

Estimation Procedures in the 1996
Medical Expenditure Panel Survey
Household Component

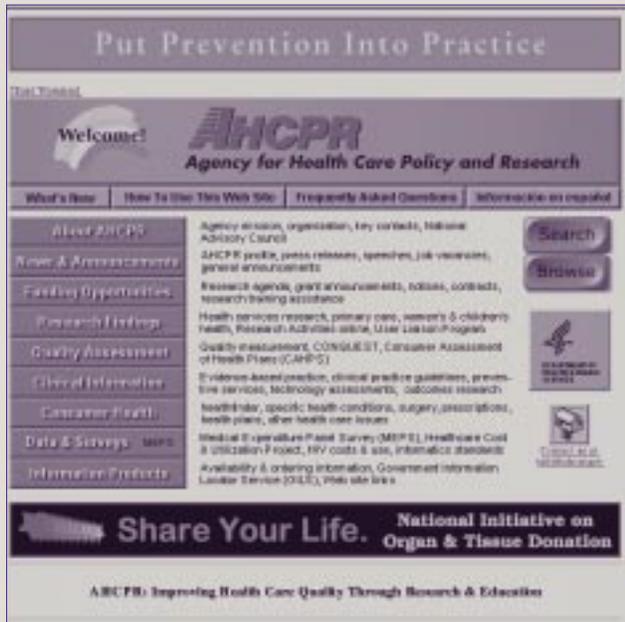
Methodology

Report 5



U.S. Department of Health and Human Services
Public Health Service
Agency for Health Care Policy and Research

Health Care Information and Electronic Ordering Through the AHCPR Web Site



The Agency for Health Care Policy and Research's Web site—<http://www.ahcpr.gov/>—makes practical, science-based health care information available in one convenient place.

Buttons correspond to major categories of Web site information, including funding opportunities, research findings, quality assessments, clinical information, consumer health, and data and surveys.

The Web site features an Electronic Catalog to the more than 450 information products generated by AHCPR, with information on how to obtain these resources. Many information products have an electronic ordering form and are mailed free of charge from the AHCPR Clearinghouse within 5 working days.

<http://www.ahcpr.gov/>

Abstract

The Medical Expenditure Panel Survey (MEPS) is the third in a series of nationally representative surveys of medical care use and expenditures sponsored by the Agency for Health Care Policy and Research (AHCPR). MEPS comprises four component surveys. The Household Component (HC) produces national and regional estimates of the health care utilization, expenditures, sources of payment, and insurance coverage of the U.S. civilian noninstitutionalized population. The HC sample design is a stratified multistage area probability design with disproportionate sampling to facilitate the selection of an oversample of minorities. This report provides an overall summary of

HC sample yields across the three rounds of data collection that cover calendar year 1996. It also provides an overview of the weighting strategies used to obtain national estimates of health care parameters for the population. Survey design complexities that require special consideration for variance estimation and analysis are discussed.

Suggested citation

Cohen SB, DiGaetano R, Goksel H. Estimation procedures in the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1999. MEPS Methodology Report No. 5. AHCPR Pub. No. 99-0027.

Estimation Procedures in the 1996
Medical Expenditure Panel Survey
Household Component

Methodology

Report 5



U.S. Department of Health and Human Services
Public Health Service
Agency for Health Care Policy and Research
AHCPR Pub. No. 99-0027
May 1999

The Medical Expenditure Panel Survey (MEPS)

Background

The Medical Expenditure Panel Survey (MEPS) is conducted to provide nationally representative estimates of health care use, expenditures, sources of payment, and insurance coverage for the U.S. civilian noninstitutionalized population. MEPS also includes a nationally representative survey of nursing homes and their residents. MEPS is cosponsored by the Agency for Health Care Policy and Research (AHCPR) and the National Center for Health Statistics (NCHS).

MEPS comprises four component surveys: the Household Component (HC), the Medical Provider Component (MPC), the Insurance Component (IC), and the Nursing Home Component (NHC). The HC is the core survey, and it forms the basis for the MPC sample and part of the IC sample. The separate NHC sample supplements the other MEPS components. Together these surveys yield comprehensive data that provide national estimates of the level and distribution of health care use and expenditures, support health services research, and can be used to assess health care policy implications.

MEPS is the third in a series of national probability surveys conducted by AHCPR on the financing and use of medical care in the United States. The National Medical Care Expenditure Survey (NMCES) was conducted in 1977, the National Medical Expenditure Survey (NMES) in 1987. Beginning in 1996, MEPS continues this series with design enhancements and efficiencies that provide a more current data resource to capture the changing dynamics of the health care delivery and insurance system.

The design efficiencies incorporated into MEPS are in accordance with the Department of Health and Human Services (DHHS) Survey Integration Plan of June 1995, which focused on consolidating DHHS surveys, achieving cost efficiencies, reducing respondent burden, and enhancing analytical capacities. To accommodate these goals, new MEPS design features

include linkage with the National Health Interview Survey (NHIS), from which the sample for the MEPS HC is drawn, and enhanced longitudinal data collection for core survey components. The MEPS HC augments NHIS by selecting a sample of NHIS respondents, collecting additional data on their health care expenditures, and linking these data with additional information collected from the respondents' medical providers, employers, and insurance providers.

Household Component

The MEPS HC, a nationally representative survey of the U.S. civilian noninstitutionalized population, collects medical expenditure data at both the person and household levels. The HC collects detailed data on demographic characteristics, health conditions, health status, use of medical care services, charges and payments, access to care, satisfaction with care, health insurance coverage, income, and employment.

The HC uses an overlapping panel design in which data are collected through a preliminary contact followed by a series of five rounds of interviews over a 2½-year period. Using computer-assisted personal interviewing (CAPI) technology, data on medical expenditures and use for 2 calendar years are collected from each household. This series of data collection rounds is launched each subsequent year on a new sample of households to provide overlapping panels of survey data and, when combined with other ongoing panels, will provide continuous and current estimates of health care expenditures.

The sampling frame for the MEPS HC is drawn from respondents to NHIS, conducted by NCHS. NHIS provides a nationally representative sample of the U.S. civilian noninstitutionalized population, with oversampling of Hispanics and blacks.

Medical Provider Component

The MEPS MPC supplements and validates information on medical care events reported in the MEPS HC by contacting medical providers and pharmacies identified by household respondents. The MPC sample includes all hospitals, hospital physicians, home health agencies, and pharmacies reported in the

HC. Also included in the MPC are all office-based physicians:

- Providing care for HC respondents receiving Medicaid.
- Associated with a 75-percent sample of households receiving care through an HMO (health maintenance organization) or managed care plan.
- Associated with a 25-percent sample of the remaining households.

Data are collected on medical and financial characteristics of medical and pharmacy events reported by HC respondents, including:

- Diagnoses coded according to ICD-9 (9th Revision, International Classification of Diseases) and DSM-IV (Fourth Edition, *Diagnostic and Statistical Manual of Mental Disorders*).
- Physician procedure codes classified by CPT-4 (Current Procedural Terminology, Version 4).
- Inpatient stay codes classified by DRG (diagnosis-related group).
- Prescriptions coded by national drug code (NDC), medication names, strength, and quantity dispensed.
- Charges, payments, and the reasons for any difference between charges and payments.

The MPC is conducted through telephone interviews and mailed survey materials.

Insurance Component

The MEPS IC collects data on health insurance plans obtained through employers, unions, and other sources of private health insurance. Data obtained in the IC include the number and types of private insurance plans offered, benefits associated with these plans, premiums, contributions by employers and employees, and employer characteristics.

Establishments participating in the MEPS IC are selected through four sampling frames:

- A list of employers or other insurance providers identified by MEPS HC respondents who report having private health insurance at the Round 1 interview.
- A Bureau of the Census list frame of private-sector business establishments.
- The Census of Governments from the Bureau of the Census.

