Outpatient Prescription Drugs: Data Collection and Editing in the 1996 Medical Expenditure Panel Survey (HC-010A)

MEPS

Methodology Report 12

U.S. Department of Health and Human Services
Public Health Service
Agency for Healthcare Research and Quality

Agency for Healthcare Research and Quality

Health Care Information and Electronic Ordering Through the AHRQ Web Site

The Agency for Healthcare Research and Quality Web site—http://www.ahrq.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.

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

http://www.ahrq.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 Healthcare Research and Quality (AHRQ). For the first time in a national expenditure survey, the 1996 MEPS included a detailed collection of information on prescription medicines obtained from pharmacy providers frequented by household sampled persons. The information was collected by means of a linked survey of pharmacy providers. This report describes the procedures adopted to collect and edit these prescription drug data for public release. It includes efforts made to retrieve complete and/or partially missing pharmacy data, the editing techniques used to fill in remaining missing data in the pharmacy database, and the matching/imputation procedure that linked every prescription drug mentioned by the respondent in the MEPS Household Component to a specific prescription drug from the Pharmacy Component (part of the Medical Provider Component).

Suggested citation

Moeller JF, Stagnitti MN, Horan E, et al. Outpatient prescription drugs: data collection and editing in the 1996 Medical Expenditure Panel Survey (HC-010A). Rockville (MD): Agency for Healthcare Research and Quality; 2001. MEPS Methodology Report No. 12. AHRQ Pub. No. 01-0002.

Outpatient Prescription Drugs: Data Collection and Editing in the 1996 Medical Expenditure Panel Survey (HC-010A)

MEPS

Methodology Report 12

U.S. Department of Health and Human Services
Public Health Service
Agency for Healthcare Research and Quality
AHRQ Pub. No. 01-0002
June 2001

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 Healthcare Research and Quality (AHRQ), formerly the Agency for Health Care Policy and Research, 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 AHRQ 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 private and public-sector employers. 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 three 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.

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. Beginning in 2000, national estimates of employer contributions to group health insurance from the MEPS IC are being used in the computation of Gross Domestic Product (GDP) by the Bureau of Economic Analysis.

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 AHRQ Publications Clearinghouse. Write or call:

AHRQ 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 AHRQ number of the document or CD-ROM you are requesting. Selected electronic files are available through the Internet on the AHRQ Web site:

http://www.ahrq.gov/

On the AHRQ 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 Healthcare Research and Quality, 2101 East Jefferson Street, Suite 500, Rockville, MD 20852 (301-594-1406).

Table of Contents

utpatient Prescription Drugs: Data Collection and Editing in the 1996 Medical Expenditure Panel Survey (HC-010A)

by John F. Moeller, Ph.D., and Marie N. Stagnitti, M.P.A., Agency for Healthcare Research and Quality; Eileen Horan and Pat Ward, Ph.D., Westat, Inc.; and Nancy Kieffer and Ed Hock, Social and Scientific Systems, Inc.

Introduction

This report describes the procedures adopted to collect and edit the 1996 Medical Expenditure Panel Survey (MEPS) prescription drug data for public release (MEPS public use file HC-010A). For the first time in a national expenditure survey, the 1996 MEPS included a detailed collection of information on prescription medicines obtained from pharmacy providers frequented by household sampled persons. The information was collected by means of a linked survey of pharmacy providers.

Because of nonresponse to the linked survey, these prescription data were not available for every sampled person in MEPS and were not necessarily complete for those persons for whom the Pharmacy Component (PC) survey data were collected. Hence, this report discusses the data collection efforts made to retrieve complete and/or partially missing pharmacy data, the editing techniques used to fill in remaining missing data in the pharmacy database, and the matching/imputation procedure that linked every prescription drug mentioned by the respondent in the MEPS Household Component (HC) to a specific prescription drug from the PC. The abbreviations used in this report are listed in Appendix A.

