel. These findings are more suggestive of occasional departures from convergence in estimates in MEPS across panels prior to the recent MEPS redesign.

When focusing attention on the comparison of these annual mean utilization measures for 2007, however, a more pronounced pattern of significant differentials in the panel specific estimates was observed. With one exception, when differentials were detected, the estimates generated by the panel with the new MEPS survey redesign in 2007 were consistently lower. There was general equivalence with respect to the factors that characterize survey attrition in MEPS in 2007 and prior years, and comparable weight adjustment strategies were applied to correct for survey nonresponse each year over the past decade. Consequently, the differentials detected in the panel specific health care utilization estimates for 2007 were suggestive of the presence of an effect attributable to the new MEPS survey design modifications. However, this analysis did not permit a determination of which design modification, CAPI upgrade or NHIS sample redesign, was the dominant factor. Additional benefits accrue to surveys such as MEPS that have adopted overlapping panel designs. In addition to affording the capacity to assess the impact of design changes on resultant survey estimates, the design allows for the “pooling” of the panel specific estimates, which facilitates an “averaging” of the design modifications in the national estimates produced from such surveys. This is the approach taken in MEPS in the derivation of the calendar year estimates for 2007.

3. Model Based Test for MEPS Survey Design Modifications on Calendar Year Estimates of Health Care Utilization: The comparisons of the stability of national care health care utilization estimates, subject to the MEPS survey design modifications initiated in 2007, were supplemented

by a model based analysis that included additional controls for other pre-dispositional factors. More specifically, a multivariate analysis was conducted to discern the influence of the MEPS survey design modifications on predicting health care utilization patterns for the entire year after controlling for socio-demographic, geographic, economic and health related factors associated with utilization behavior. Separate analyses were conducted for each of the following health care services: ambulatory visits; in-patient stays, ER visits, dental visits and prescribed medicine purchases. Building on previous research efforts that have identified salient factors associated with health care utilization and related medical care expenditures, a regression model was developed to consist of the subset of significant predictors that distinguished the uninsured from those with either public or private coverage .

Based on prior studies that identified factors determined to be significant correlates in distinguishing between individuals with limited or no health care use for a given service from their counterparts with more intensive utilization of services in a calendar year, a detailed set of pre-dispositional factors were given consideration in this study. The measures under consideration included demographic, socio-economic and geographical characteristics, health insurance coverage, health status and specific health conditions, accidental events, health care utilization indices and total health care spending. Once these measures were controlled for in the multivariate regression model, it was possible to determine whether an individual's classification with respect to MEPS Panel, which varied significantly in terms of the CAPI platform and by sample design, influenced the prediction of the level of intensity in the utilization of specific health care services in a calendar year.

Under the assumption that the two distinct MEPS panels that are combined to produce annual survey estimates were characterized by the same survey design, one would not expect to observe a significant Panel effect. In that vein, a comparable model-based analysis testing for panel effects in the 2004-2006 MEPS when both panels shared the same survey design features, served as a test of the impact of survey attrition on health care utilization estimates for specific healthcare services. As in prior evaluations and consistent with the findings from the bivariate analyses, the results of these multivariate regression analyses revealed few systematic significant effects for MEPS Panel classification in distinguishing the level of intensity in health care service utilization patterns for ambulatory visits; in-patient stays, ER visits, and dental visits. For the years 2004-2006, other than an observed panel effect for prescribed medicine purchases with higher use profiles usually associated with the newer panel, the findings provided limited evidence of nonresponse bias attributable to survey attrition affecting the national utilization estimates derived from the older panel, when contrasted with the new panel estimates.

Given the design modifications to the MEPS CAPI interview and sample design introduced in 2007, a test for a MEPS Panel effect affords the opportunity to assess the joint influence of these survey design modifications on health care utilization estimates in a modeling context. Here, the results of the multivariate regression analysis were mixed. No significant effect for MEPS Panel classification was detected in distinguishing the level of health care service utilization for annual estimates of outpatient visits, ER visits, inpatient stays or prescribed medicine purchases, when testing at an alpha level of .05. These results further reinforce the efficacy of the survey design and estimation strategies adopted in the MEPS to limit the impact of recent design modifications on specific health care utilization estimates and related model based studies. Alternatively, a MEPS Panel effect was detected for the model based analyses of annual office based visits and dental visits, with higher estimates associated with the older panel.

4. *Evaluation of MEPS sample redesign on national utilization estimates:* In addition to the within MEPS studies, the linkage of the MEPS to the NHIS permits a related set of analyses to be

conducted to discern the impact of the MEPS sample redesign initiated in 2007 and associated survey attrition on national estimates. The MEPS and NHIS linked design permits appending the health care utilization data profiles collected in the NHIS for the prior year to the responding MEPS sample, for a respective calendar year. Using the NHIS data in concert with the restricted sample of MEPS respondents permits the derivation of NHIS national health care utilization estimates for the prior year based on a subsample characterized by a lower response rate and a modified MEPS sample redesign. With this design feature, the NHIS national estimates derived from the MEPS sample may be compared to the national estimates obtained from the full NHIS, prior to its linkage to MEPS. Since the NHIS data that are being analyzed are not the variables collected in MEPS, they are not affected by the other 2007 design modification to the MEPS CAPI instrument. Consequently, this targeted analysis serves to permit a direct assessment of the effect of sample design modifications and adjustments for survey attrition on utilization estimates obtained from the MEPS independent of the CAPI design modifications. When viewed in this light, the findings from these analyses reveal the capacity and degree of specific MEPS sample design and nonresponse adjustment strategies to yield comparable national estimates of the NHIS health care utilization measures under investigation that align with the estimates produced off the full NHIS for the prior year.