- An Internal Revenue Service list of the self-employed.

To provide an integrated picture of health insurance, data collected from the first sampling frame (employers and other insurance providers) are linked back to data provided by the MEPS HC respondents. Data from the other three sampling frames are collected to provide annual national and State estimates of the supply of private health insurance available to American workers and to evaluate policy issues pertaining to health insurance.

The MEPS IC is an annual panel survey. Data are collected from the selected organizations through a prescreening telephone interview, a mailed questionnaire, and a telephone followup for nonrespondents.

Nursing Home Component

The 1996 MEPS NHC was a survey of nursing homes and persons residing in or admitted to nursing homes at any time during calendar year 1996. The NHC gathered information on the demographic characteristics, residence history, health and functional status, use of services, use of prescription medications, and health care expenditures of nursing home residents. Nursing home administrators and designated staff also provided information on facility size, ownership, certification status, services provided, revenues and expenses, and other facility characteristics. Data on the income, assets, family relationships, and caregiving services for sampled nursing home residents were obtained from next-of-kin or other knowledgeable persons in the community.

The 1996 MEPS NHC sample was selected using a two-stage stratified probability design. In the first stage, facilities were selected; in the second stage, facility residents were sampled, selecting both persons in residence on January 1, 1996, and those admitted during the period January 1 through December 31.

The sampling frame for facilities was derived from the National Health Provider Inventory, which is updated periodically by NCHS. The MEPS NHC data were collected in person in three rounds of data collection over a 1½-year period using the CAPI system. Community data were collected by telephone using computer-assisted telephone interviewing (CATI) technology. At the end of three rounds of data collection,

the sample consisted of 815 responding facilities, 3,209 residents in the facility on January 1, and 2,690 eligible residents admitted during 1996.

Survey Management

MEPS data are collected under the authority of the Public Health Service Act. They are edited and published in accordance with the confidentiality provisions of this act and the Privacy Act. NCHS provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports and microdata files. Summary reports are released as printed documents and electronic files. Microdata files are released on CD-ROM and/or as electronic files.

Printed documents and CD-ROMs are available through the AHCPR Publications Clearinghouse. Write or call:

AHCPR Publications Clearinghouse
Attn: (publication number)
P.O. Box 8547
Silver Spring, MD 20907
800-358-9295
410-381-3150 (callers outside the United States only)
888-586-6340 (toll-free TDD service; hearing impaired only)

Be sure to specify the AHCPR number of the document or CD-ROM you are requesting. Selected electronic files are available through the Internet on the AHCPR Web site:

<http://www.ahcpr.gov/>

On the AHCPR Web site, under Data and Surveys, click the MEPS icon.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Cost and Financing Studies, Agency for Health Care Policy and Research, 2101 East Jefferson Street, Suite 500, Rockville, MD 20852 (301-594-3075).

Table of Contents

Introduction	1
MEPS Household Component Sample Design	1
Round 1 Field Results	4
Sampling Weight Specifications	6
Variance Estimation	15
Summary	15
References	16

Tables

1. Response rates for the 1996 MEPS Household Component, Round 1	5
2. Population estimates by sex and age for December 1996 and March 1997: United States.	14

Estimation Procedures in the 1996 Medical Expenditure Panel Survey Household Component

by Steven B. Cohen, Ph.D., Agency for Health Care Policy and Research, and Ralph DiGaetano, M.S., and Huseyin Goksel, M.S., Westat, Inc.

Introduction

The 1996 Medical Expenditure Panel Survey (MEPS) Household Component (HC) was designed to produce national and regional estimates of the health care utilization, expenditures, sources of payment, and insurance coverage of the U.S. civilian noninstitutionalized population for calendar year 1996. MEPS includes surveys of medical providers, employers, and other health insurance providers to supplement the data provided by household respondents. The design of MEPS permits both person-based and family-level estimates. Government agencies, legislative bodies, and health professionals need comprehensive national estimates in order to formulate and analyze national health policies. The scope and depth of this data collection effort address that need.

MEPS is the third in a series of national probability surveys conducted by the Agency for Health Care Policy and Research (AHCPR) on the financing and utilization of medical care in the United States. Previous surveys were the 1977 National Medical Care Expenditure Survey (NMCES) and the 1987 National Medical Expenditure Survey (NMES).

The MEPS HC is an ongoing annual survey, with each sample panel collecting data over a 30-month period to obtain information that covers two consecutive calendar years. MEPS collects data on the specific health services that Americans use, how frequently they use them, the cost of these services, and how they are paid for. MEPS also collects data on the cost, scope, and breadth of private health insurance held by and available to the U.S. population. MEPS is unparalleled for the degree of detail in its data, as well as its ability to link medical expenditures and health insurance data to survey respondents' demographic, employment, economic, health status, utilization of health services, and other characteristics. Moreover, MEPS is the only federally sponsored national survey that provides a foundation for estimating the impact of changes in sources of payment

and insurance coverage on different economic groups or special populations of interest, such as the poor, elderly, families, veterans, the uninsured, and racial and ethnic minorities.

The survey is cosponsored by AHCPR and the National Center for Health Statistics (NCHS). Westat, Inc., and the National Opinion Research Center (NORC) are the data collection organizations for the 1996 MEPS HC.

The MEPS HC sample design is a stratified multistage area probability design with disproportionate sampling to facilitate the selection of an oversample of minorities (Cohen SB, 1997a). This report provides an overall summary of sample yields across the three rounds of data collection that cover calendar year 1996. It also provides an overview of the weighting strategies adopted to obtain national estimates of health care parameters for the U.S. civilian noninstitutionalized population. In addition, survey design complexities that require special consideration for variance estimation and analysis are discussed.

MEPS Household Component Sample Design

The set of households selected for the 1996 MEPS is a subsample of those participating in the National Health Interview Survey (NHIS). NHIS is an ongoing annual household survey of approximately 42,000 households (109,000 individuals). NCHS conducts NHIS to obtain national estimates for the U.S. civilian noninstitutionalized population on health care utilization, health conditions, health status, insurance coverage, and access. In addition to the cost savings achieved by eliminating the need to list and screen households independently, selecting a subsample of NHIS participants for MEPS has resulted in survey data with enhanced analytical capacity. Use of the 1995 NHIS data in concert with the data collected for the 1996 MEPS provides the capacity for longitudinal analyses not

otherwise available. Furthermore, the greater number and dispersion of the primary sampling units (PSUs) that compose the MEPS national sample have resulted in improvements in precision over previous expenditure survey designs (Arnett, Hunter, Cohen, et al., 1996; Cohen SB, 1996).

The MEPS HC consists of an overlapping panel design in which each sample panel is interviewed a total of five times over 30 months to yield annual use and expenditure data for two calendar years. Design specifications for the 1996 MEPS required that the full series of interviews necessary to acquire information for calendar year 1996 be completed in approximately 9,000 households. The same panel of households was interviewed in person three times over the course of the survey to obtain data on members' health care experience for 1996 (Cohen JW, 1997).

The 1996 MEPS HC sample was selected from households that responded to the 1995 NHIS. More specifically, the 1996 MEPS HC sample was selected from a nationally representative NHIS subsample from two out of four NHIS panels to represent the Nation. It encompassed half of the participating households in the NHIS sample during the second and third quarters of 1995. NHIS was designed to permit nationally representative subsamples to be selected by restricting the sample to one of four distinct panels. Any combination of one to four panels will provide a nationally representative sample of households. Furthermore, each NHIS panel subsample for a given quarter of a calendar year is nationally representative.