The MEPS prescription drug use and expenditure data comprise a critical component of health expenditures collected in the survey. Recently, there has been considerable policy interest in developing a prescription drug benefit for the Medicare population. In addition, the rising cost of prescription drugs has been singled out as a leading contributor to the escalating premiums for plans covering the non-Medicare population. Apart from coverage issues, it is critical to have accurate, detailed information on drug therapies in order to analyze alternative treatment regimens for disease and to investigate the potential for adverse drug interactions in treating chronic illnesses.

Previous household health expenditure surveys have been criticized for underestimating utilization and expenditures on prescribed medicines because of respondent underreporting of prescription data. The potential for this problem is understandable when considering the length of the recall period in the expenditure surveys, the burden placed on the respondent to report details of numerous medication purchases for household members, the irregular frequency with which prescriptions are purchased for treating acute health conditions during a survey period, and the complexity of reimbursement mechanisms for prescription medicines.

The general approach used in MEPS to address the underreporting issue was to relieve the household of the burden of reporting detailed financial information for every prescription purchase during each round of the survey. Instead, computerized printouts from respondents' pharmacy providers that contained such information were used when they were available. When computerized printouts were unavailable from a pharmacy provider, completed written data forms were secured, when possible. These printouts or forms also provided detailed information about the pharmacology of the prescription summarized by the National Drug Code (NDC). Heretofore, this information had not been available from household health expenditure surveys. In addition, efforts were made to improve the reporting accuracy of prescription utilization by asking the household respondent about medications prescribed in conjunction with other medical events, such as hospital stays, emergency room visits, and doctor visits. This information was gathered at the same time during the survey that the respondent was queried about other nonprescription events.

The challenge was to match prescription mentions by the household to the prescription purchases on their computerized printouts. When computerized printouts were not available for MEPS households from their pharmacy providers, detailed information on specific prescription purchases by participating individuals in the PC had to be imputed to the drug mentions of the HC nonparticipating individuals. (HC nonparticipating

individuals are respondents whose pharmacy providers were not contacted because permission forms were not signed or whose pharmacy providers were contacted but did not provide information about the individual's prescriptions.) Once the pricing and payment data from the PC were either matched or imputed to all of the household prescription mentions, the prescription expenditures and payments could be "rolled up" from the event or purchase level to the person level for estimating national prescription expenditures and payment sources for the entire U.S. civilian noninstitutionalized population in 1996.

Data Collection

Household Component

Prescription drug data were collected in the MEPS HC questionnaire and in the linked MEPS PC. During each round of the MEPS HC, all respondents were asked to supply the name(s) of any prescribed medication that they or their family members purchased or otherwise obtained during that round.

Respondents were first given an opportunity to mention prescribed drugs when they were surveyed about other (nonprescription) health service events. They were given a last opportunity to mention prescribed medicines during the prescribed medicines section of the HC. The order in which health care events are described in the HC is simply an artifact of the design of the HC. During each round of the HC, detailed information was obtained on various types of health care service events, including prescribed medicines. When respondents provided information on a health care event that was not a prescribed medicine (e.g., an emergency room visit), they were asked to supply information describing the health care event itself, as well as the names of any medications that were prescribed during that event. In addition, prescribed medicine mentions could be added when respondents went through the prescribed medicines section of the household questionnaire.

These are the ways in which a respondent "created" (possible during Rounds 1-5) and/or "selected" (possible during Rounds 2-5) prescribed medicine mentions for their prescribed medicines roster. "Created" means that the respondent had not mentioned the prescribed medicine in any previous round of the survey, while "selected" means that the respondent had mentioned the prescribed medicine during a previous round. This roster served as a "base" prescribed medicines roster for that respondent throughout all of his or her rounds in MEPS. In each round, respondents had a final opportunity to add any additional medication names to their roster of prescribed medicines in the prescription drug section of the MEPS HC.