To facilitate this analysis, the following NHIS measures of health care utilization and access were selected in support of these analyses: during the past 12 months, have you been hospitalized overnight?; how many different times did you stay in any hospital overnight or longer?; how many nights were you in the hospital? did you receive care from doctors or other health care professionals 10 or more times?; have you delayed seeking medical care because of worry about the cost?; and was there any time when you needed medical care, but did not get it because you/the family couldn't afford it? We initially examined the level of convergence in health care utilization related estimates based on the 2004-2006 MEPS, with survey design features that pre-date the modifications to the CAPI and the NHIS sample redesign. As can be observed from a review of the comparisons of the MEPS-based prior year NHIS utilization related estimates and the full NHIS generated estimates, only a negligible number of significant differences in estimates (other than for missing values) are evident, when testing at an alpha level of .05. A comparison of the NHIS derived and the MEPS derived utilization estimates for the total population and for age groupings <18 and 18-64 revealed similar levels of convergence. The results held for each of the NHIS measures of health care utilization and access to care under consideration, whether derived from the linked MEPS Round 1 sample or the MEPS full year sample which has a lower response rate due to greater sample attrition over time. Similar findings were observed in related studies of the capacity of the MEPS subsample to reproduce prior year NHIS health insurance coverage estimates. Taken together, the results present no evidence of nonresponse bias attributable to survey attrition affecting these national health care utilization and access estimates when subject to the more restrictive response rate experience in MEPS.

The targeted analyses in this section served to isolate the effect of MEPS sample design modifications and adjustments for survey attrition on health care utilization related estimates independent of the CAPI design modifications introduced in 2007. A synthesis of these findings provides evidence in support of the continued capacity of the MEPS survey after experiencing sample redesign modifications to reproduce national estimates for many of the NHIS health care utilization related measures based on the complete NHIS sample for the prior year. Furthermore, the limited set of departures observed in 2007 for the MEPS derived NHIS utilization estimates serve to identify measures that may further enhance the impact of poststratification adjustments to the MEPS estimation weights with corresponding gains in accuracy.

5. *Summary:* The first arm of the study examined the alignment in MEPS health care utilization estimates across panels, controlling for design features. Based on the findings from the panel specific comparisons of the calendar year MEPS utilization estimates expressed as both population means and population totals, there was some evidence of differentials in estimates attributable to the joint effects of the 2007 MEPS survey redesign. With some exceptions, when differentials were detected, the estimates generated by the new panel experiencing the MEPS survey redesign in 2007 were consistently lower. Alternatively, when focusing attention on comparing the calendar year health care utilization related estimates expressed as totals across panels for years prior to the MEPS survey redesign (2002-2006), other than for prescribed medicines no statistically significant differentials in estimates were detected both overall and when distinguished separately by age for children and adults. By pooling the two panel specific estimates in MEPS, any extant effects attributable to the survey redesign are mitigated.

These descriptive analyses were supplemented by a model-based analysis of the impact of recent MEPS design modifications on healthcare utilization estimates, controlling for pre-dispositional factors associated with healthcare use. When testing for the joint influence of 2007 MEPS survey design modifications on health care utilization estimates, the results of the regression analyses varied by the type of health care service measure under consideration. While no significant effect for MEPS Panel classification was detected in distinguishing the level of health care service utilization for annual estimates of outpatient visits, ER visits, inpatient stays or prescribed medicine purchases, a panel effect was detected for the predictions of annual office based visits and dental visits. When significant differentials were operational in 2007, higher model based utilization estimates were associated with the older panel. The results provided important data to illustrate the level of impact of the recent MEPS design modifications have had on resultant calendar year health care utilization estimates and related model based studies.

The final series of analyses attempted to isolate the effects of MEPS sample design modifications and adjustments for survey attrition on utilization estimates from those attributable to the CAPI design modifications introduced in 2007. The survey integration between the MEPS and the NHIS facilitated this type of investigation. For each of the individual MEPS Panels operational in 2007, national estimates of prior year NHIS health care utilization related measures were derived from the MEPS first part of year and annual responding samples, and compared with those obtained from the full NHIS in 2006. A review of the results of the MEPS and NHIS generated estimates based on the same NHIS measures revealed only modest differences in estimates. Taken in concert, these findings are indicative of the level of stability in utilization related estimates attributable to sample design modifications and the effectiveness of the MEPS nonresponse adjustments..

To the extent the source(s) of the observed differences in estimates can be attributed to specific survey design differentials, estimation strategies could be developed to bridge the redesign-based estimates with those of the original design for analyses of trends over time. In this study, a set of options are presented for consideration when attempting to aligning redesign-based estimates with the original design to enhance analyses of trends over time. One of these approaches employs direct standardization technique, where the overall estimated utilization totals derived from the new panel are “aligned” to converge with the national utilization estimates derived from the old panel via adjustments to the survey estimation weights. The analysis of the impact this bridging strategy on the overall MEPS estimates of utilization totals by event type for calendar year 2007 revealed a modest, but non-significant increase in the total number of office based provider visits.

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 is intended or should be inferred.