The complete 1995 NHIS sample (panels 1-4) consists of 358 PSUs (i.e., counties or groups of contiguous counties) with a targeted sample of approximately 42,000 responding households. The sample PSUs selected for NHIS were stratified by geographic area (Census region and State), metropolitan statistical area (MSA), and sociodemographic measures (Judkins, Marker, and Waksberg, 1994). In order to oversample areas with high population concentrations of blacks and Hispanics, a sample of blocks (segments) was selected within sample PSUs after being stratified by measures of minority population density. A nationally representative sample of approximately 71,000 addresses within sampled blocks was selected and targeted for further screening to include an oversampling of households containing blacks and Hispanics as part of the 1995 NHIS interview.

The nationally representative 1995 NHIS subsample selected for the 1996 MEPS consists of 195 PSUs. In the two targeted quarters of 1995, these PSUs included 1,675 sample segments (second-stage sampling units) and 10,597 households responding to NHIS. This NHIS sample reflects an oversampling of households with Hispanics and blacks at the following approximate ratios of representation relative to the remaining households: Hispanics, 2.0:1.0; blacks, 1.5:1.0. The sample size for the 1996 MEPS was targeted at approximately 9,000 reporting units (generally, families or single persons) yielding the complete series of core interviews (i.e., Rounds 1-3) to obtain health care use and expenditure data for calendar year 1996.

Procedures for Data Collection

Five in-person interviews were conducted with each NHIS panel selected for the MEPS HC at 3- to 4-month intervals over an approximately 30-month field period. The first three of these rounds (Rounds 1-3) of the 1996 MEPS HC served to collect the main body of utilization and expenditure data for calendar year 1996. Rounds 3-5 for the 1996 MEPS panel cover 1997. Rounds 3-5 will be combined with Rounds 1-3 of the 1997 MEPS panel to yield the full sample base for the 1997 MEPS HC and be the source of annual estimates for that calendar year.

All interviews were conducted in person, using a computer-assisted personal interview (CAPI) as the principal data collection mode. Round 1 asks about the period from January 1 of the calendar year to the date of that interview; Round 2 asks about the time from the Round 1 interview through the date of the Round 2 interview; and Round 3 collects data from the date of the Round 2 interview in 1996 through the date of the Round 3 interview in 1997. Thus, Round 3 covers both calendar years, and the data are partitioned accordingly for estimation purposes.

Sampling Unit Definitions

The definitions of dwelling units and group quarters in the MEPS HC are generally consistent with the definitions used for NHIS. While the MEPS sample is a subsample of NHIS dwelling units (referred to as households), a reporting unit for MEPS data collection

purposes was defined as a person or group of persons in a sampled dwelling unit who are related by blood, marriage, adoption, or other family associations. All persons in a sampled reporting unit were to be interviewed at the same time. When unrelated persons lived in the same dwelling unit, sample households were split into multiple reporting units. Examples of the relationship between sample dwelling units and corresponding reporting units include the following:

- A married daughter and her husband living in the same dwelling with her parents are considered one reporting unit.
- A husband and wife and their unmarried daughter, age 18, who is living away from home at college constitute two reporting units operationally (but only one family unit analytically).
- Three unrelated persons living in the same dwelling unit represent three reporting units.

The first round of the 1996 MEPS averaged 1.09 reporting units per sample dwelling unit. This low average reflects the fact that a substantial proportion of sample dwelling units contained only one reporting unit.

In MEPS, analyses are conducted at both the individual and family levels. Through the re-enumeration section of the Round 1 questionnaire, the status of each individual sampled at the time of the NHIS interview is classified as “key or non-key” and “in-scope or out-of-scope.” The “key” and “scope” indicators, together, define the target sample to be used for person-level national estimates. They are discussed in detail next.

In-Scope Persons

Individuals are in scope if they are members of the civilian noninstitutionalized population. Because a person’s eligibility for the survey may have changed since the NHIS interview, sampling re-enumeration takes place in each subsequent interview for persons in all households selected into the core survey.

Key Persons

A “key” survey participant is one whose chance of selection for MEPS is linked to the sample of households originally selected for NHIS. (College

students interviewed in dorms for NHIS are not included in MEPS; college students living away from home are included in MEPS if, during MEPS enumeration, their parents identified the students’ usual place of residence as somewhere away from home.) A person must be key in order to be eligible to receive a person-level weight. (Other conditions must be met as well.)

Key survey participants include all civilian noninstitutionalized individuals who resided in households that responded to the nationally representative NHIS subsample reserved for MEPS. Members of the Armed Forces who are on full-time active duty and reside in responding NHIS households that include other family members who are civilian noninstitutionalized individuals are also considered key persons. However, they are considered out of scope for person-level estimates derived from the survey unless they re-enter the civilian noninstitutionalized population for some time during 1996.

Individuals who join the NHIS reporting units that define the 1996 MEPS HC sample (in Round 1 or later MEPS rounds) and did not have an opportunity for selection during the time of the NHIS interview are also considered key persons. These include newborns, individuals who were in an institution or outside the United States, and military personnel previously residing on military bases who join MEPS reporting units to live in the community.

College students under age 24 who were interviewed at dormitories in the 1995 NHIS were considered ineligible for the 1996 MEPS sample. If an unmarried college student under age 24 responded to the 1995 NHIS interview while living away at school (but not in a dormitory) and, during the MEPS Round 1 interview, was determined to be unmarried, under age 24, and a student with parents living elsewhere who resided at his/her current housing only during the school year, he or she was excluded from the sample. If, on the other hand, the person’s status at the time of the MEPS Round 1 interview was no longer that of an unmarried student under age 24 living away from home, then the person was retained in the 1996 MEPS sample as a key person.

At the time of the MEPS Round 1 interview with NHIS sample respondents, a question was asked to determine if any related college students under age 24

usually lived in the sampled household but were currently living away from home and going to school. These college students were considered key persons. They were identified and interviewed at their college address but were linked to the sampled household for family analyses. Some of these college students living away from home at the time of the Round 1 interview had been identified as living in a sampled household at the time of the 1995 NHIS interview. The remainder were identified at the time of the MEPS Round 1 interview.

When key in-scope MEPS participants move out (in Round 1 or later rounds) and join or create another family, data on all members who are related by blood, marriage, adoption, or foster care to the key in-scope MEPS participants are obtained from the time that the key in-scope person(s) joined the family. Similarly, data are collected (in Round 1 and later rounds) on all related persons who join families already participating in MEPS, whether the new persons are key or non-key.

If key in-scope MEPS participants enter a nursing home, thus leaving the civilian noninstitutionalized population, data are collected during their stay in the nursing home. Data are not collected for other key in-scope persons after they leave the civilian noninstitutionalized population of the United States. If they return to the U.S. civilian noninstitutionalized population, these persons are once again eligible for data collection in MEPS.

Non-Key Persons

Persons who were not living in the original sampled dwelling unit at the time of the 1995 NHIS interview but were part of the civilian noninstitutionalized population at that time are considered non-key. If such persons happen to be living in a MEPS sampled household in Round 1 or later rounds, MEPS data (e.g., utilization and income) will be collected for the period of time they are part of the sampled unit to permit family analyses. Non-key persons who leave a sample family without an accompanying key in-scope person will not be recontacted for subsequent interviews. Non-key individuals are not part of the target sample used to obtain person-level national estimates.

Eligible Persons

People are eligible for data collection in MEPS if they are key and in scope. Individuals who are non-key and in scope, as well as members of a family with at least one member who is key and in scope, are also eligible for data collection. Out-of-scope individuals who are full-time active-duty members of the Armed Forces are also eligible for MEPS data collection for the time period during which they are members of a family with at least one member who is key and in scope.