The following information was collected in the prescribed medicines section of the questionnaire for each medication listed on the roster in each round of MEPS: whether any free samples of the medication were obtained; the name(s) of any health problem(s) for which the medication was prescribed; the number of times the prescription drug was obtained or purchased; the year, month, and day on which the person first used the medication; and a list of the names, addresses, and types of pharmacies that filled the household's prescriptions. In addition, all the HC respondents were asked if they send in claim forms for their prescriptions (this type of person is referred to as a self-filer, or SF) or if their pharmacy providers do this automatically for them at the point of purchase (this type of person is designated as a non-self-filer, or NSF). Uninsured persons were treated in the same manner as NSFs.[1] The uninsured were not asked charge and payment questions during the HC. Only SFs were asked for charge and payment information about their prescription purchases on the household questionnaire. Payments by private third parties for SF prescription purchases would not be available from the pharmacy provider.

When diabetic supplies and equipment (such as syringes and insulin) were mentioned in the section of the MEPS HC on other medical expenses, the interviewer was directed to collect information on these items in the prescription medicines section. To the extent that these items are purchased without prescription, they represent a nonprescription addition to the MEPS prescription drug expenditure and utilization data.

[1] The uninsured are included in the NSF group throughout this report.

Pharmacy Component

The PC was designed as a mail survey of the pharmacy providers identified by household respondents during the series of MEPS interviews covering calendar year 1996. (See Appendix B for a facsimile of the mailed PC survey booklet.) During the last of these interviews, the household respondents were asked to sign permission forms (Appendix C) authorizing the project to contact their pharmacies and authorizing the pharmacies to release a respondent's pharmacy records. Only those pharmacies for which a household respondent signed this permission form were included in the linked followback survey.

The data collection protocol consisted of an initial mailing to all of the nominated pharmacies for which one or more permission forms had been obtained, a second mailing to nonrespondents, and followup telephone prompting of pharmacies that did not respond to the mailings. The initial mailing (Appendix B) was designed in the form of a printed booklet containing an introductory letter from the Agency for Healthcare Research and Quality (AHRQ) and the National Center for Health Statistics (NCHS), a brief description of MEPS and the PC, answers to frequently asked questions about the study, and an explanation of the data items being requested. A computer printout listing the persons for whom the pharmacy was being asked to provide information and copies of the signed permission forms were inserted inside the back cover of the booklet. The data request offered pharmacies two main options for responding. If available, pharmacies were invited to send computerized printouts of the data for the identified patients. This was seen as a response option imposing minimal burden on pharmacies, many of which routinely provide such printed listings to customers who request them. Alternatively, pharmacies could fill in the requested information on data forms, which were inserted in the booklet along with the patient list and permission forms.

The pharmacies were asked to provide information about each prescription filled or refilled for the named patients during calendar year 1996. For each medication, they were asked to provide:

- The date the prescription was filled or refilled.
- The NDC.
- The medication name (generic or brand).
- The strength of the medicine.
- The quantity dispensed.
- The total charge.
- The sources of payment.
- The amount of payment made by each source.

The initial mailings were directed to the individual retail pharmacies or other specific locations identified by the household respondents as the places from which household members had obtained their prescriptions. Although it was expected that some pharmacy chains might require corporate permission before allowing their individual locations to participate and that some would prefer to provide information from regional or corporate resources, the plan was to make the first contacts at the individual locations, working up the corporate ladder only after being referred there by the local pharmacies. "Chain" codes were assigned to the individual pharmacies, creating a mechanism for associating local establishments with a shared corporate parent.

The final round of household interviews through which the pharmacy sample was identified ended in July 1997. After a period of sample preparation, the first waves of mailings were released in September 1997. The bulk of the mailings, grouped in nine waves defined by groups of States, was completed by the end of October 1997; a final wave of cases that had required problem resolution was released in January 1998. The followup mailings began in late September 1997 and continued into December 1997. Calls to prompt nonresponding pharmacies began in mid-December 1997.

The response to the first wave of mailings was promising, with replies received from as many as 40-50 percent of the pharmacies. The second mailing and the telephone prompts, however, added only marginally to the initial response.