Round 1 Field Results

The 1995 NHIS subsample eligible for the 1996 MEPS consisted of 10,639 responding NHIS dwelling units. Of these, 10,597 (99.6 percent) had sufficient information to permit MEPS data collection. Table 1 summarizes response rates for MEPS (conditional on response to NHIS) at both the dwelling-unit and reporting-unit levels. The 10,509 sample dwelling units that had sufficient address information from NHIS and were considered eligible for MEPS contained a total of 11,424 reporting units. Of these reporting units, 83.1 percent responded to the first MEPS interview, 2.2 percent could not be located, and 14.7 percent were located but did not participate in the MEPS interview.

For a reporting unit to be eligible for MEPS data collection, it had to include at least one key individual who was a member of the civilian noninstitutionalized population between January 1, 1996, and the date of the MEPS interview. Within the 10,597 sampled dwelling units that constitute the MEPS Round 1 sample, 11,590 reporting units were identified and targeted for data collection. Of these, 166 reporting units were determined to be ineligible for the 1996 MEPS. MEPS sample ineligibility for Round 1 was a consequence of the following situations:

- All members of the reporting unit died prior to January 1, 1996 (21 cases).
- All members of the reporting unit were full-time active-duty members of the military prior to January 1, 1996 (4 cases).
- All members of the reporting unit were institutionalized prior to January 1, 1996 (26 cases).

Table 1. Response rates for the 1996 MEPS Household Component, Round 1

Item	Dwelling units		Reporting units	
	Number	Percent	Number	Percent
Sample cases	10,597	—	11,590	—
Sample eligibles ¹	10,509	100.0	11,424	100.0
Respondents ²	8,793	83.7	9,488	83.1
Nonrespondents	1,716	16.3	1,936	16.9
Unable to locate	—	—	251	2.2
Nonparticipants	—	—	1,685	14.7

¹88 sample dwelling units and 166 reporting units were deemed ineligible for MEPS. To be eligible for MEPS, a unit must contain at least 1 individual from an NHIS sample household who was also a member of the civilian noninstitutionalized population between January 1, 1996, and the date of the MEPS interview. Most of the ineligible dwelling units consisted of individuals who died or were institutionalized prior to January 1, 1996.

²A dwelling unit is classified as a respondent if at least 1 member of the reporting unit responded to the survey. A reporting unit is classified as a responding unit if all members responded.

Note: MEPS is the Medical Expenditure Panel Survey. NHIS is the National Health Interview Survey. MEPS response rates are conditional on NHIS response.

Source: Center for Cost and Financing Studies, Agency for Health Care Policy and Research: Medical Expenditure Panel Survey Household Component, 1996 (Round 1).

- All members of the reporting unit left the United States prior to January 1, 1996 (29 cases).
- All members of the reporting unit were ineligible for data collection for two or more of the above reasons (e.g., one member died prior to January 1, 1996, and the other was institutionalized prior to January 1, 1996) (86 cases).

Of the 11,424 eligible reporting units targeted for interviews in Round 1, 9,488 (83.1 percent) responded to the first MEPS interview (Table 1). The remaining 1,936 eligible reporting units (16.9 percent) were classified as nonrespondents because of:

- Refusal to complete the interview (1,506 cases).
- Unavailable during the field period (43 cases).
- Unable to locate (251 cases).

- Illness (27 cases).
- Other nonresponse (109 cases).

Since the MEPS sample is a nationally representative subsample of households that were part of the 1995 NHIS, the response rate that has implications for the development of national estimates from MEPS is a function of the response rates to both surveys. Specifically, the overall Round 1 MEPS response rate is the product of the following three components:

- The NHIS response rate achieved for households eligible for MEPS (93.9 percent).
- The proportion of NHIS units selected that had sufficient information to permit inclusion in MEPS (99.6 percent).

- The conditional response rate for MEPS Round 1 reporting units (83.1 percent).

The combination of these factors resulted in a response rate of 77.7 percent ($.939 \times .996 \times .831$) for the 1996 MEPS Round 1 HC (Cohen and Machlin, 1998).

In Table 1, the conditional response rates for the 1996 MEPS Round 1 survey are shown at both the dwelling-unit level and the reporting-unit level. Since there is generally a one-to-one correspondence between a dwelling unit and a reporting unit, the conditional response rates for both are similar (83.7 and 83.1 percent, respectively). While the reporting-unit response rate is more meaningful from an operational perspective, the dwelling-unit response rate is also provided because the MEPS estimation weights are initially adjusted for nonresponse at this level.

Sampling Weight Specifications

Because of the complex design of the MEPS HC, the MEPS sample data must be appropriately weighted to obtain approximately unbiased national estimates for the U.S. civilian noninstitutionalized population. The sampling weights developed for this purpose reflect the disproportionate sampling adopted in NHIS to oversample minority populations. They also reflect adjustments for:

- Complete nonresponse of eligible sample units.
- Partial response of survey participants providing data for only a portion of the time in 1996 during which they were eligible to respond.
- Poststratification to more accurate population totals obtained from the Current Population Survey (CPS).

The 1996 MEPS estimation weights are built from the estimation weights developed for the 1995 NHIS. To reduce the impact of large sampling weights on resultant variances of survey estimates, the MEPS estimation weights reflect a weight-trimming adjustment. The 1996 MEPS dwelling-unit weights also include an initial ratio adjustment to population estimates for selected socioeconomic measures derived from the full 1995 NHIS and subsequent adjustments for nonresponse to the first round of MEPS. In addition, the MEPS estimation weights developed at the person and family

levels reflect additional adjustments that poststratify the MEPS estimates to more accurate population totals obtained from CPS. The details of the development of MEPS estimation weights are described in this section.

Base Weights

Because the 1995 NHIS and MEPS are linked, the sampling weights developed for NHIS serve as the base weights for the 1996 MEPS. More specifically, the base weight for a dwelling unit selected in the 1996 MEPS is the nonresponse-adjusted 1995 NHIS quarter-specific estimation weight of the reference person in the primary reporting unit of a sampled dwelling unit. The reference person is the person who owns or rents the house. This NHIS estimation weight reflects the household's probability of selection in NHIS and adjustments for NHIS survey nonresponse.

More specifically, if

- $P(i)$ is the i th dwelling unit's probability of selection in NHIS to represent the Q th quarter of 1995 and includes disproportionate values associated with the oversampling of minorities, and
- $A(c)$ adjusts for NHIS nonresponse within nonresponse-adjustment class c , of which dwelling unit i is a member,

then the NHIS estimation weight $NHISWTQ(i)$, for the i th dwelling unit selected for MEPS in quarter $Q = 2$ or 3 , would take the form

$$NHISWTQ(i) = (1/P(i)) \times A(c)$$

The NHIS quarter-specific base weight was obtained by using the final-quarter basic NHIS weight from the 1995 NHIS analytical file delivered to AHCPR. The available estimation weight (HIS.WT.BF) also included a first-stage ratio adjustment (HIS.ADJ) that adjusts the initial NHIS population estimates to Census Bureau estimates for cross-classification of the population based on race/ethnicity (Hispanic, non-Hispanic black, other), Census region (East, Midwest, South, and West), and MSA classification (MSA/non-MSA). This component needed to be factored out of the NHIS estimation weight, since the first-stage ratio adjustment was implemented in

NHIS without reflecting the subsampling of NHIS PSUs for MEPS that occurred by MSA classification. Consequently, the initial MEPS base weight (WT.MEPI) was specified as

$$WT.MEP.I = HIS.WT.BF/HIS.ADJ$$

Use of the NHIS quarter-specific estimation weight across multiple quarters of 1995 to produce a national estimate required that the weight be divided by the number of quarters being pooled. Since the MEPS sample was confined to quarters 2 and 3 of calendar year 1995, the NHIS quarter weight, HIS.WT.BF, was initially divided by 2, HIS.WT.BF/2. Since the MEPS sample was restricted to panels 1 and 3 out of a four-panel NHIS design, it represented a 50-percent subsample of NHIS. Consequently, the NHIS quarter weight, HIS.WT.BF/2 (representing quarters 2 and 3), was further multiplied by 2 to reflect the 50-percent subsample considered for MEPS. Consequently, specification of the initial MEPS base weight as

$$WT.MEP.I = HIS.WT.BF/HIS.ADJ$$

reflects the restriction of the NHIS sample to quarters 2 and 3 and a 50-percent sample for the 1996 MEPS.