The majority of the returns received were in the form of printouts. When reviewed for processing, many of these proved to be incomplete or unclear in their presentation of key data items. The identification of third-party payers and the amounts paid by third parties were the data items most frequently missing. Variations in the way the printouts were formatted and the manner in which data items were labeled frequently resulted in ambiguity about the meaning of specific items on the printouts, requiring some followup contact with the pharmacy for clarification.

To improve response at the levels of both the pharmacy and the individual data items, the initial data collection protocol was supplemented with a two-pronged telephone data collection effort. One group of telephone interviewers concentrated on data retrieval calls to pharmacies that had responded. Retrieval telephone calls to collect missing data items or to clarify data on the printouts were needed for nearly 70 percent of the responding pharmacies. A second group of telephone interviewers concentrated on primary data collection from the nonresponding pharmacies, adopting an approach similar to that used for the MEPS Medical Provider Component (MPC) (Cohen, Monheit, Beauregard, et al., 1996). These interviewers contacted pharmacies by telephone, explained the data collection request, and faxed copies of the permission forms and other relevant materials to the pharmacies. Within several days of the faxing, they placed additional calls to collect the requested data over the phone or to prompt their pharmacy contact to send in the printed patient printouts. Most responding pharmacies chose to mail or fax the printouts to the study.

The bottom-up approach adopted for dealing with the large pharmacy chains yielded mixed results. Many of the individual pharmacies associated with chains responded directly to the initial mailed request for data. However, as the telephone followup work progressed, pharmacies for a number of chains referred interviewers to regional or corporate offices. Corporate contacts reacted to the data requests in several ways: some referred the interviewers back to the individual pharmacies, with or without corporate endorsement of the study; some chose to consolidate the project's requests and provide data from a centralized location; and several of those that undertook an effort to provide the information were unable to provide it on a timely basis or abandoned the effort as too burdensome.

The permission form response rate (that is, the rate at which household respondents who were asked to sign permission forms actually signed them) is shown below.

Eligible person-pharmacy pairs.....20,023
Signed permission forms.....14,531
Permission form signing rate.....72.6

Table 1 shows response rates for the pharmacy data collection. Response rates are shown at both the pharmacy level (72.2 percent) and the household patient-pharmacy pair level (67.1 percent).

Data Editing, Imputation, and Matching

The general approach to preparing the household prescription data for public release was to impute information collected from pharmacy providers to the household drug mentions. For SFs, information on payment sources was retained if these data were reported in the charge and payment section of the household questionnaire. A matching program was developed to link drugs and drug information from the PC to HC drug mentions. To improve the quality of these matches, all drugs on the household files were assigned numeric codes from a proprietary database on the basis of the medication names provided by the household. These codes were also assigned to the prescriptions in the PC by using the NDC, when available, and medication names reported by the pharmacy providers. Considerable editing was done prior to the matching to identify free samples among household drug mentions, to correct data inconsistencies in both data sets, and to fill in missing data and correct outliers on the pharmacy file. After the matching, household drug mentions in Round 3 of MEPS, which spanned portions of both 1996 and 1997, had to be allocated to each year to produce the final annual prescription use and expenditure data for survey year 1996.

Drug Coding and Flat Files

The initial task for the drug editing was to assign a common set of drug codes to the household drug mentions and to the prescription drugs reported by the pharmacy providers. Westat (the MEPS data collection contractor responsible for collecting the pharmacy data from the participating pharmacies and producing the initial HC and PC files containing each separate drug purchase) contracted with Aspen Systems Corporation to provide coding support services as a subcontractor for the MEPS project. First DataBank's proprietary 1998 Master Drug Data Base (MDDB), which contains the Generic Product Identifier (GPI) code, was selected for this task. The GPI is a 14-digit code that contains 7 pairs of digits. The first pair of digits represents the drug group. Successive paired digits represent the drug class, drug subclass, drug name, drug name extension,

Table 1. Pharmacy data collection response rate in the 1996 Medical Expenditure Panel Survey