As mentioned previously, unmarried students ages 17-23 who were living in dormitories and had been respondents in the 1995 NHIS were not eligible for the 1996 MEPS. Furthermore, a very small set of NHIS dwelling units (54) that were determined to be eligible for MEPS at the time of sample selection could not be linked back to the 1995 NHIS analytical file that was provided a year later. The following strategy was implemented to obtain a base weight for these dwelling units. Median values of WT.MEPI were determined for the dwelling units eligible for MEPS that linked back to the NHIS analytical file, based on classes defined by a cross-classification of the minority status of the dwelling unit (1 if the dwelling unit has a Hispanic or black member, 2 otherwise) and 20 mutually exclusive and exhaustive distinct sampling strata defined for NHIS at the segment level for oversampling purposes. MEPS base-weight assignments for the nonlinked cases were made based on the median value of WT.MEPI for the class with which they were associated.

Trimming Base Weights

An initial examination of the distribution of the MEPS base weights identified a high level of variability. To correct for the impact of large sampling weights on resultant MEPS variance estimates, the initial MEPS base weights were trimmed according to the following specifications.

In each of the 40 classes (c) determined by a cross-classification of the dwelling unit's minority status and the 20 NHIS sampling strata defined at the segment level for oversampling purposes, the mean of the initial MEPS base weight, MEANDUWT(c) = MEAN(WT.MEPI(i&c)) was computed. If the initial MEPS base weight for the dwelling units within a given class c was greater than 3 times the mean of the base weights, the weight was truncated to that value. Otherwise, the sampling weight retained its initial value. More specifically, for class c, where c = 1,...,40,

If $WT.MEPI(i) > 3 \times MEANDUWT(c)$,
then $TRIMFAC = 3 \times (MEANDUWT(c) / WT.MEPI(i))$;

If $WT.MEPI(i) \leq 3 \times MEANDUWT(c)$,
then $TRIMFAC = 1$.

Consequently, the trimmed MEPS weights were specified as

$$TRIMDUWT(i) = TRIMFAC \times WT.MEP.I(i)$$

While only a few sampling weights were modified, the largest MEPS base weight was reduced by nearly 50 percent.

Ratio Adjustment of Trimmed Base Weights

To improve the accuracy of the MEPS estimates, the trimmed dwelling unit weights were subsequently ratio adjusted to population estimates derived from the full 1995 NHIS, using data from the first three quarters of the 1995 NHIS (all of the 1995 NHIS that was available at the time of MEPS sampling weights development). The following measures were used in specifying the

ratio-adjustment cells to facilitate the adjustment at the dwelling unit (DU) level:

- MSA status (central city, MSA; not central city, MSA; non-MSA).
- Family income classification of reference person (below poverty; under \$35,000 but above poverty; equal to or greater than \$35,000; unknown).
- Employment status of reference person (employed; not employed).
- Race/ethnicity of reference person (Hispanic; black non-Hispanic; other).
- Dwelling-unit level measure of activity limitations (at least one person in DU either cannot perform major activity or is limited in kind or amount of major or other activities; no member of DU has an activity limitation).

These measures were selected to represent a set of measures that related to the oversampling done in NHIS (DU minority status and MSA status). They also represented socioeconomic and health-specific measures potentially associated with health care use, expenditures, and insurance coverage that are not used to define the family- and person-level poststratification adjustments in MEPS. For dwelling units associated with more than one reporting unit, the reference person and family income of the primary reporting unit were used for classification purposes. Cross-classification of these measures yielded 144 weighting class cells to implement the ratio adjustment to more accurate national estimates at the household level based on the entire 1995 NHIS sample for quarters 1-3.

More specifically, the DU-level ratio adjustment for the c th weighting class takes the form

$$A(c) = \frac{NHISDU(c)}{\sum_{i \in c} TRIMDUWT(i)}$$

where $i \in c$ represents all NHIS dwelling units in c selected for the 1996 MEPS, $TRIMDUWT(i)$ represents the trimmed initial NHIS base weight for the i th dwelling unit selected for MEPS, and $NHISDU(c)$ represents the national population estimate at the dwelling-unit level for weighting class c , derived from the 1995 NHIS using data from quarters 1-3.

Consequently, the ratio-adjusted MEPS dwelling unit weight $DUPSWT(i)$ for the i th dwelling unit associated with class c , adjusted to population estimates derived from the full 1995 NHIS, takes the form

$$DUPSWT(i) = A(c) \times TRIMDUWT(i)$$

The sum of the ratio-adjusted and trimmed MEPS dwelling unit weights for the 10,597 NHIS dwelling units fielded for the 1996 MEPS was 104,002,800.

Round 1

Nonresponse-Adjusted Dwelling Unit Weights

Of the 10,509 dwelling units eligible for data collection in the first round of the 1996 MEPS, 8,793 (83.7 percent) contained at least one reporting unit that responded to the MEPS interview. Since survey nonresponse is potentially a significant source of bias in survey estimates, the MEPS dwelling unit sampling weights included an adjustment to help reduce its potential for bias. In general, the greater the difference among subgroups in response rates and the analytic characteristic(s) of interest, the greater is the need to adjust survey weights for nonresponse. In MEPS, a weighting class nonresponse adjustment was implemented, under the assumption that nonresponding sampling units would have responded in a manner similar to that of respondents with similar sociodemographic and economic characteristics within the same adjustment class. Properly designed, a weighting class nonresponse adjustment strategy can result in reduced nonresponse bias. The technique requires that the sample be partitioned into mutually exclusive classes, with classification information available for both responding and nonresponding units (Cox and Cohen, 1985).

Analyses were conducted of characteristics associated with differential nonresponse in MEPS. These analyses identified the most important measures to use in developing a nonresponse adjustment to the MEPS sampling weights to correct for potential nonresponse bias at the dwelling-unit level (Cohen and Machlin, 1998; DiGaetano and Goksel, 1996). To facilitate these comparisons, the demographic,

socioeconomic, health-related, and interview-specific profiles of respondents and nonrespondents were examined, based on available data for both groups from the 1995 NHIS.

Based on the results of these analyses, weighting classes were specified for the MEPS Round 1 dwelling unit nonresponse adjustments. They were defined by cross-classifications of the following measures:

- Family income of primary reporting unit (less than \$10,000; \$10,000-\$19,999; \$20,000-\$34,999; \$35,000 or more; unknown).
- Size of dwelling unit (one; two; three; four; five or more).
- MSA size (MSA, population 500,000 or more; MSA, population less than 500,000; non-MSA).
- Region (Northeast; Midwest; South; West).
- Employment classification of reference person (government; private sector; not in labor force/never worked/worked without pay; unknown or under 18 years of age).
- DU-level personal help measure (units with at least one member unable to perform personal care activities or other routine needs; remaining units with person 70 and over; remaining units with no limitations).
- Propensity to cooperate, based on providing phone number during NHIS (phone number provided; phone present but no number provided; no phone; unknown).
- Age of reference person (under 25; 25-34; 35-44; 45-64; 65 and over).
- Race/ethnicity of reference person (Hispanic; black non-Hispanic; other).
- Sex of reference person.
- Marital status (married, spouse present; other).

Overall, 49 cells were identified based on cross-classifications of these measures, with cell collapsing specified according to a hierarchy determined by significance level.

More specifically, the nonresponse adjustment for the c th weighting class takes the form

$$B(c) = \frac{\sum_{i \in c} E(i)DUPSWT(i)}{\sum_{i \in c} R(i)DUPSWT(i)}$$

where

$DUPSWT(i)$ is the initial MEPS Round 1 dwelling unit weight for the i th sample dwelling unit, which reflects the reciprocal of the dwelling unit's selection probability for MEPS and a poststratification adjustment to 1995 NHIS population totals;

$E(i) = 1$ for all MEPS dwelling units selected for the Round 1 interview; $E(i) = 0$ otherwise;

$R(i) = 1$ for all selected MEPS dwelling units responding in Round 1, $R(i) = 0$ otherwise; and

$i \in c$ represents eligible dwelling units classified in weighting class c .

Consequently, the estimation weight adjusted for MEPS Round 1 dwelling unit nonresponse, $WGTDU1(i)$, for the i th dwelling unit associated with class c , takes the form

$$WGTDU1(i) = B(c) \times DUPSWT(i)$$

The sum of the nonresponse-adjusted MEPS dwelling unit weights for the 8,793 eligible dwelling units with at least one responding reporting unit in Round 1 of MEPS was 102,892,600.

Family-Level Estimation Weights

In MEPS, a family is defined as a person or group of persons who are living together and are related by blood, marriage (or partnerships that are viewed as such), adoption, or other family associations. Any related college students under age 24 who usually live in the sampled household but are currently living away from home and going to school full time are considered to be members of the family. These college students are considered key persons in MEPS and are interviewed at their college address but linked to the sampled household for family analyses. Families in MEPS without college students living away from home were identified as single reporting units. Families in MEPS with college students living away from home were identified by linking the student reporting unit(s) back to their parent(s)' reporting unit.

To be considered a responding family in MEPS for the Round 1 interview, the family needed to include at least one person who was key, in scope, and eligible for data collection. Furthermore, all such key, in-scope, and eligible persons had to have responded for their entire period of eligibility in 1996 covered by the Round 1 interview. Each family in MEPS characterized as responding was assigned the weight of its corresponding dwelling unit (adjusted for nonresponse):

$$WGTFAM(i) = WGTDU1(i)$$

Overall, 9,488 reporting units responded in the first round of MEPS. After linking the responding student reporting units back to their parent(s)' family, this translated to 9,388 responding families.

The initial weights at the family level were then further poststratified to reflect population estimates obtained from the March 1996 CPS for unrelated individuals plus families. This poststratification also served as an adjustment for nonresponse at the family level. The weighting classes that were considered for the family-level poststratification adjustment were defined by a cross-classification of the following variables, defined at the time of the MEPS Round 1 interview:

- Family type—reference person (oldest person when there is no reference person) married and spouse present; male reference person and spouse not present; female reference person and spouse not present.
- Race/ethnicity of reference person (oldest person when there is no reference person)—Hispanic; black non-Hispanic; other.
- Region—Northeast; Midwest; South; West.
- MSA status—MSA; non-MSA.
- Number of persons in family—one; two; three; four; five or more.
- Age of reference person (oldest person when there is no reference person)—under 35; 35-44; 45-64; 65 and over.

More specifically, the family-level poststratification adjustment for the c th weighting class takes the form

$$C(c) = \frac{CPSFAM(c)}{\sum_{i \in c} WGTFAM(i)}$$

where $CPSFAM(c)$ represents the national population estimate at the family level for weighting class c , derived from the March 1996 CPS; $i \in c$ represents all MEPS family units classified in c that responded to the Round 1 interview; and $WGTFAM(i)$ represents the initial MEPS family-level weight for the i th family unit responding in the 1996 MEPS. Consequently, the Round 1 poststratified MEPS family unit weight $WGTRU1(i)$, for the i th family unit associated with class c , adjusted to population estimates derived from the March 1996 CPS, takes the form of:

$$WGTRU1(i) = C(c) \times WGTFAM(i)$$

The weighted estimate of the number of family units (including single-person units) containing at least one member of the U.S. civilian noninstitutionalized population is 110,206,950. It is obtained by summing the poststratified MEPS family unit weights for the 9,388 MEPS family units that responded to the Round 1 interview. In the development of family-level attributes in MEPS, all eligible and responding persons who are in scope, both key and non-key individuals, should be included.

Person-Level Estimation Weights

Estimation weights were assigned for key in-scope MEPS participants who were in responding Round 1 reporting units and for whom data were obtained for their entire Round 1 period of eligibility in 1996. To be considered a responding survey participant in MEPS for the Round 1 interview, the person needed to be in scope, with data provided for their entire Round 1 period of eligibility in 1996. All key, in-scope, and eligible sample participants in MEPS who satisfied these criteria for survey response were assigned estimation weights. The initial person-level estimation weight assigned to these MEPS survey respondents was the corresponding poststratified estimation weight for the family unit of which they were a member,

$$WGTPER(i) = WGTRU1(i)$$

Overall, 23,612 key, in-scope, and eligible individuals were classified as survey respondents in the first round of MEPS. The Round 1 MEPS person-level

weights were then poststratified to population totals obtained from the March 1996 CPS.

To establish consistency between family-level and person-level estimates, the final person weight for the reference person for each family (oldest person when there is no reference person), in addition to married couples living together, was the value of the MEPS family-unit weight:

$$WGTSP1(i) = WGTRU1(i)$$

The person-level estimation weights of all other MEPS key, in-scope, and eligible survey respondents (e.g., children of reference persons) were poststratified to population totals obtained from the March 1996 CPS. Weighting classes were defined by a cross-classification of the following variables:

- Region (Northeast, Midwest, South, West).
- Race/ethnicity (Hispanic, black non-Hispanic, other).
- Sex (male, female).
- Age at interview date (under 1, 1-4, 5-9, 10-14, 15-19, 20-29, 30-34, 35-44, 45-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80 and over).

This adjustment served as both a nonresponse and poststratification adjustment at the person level. The person-level poststratification adjustment factor for the *c*th weighting class takes the form

$$D(c) = 1$$

for the reference person for each family (oldest person when there is no reference person), in addition to married couples living together (denoted by $L(i) = 1$; $L(i) = 0$ otherwise). For others, it takes the form

$$D(c) = \frac{CPSPER(c) - \sum_{i \in c} WGTPER(i)L(i)}{\sum_{i \in c} WGTPER(i)(1 - L(i))}$$

where $CPSPER(c)$ represents the national population estimate at the person level for weighting class *c*, derived from the March 1996 CPS; $i \in c$ represents all MEPS key and in-scope survey participants classified in *c* who responded to the Round 1 interview; and

$WGTPER(i)$ represents the initial MEPS person-level weight for the *i*th person responding in the 1996 MEPS. Consequently, the Round 1 poststratified MEPS person weight $WGTSP1(i)$, for the *i*th person associated with class *c*, adjusted to population estimates derived from the March 1996 CPS, takes the form

$$WGTSP1(i) = D(c) \times WGTPER(i).$$

The weighted estimate of the number of persons who were members of the U.S. civilian noninstitutionalized population as of spring 1996 is 263,515,813. It can be derived by summing the poststratified MEPS person weights for the 23,612 MEPS key and in-scope survey participants classified as respondents for the Round 1 interview, as indicated on MEPS HC Public Use File HC-001: 1996 Panel Round 1 Population Characteristics (AHCPR Pub. No. 97-DP20, available on CD-ROM or through the MEPS Web site at <<http://www.meps.ahcpr.gov/>>).

Full-Year 1996

Part-Year Nonresponse Adjustment

The MEPS Round 1 person-level weight was developed to make estimates of the health care experience and insurance coverage of the U.S. civilian noninstitutionalized population for the first half of 1996. An annual person-level weight for 1996 was developed for deriving person-level estimates that cover all of calendar year 1996. Application of this weight permits the derivation of national estimates of the health care use, expenditures, insurance coverage, and sources of payment for the civilian noninstitutionalized population for calendar year 1996.

In order to be considered a responding survey participant in MEPS for the purpose of deriving annual 1996 estimates, the person had to be key and in scope, with data provided for their entire period of eligibility in 1996. If all the key, in-scope, and eligible sample participants in MEPS with positive values for the MEPS Round 1 person-level weight—in addition to new key and in-scope respondents who joined a responding household in 1996 after Round 1—had responded for their entire period of eligibility in 1996, no additional

adjustment for part-year survey nonresponse over the course of Rounds 1 through 3 would have been necessary. (New respondents acquired the sampling weight of the family they joined.) However, of 23,881 sample participants identified in MEPS, 21,571 (90.33 percent) provided data for their entire period of eligibility in 1996. Consequently, after factoring in the impact of survey attrition, the overall MEPS person-level response rate for deriving annual estimates was 70.2 percent ($.777 \times .903$).

Since survey nonresponse is potentially a significant source of error or bias in survey estimates, the MEPS full-year sampling weights included an adjustment for survey attrition to help reduce its potential impact. The characteristics that distinguish MEPS respondents with survey response for their entire period of eligibility in 1996 from Round 1 participants who discontinued survey participation were analyzed. This analysis identified the most important variables to incorporate into the nonresponse adjustments to the MEPS sampling weights to correct for part-year survey nonresponse. It was based only on data from the first two rounds of the survey because relevant Round 3 data for 1996 were unavailable at the time of the analysis. A logistic regression analysis identified the most important measures to include in specifying a nonresponse adjustment to the MEPS sampling weights to correct for part-year response for calendar 1996 at the person level. Based on the results, weighting classes were specified for the MEPS full-year person-level nonresponse adjustments. They were defined by cross-classifications of the following measures as of Round 1 (or as of the first eligible round in MEPS for key and in-scope respondents who joined a household after Round 1):

- Round 1 interview classification (no initial refusal, initial refusal).
- Size of MEPS family (one, two, three, four, five or more).
- MSA (MSA, non-MSA).
- Age (under 20, 20-29, 30-44, 45-64, 65 and over).
- Marital status of reference person (married, widowed, divorced, separated, never married).

Overall, 218 cells were identified based on cross-classifications of these measures. Cell collapsing was

specified according to a hierarchy determined by significance level.

The nonresponse adjustment for the *c*th weighting class takes the form

$$F(c) = \frac{\sum_{i \in c} E(i)WGTSP1(i)}{\sum_{i \in c} R(i)WGTSP1(i)}$$

where

$WGTSP1(i)$ is the MEPS Round 1 poststratified person-level weight for the *i*th Round 1 respondent, and $WGTSP1(i) = WGTRU1(i)$ for key and in-scope respondents who joined a household in 1996 after Round 1 (here, the new respondent acquires the sampling weight of the family joined);

$E(i) = 1$ for all MEPS Round 1 respondents with positive values of $WGTSP1(i)$ and for key and in-scope respondents who joined a responding household in 1996 after Round 1 with positive values of $WGTSP1(i)$; $E(i) = 0$ otherwise;

$R(i) = 1$ for all persons with $E(i) = 1$ who responded for their entire period of eligibility in 1996; $R(i) = 0$ otherwise; and

i ∈ *c* represents all key and in-scope MEPS full- and part-year respondents classified in weighting class *c*.

Consequently, the estimation weight adjusted for survey attrition in MEPS covering calendar year 1996, $WGTSP2(i)$, for the *i*th person associated with class *c*, takes the form

$$WGTSP2(i) = F(c) \times WGTSP1(i)$$

for the 21,571 key and in-scope survey participants who responded for their entire period of eligibility in 1996.

Person-Level Estimation Weights

The subset of the 21,571 key and in-scope survey participants who responded for their entire period of eligibility in 1996 and who were also in scope on December 31, 1996, had their part-year nonresponse-

adjusted annual estimation weights further poststratified to Census Bureau population estimates as of December 1996. The person-level estimation weights, $WGTSP2(i)$, of the 21,326 sample participants that meet these criteria were poststratified to population totals obtained from the March 1997 CPS and further scaled to reflect Census Bureau population estimates as of December 1996 within weighting classes defined by a cross-classification of the following variables:

- Sex (male, female).
- Age as of December 31, 1996 (under 1, 1-4, 5-9, 10-14, 15-19, 20-24, 25-29, 30-34, 35-44, 45-54, 55-59, 60-64, 65-69, 70-74, 75-79, 80 and over).
- Race/ethnicity (Hispanic, black non-Hispanic, other).
- Region (Northeast, Midwest, South, West).
- MSA (MSA, non-MSA).

Within each of the weighting classes ($c \in C$) associated with a given age-by-sex cross-classification (C), the population totals derived from the March 1997 CPS were further adjusted by the factor $SCALE(C)$, which was defined as the ratio of the December 1996 Census Bureau population estimate to the March 1997 population estimate derived from the CPS (Table 2).

More specifically, the person-level poststratification adjustment for the c th weighting class takes the form

$$G(c) = \frac{MAR97CPS(c) \times SCALE(C)}{\sum_{i \in c} WGTSP2(i)}$$

where

$MAR97CPS(C)$ represents the national population estimate at the person level for weighting class c , derived from the March 1997 CPS;

$SCALE(C)$ represents the ratio of the December 1996 Census Bureau population estimate to the March 1997 population estimate derived from the CPS for the specific cross-classification of age and sex associated with cell C ;

$i \in c$ represents all key and in-scope survey participants associated with cell c ($c \in C$) who responded for their entire period of eligibility in 1996 and were also in scope on December 31, 1996; and

$WGTSP2(i)$ represents the annual person-level estimation weight for calendar year 1996 adjusted for survey attrition.

Consequently, the MEPS full-year 1996 person-level weight $WGTSP96(i)$ for the i th key, full-year survey participant in scope as of December 31, 1996, who is associated with class c , adjusted to population estimates derived from the March 1997 CPS and further scaled to Census Bureau estimates for December 1996, takes the form

$$WGTSP96(i) = G(c) \times WGTSP2(i)$$

The remaining 245 key in-scope MEPS survey participants responded for their entire period of eligibility in 1996 but were not in scope as of December 31, 1996 (e.g., persons who died during the survey year). These participants maintained their estimation weight, adjusted for survey attrition. Consequently, their MEPS full-year person-level weight was specified as

$$WGTSP96(i) = WGTSP2(i)$$

The weighted estimate of the number of persons in the U.S. civilian noninstitutionalized population as of December 1996 is 265,439,511. It can be derived by summing the poststratified MEPS person weights for the 21,326 MEPS key and in-scope survey participants classified as respondents and in scope as of December 31, 1996, as indicated on the MEPS HC Public Use File HC-003: 1996 Panel Population Characteristics and Utilization Data for 1996 (AHCPR Pub. No. 98-DP12, available on CD-ROM or through the MEPS Web site at <http://www.meps.ahcpr.gov/>). Once the full-year insurance coverage measures are available for release, analysts can produce cross-sectional national insurance coverage estimates as of December 1996 with the MEPS data. However, they should restrict their sample to this set of survey participants who were in scope as of December 31, 1996. Similarly, the weighted estimate of the number of persons in the U.S. civilian noninstitutionalized population over the course of 1996 is 268,130,477. It can be derived by summing the final poststratified MEPS person weights for the 21,571 MEPS key and in-scope survey participants who responded for their entire period of eligibility in 1996.

Table 2. Population estimates by sex and age for December 1996 and March 1997: United States

Age in years	December 1996 ¹		March 1997 ²	
	Male	Female	Male	Female
Total	129,578,301	135,861,210	129,842,460	136,084,232
Under 1	1,989,676	1,902,075	1,959,414	1,856,886
1-4	8,162,745	7,787,372	8,118,381	7,846,632
5-9	10,322,287	9,848,773	10,428,712	9,842,699
10-14	9,990,073	9,526,892	9,970,793	9,534,281
15-19	9,723,221	9,327,806	9,766,326	9,377,121
20-24	8,728,774	8,730,416	8,635,045	8,719,707
25-29	9,354,544	9,674,812	9,450,287	9,636,289
30-34	10,339,061	10,652,556	10,242,979	10,548,081
35-44	21,458,434	22,046,082	21,546,881	22,147,277
45-54	16,001,003	16,759,899	16,063,659	16,891,429
55-59	5,446,069	5,892,934	5,569,213	6,003,373
60-64	4,644,237	5,191,637	4,686,964	5,206,925
65-69	4,415,868	5,236,207	4,321,136	5,179,642
70-74	3,721,987	4,787,277	3,764,159	4,750,254
75 and over	5,280,322	8,496,472	5,318,511	8,543,637

¹Obtained from the U.S. Bureau of the Census.

²Computed from the Current Population Survey March 1997 data file.

Note: Civilian noninstitutionalized population.

Analysts who want to produce annual 1996 health care utilization and expenditure estimates should include all of these 21,571 key and in-scope MEPS survey participants for the purposes of estimation. A future MEPS data release will include an updated person-level weight further poststratified by poverty status classification and estimation weights to support annual 1996 family-level health care use and expenditure estimation.

Variance Estimation

The sample design of the MEPS HC includes stratification, clustering, multiple stages of selection, and disproportionate sampling. This complex sample design results in serious departures from the assumptions of simple random sampling. Furthermore, the MEPS sampling weights reflect differential adjustments for survey nonresponse and poststratification. These survey design and estimation complexities require special consideration with regard to variance estimation and analysis. To obtain accurate estimates of the standard errors associated with MEPS person- and family-level survey estimates, for either descriptive statistics or more sophisticated analyses based on multivariate models, the MEPS survey design complexities need to be taken into account. Several methods for estimating sampling variances that adjust for survey design complexities have been developed that are appropriate for analytical applications tied to MEPS (Cohen SB, 1997b). These variance estimation strategies include the Taylor-series linearization method, balanced repeated replication, and the jack-knife method.

The MEPS public use files include variables necessary for implementing a Taylor-series variance estimation approach for survey estimates. When using such an approach, the sampling strata and associated PSUs that define the MEPS survey design need to be specified. The corresponding variables on the MEPS Round 1 database (AHCPR Pub. No. 97-DP20) are VARSTR1 and VARPSU1, respectively. Similarly, the corresponding variables on the MEPS full-year 1996 utilization estimates database (AHCPR Pub. No. 98-DP12) are VARSTR96 and VARPSU96. Specifying a “with replacement” design in a variance estimation software package appropriate for the analysis of

complex survey data that utilizes the Taylor-series approach, such as SUDAAN (Shah, Barnwell, Bieler, et al., 1996) or Stata (StataCorp, 1997) will yield standard errors that have been appropriately adjusted for survey design complexities.

The number of degrees of freedom associated with estimates of variability indicated by such statistical software packages may not appropriately reflect the actual number. For purposes of variance estimation, it is estimated that MEPS sample estimates for the general population derived at the national level have approximately 100 degrees of freedom.

Summary

Because of the national scope and depth of the MEPS data collection effort and because MEPS survey estimates are used to inform national health policies, it is particularly important to use estimation strategies that improve the quality and accuracy of survey estimates. Research was conducted to help ascertain potential sources of bias that were attributable to MEPS dwelling unit nonresponse; these findings were incorporated in the specification of the MEPS weights adjustment strategy to help reduce the impact of nonresponse bias. The MEPS sample is linked to NHIS, so detailed information on the sociodemographic and health characteristics of the eligible MEPS sample was available to inform the investigation. This investigation revealed that the dwelling units responding to the first round of the MEPS HC differed from the nonrespondents in a number of dimensions. The multivariate analysis showed that family income, dwelling unit size, health status of household members (as measured by personal help needs), phone availability, MSA size, and item nonresponse for employment classification were significant factors in distinguishing MEPS respondents. The measures most significant in differentiating MEPS survey response status were used in the specification of the MEPS Round 1 dwelling unit nonresponse adjustments. Identification of weighting classes in MEPS that capture the greatest variation in response rates across subgroups should reduce the bias attributable to survey nonresponse.

An analysis was done of the characteristics that distinguish MEPS respondents with survey response for

their entire period of eligibility in 1996 from the Round 1 participants who discontinued survey participation. This analysis identified the most important measures to include in specifying a nonresponse adjustment to the MEPS estimation weights to correct for part-year survey nonresponse. A logistic regression analysis identified characteristics that distinguished the MEPS full-year respondents from their part-year respondent counterparts. These are the most important measures to include in specifying a nonresponse adjustment to the MEPS sampling weights to correct for part-year response for calendar year 1996 at the person level. Family size, residence by MSA classification, age, marital status, and reluctance to participate were found to be important factors in distinguishing the MEPS part-year respondents from their full-year counterparts.

The overall MEPS person-level response rate for deriving annual 1996 estimates was 70.2 percent, after adjusting for the multiplicative effects of nonresponse to NHIS, nonresponse to the first round of MEPS, and survey attrition. To further improve the accuracy of MEPS survey estimates, additional poststratification adjustments were incorporated in the development of the annual MEPS estimation weights. The poststratification adjustments relied on population estimates derived from the CPS and other Census Bureau sources. A detailed summary of the MEPS estimation weight specifications has been provided in this report to ensure a better understanding of the estimation procedures that were adopted.

References

- Arnett RA, Hunter E, Cohen S, et al. The Department of Health and Human Services' Survey Integration Plan. In: Proceedings of the American Statistical Association (ASA). Section on Government Statistics; 1996 Aug; Chicago. Alexandria (VA): American Statistical Association; 1996.
- 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 SB. The redesign of the Medical Expenditure Panel Survey, a component of the DHHS Survey Integration Plan. In: Proceedings of the Council of Professional Associations on Federal Statistics (COPAFS) Seminar on Statistical Methodology in the Public Service. 1996 Nov; Bethesda (MD).
- Cohen SB. Sample design of the 1996 Medical Expenditure Panel Survey Household Component. Rockville (MD): Agency for Health Care Policy and Research; 1997a. MEPS Methodology Report No. 2. AHCPR Pub. No. 97-0027.
- Cohen SB. An evaluation of alternative PC-based software packages developed for the analysis of complex survey data. *The American Statistician* 1997b;51(3):285-292.
- Cohen SB, Machlin SR. Characteristics of nonrespondents in the MEPS Household Component. Proceedings of the American Statistical Association. Section on Survey Research Methods; 1998 Aug; Dallas. Alexandria (VA): American Statistical Association; 1998.
- Cox BG, Cohen SB. Methodological issues for health care surveys. New York: Marcel Dekker; 1985.
- DiGaetano R, Goksel H. Internal report on factors associated with MEPS Round 1 dwelling unit nonresponse. Unpublished report. Rockville (MD): Westat, Inc.; 1996.
- Judkins D, Marker D, Waksberg J. National Health Interview Survey: research for the 1995 redesign. Unpublished report. Prepared under contract for the National Center for Health Statistics. Rockville (MD): Westat, Inc.; 1994.
- Shah BV, Barnwell BG, Bieler GS, et al. Technical manual: statistical methods and algorithms used in SUDAAN release 7.0. Research Triangle Park (NC): Research Triangle Institute; 1996.
- StataCorp. Stata Statistical Software: Release 5.0. College Station (TX): Stata Corporation; 1997.