Sample type	Pharmacies		Person-pharmacy pairs	
	Number	Percent	Number	Percent
Initial sample	6,109	100.0	14,531	100.0
Out of scope[a]	788	12.9	2,385	16.4
Net sample	5,321	87.1	12,146	83.6
Complete	3,840	72.2	8,149	67.1
Refusal	325	6.1	1,114	9.2
Other nonresponse	1,156	21.7	2,883	23.7

[a] The category "out-of-scope pharmacies" included establishments reported by a household respondent that were not pharmacies (e.g., when the medicine was given as a free sample in a physician's office), pharmacies located outside the United States, originally reported pharmacies that merged with another reported pharmacy during the data collection period, and pharmacies whose associated sampled persons were out of scope. It also included pharmacies that did not fill prescriptions for the sampled persons in 1996 but may have done so in later years. A person-pharmacy pair was treated as out of scope if the pharmacy to which the person was linked was out of scope.

Note: The sample for the Pharmacy Component (PC) included all persons reported to have had prescriptions filled or refilled during the first three rounds of the Household Component (HC). For pharmacies reported in the HC in Round 3, which overlapped the end of 1996 and the start of 1997, the sample processing did not have respondents differentiate between prescriptions obtained in 1996 and those obtained in 1997. This was not true for what the PC pharmacies were asked and reported. PC pharmacies were asked for 1996 data only. If, during data collection for 1996 prescriptions, a pharmacy acknowledged the person as a customer but reported having filled no prescriptions for the person in 1996, the person was treated as out of scope for 1996 pharmacy data collection (but in scope for the 1997 data collection).

Source: Center for Cost and Financing Studies, Agency for Healthcare Research and Quality: Medical Expenditure Panel Survey, 1996, public use file HC-010A.

dosage form, and strength. Coders filled in as many digits of the GPI as possible based on the medication name (and any supplementary information appended to the name) provided by the household, and the NDC, medication name, and other information provided by the pharmacy provider. Typically, 8 to 10 digits were coded for household-reported drugs, and all 14 digits were filled in for pharmacy-reported drugs.

The second task for the prescription editing work was to determine the MEPS variables to add to the household and pharmacy event-level files for later use. Separate household event files were constructed for drugs reported by households classified as SFs and for drugs reported by households classified as NSFs. The additional data on prescription charges and payments collected for the SFs, and the additional processing this would entail, made this necessary. To maintain consistency between the PC and HC databases, household-reported drug mentions needed to be unfolded into individual records; that is, each prescription constituted an individual record, whether it was a refill or an initial purchase. An individual record on the Pharmacy Component file represented a single prescription purchase regardless of whether the purchase represented an initial prescription or a refill. For matching purposes, it was therefore necessary to create separate records for each prescription, whether initial purchase or refill, on the household event files.

Variables added to the household SF and NSF flat files, as well as to the pharmacy flat file, were critical to the matching and editing/imputation processes. (Flat files are files in which each record represents one prescribed medicine event, whether it be an original prescription, a refill, or a free sample.) These variables were:

- Prescription event and person identifiers.
- Beginning and ending reference period dates for each round.

- An indicator of private health insurance coverage for prescription drugs by round.
- Potential source-of-payment indicators by round.
- Whether a person was in a health maintenance organization (HMO) by type of HMO (Medicare HMO, Medicaid or other public HMO, or private HMO) by round.
- Conditions associated with prescription drugs by round.
- Geographic division and region, in addition to metropolitan statistical area (MSA) status.
- Various health status and demographic characteristics.

Additional variables added to the pharmacy flat files included round-specific pharmacy identifiers and names and types of pharmacy providers.

Preliminary HC Event Edits

The three flat files contained 15,804 SF drug events for 1,562 persons, 188,422 NSF drug events for 14,629 persons, and 84,029 PC drug events for 6,874 individuals. The individuals in the SF and NSF groups were not mutually exclusive because 941 persons had events in both files; however, additional analysis showed that no single person was classified as both an SF and NSF in the same round. Excluded from the flat files were 899 SF drugs, 6,691 NSF drugs, and 32 PC drugs for persons in the preliminary MEPS Round 1 and Round 2 samples who were not in the full-year population. These persons did not have positive-valued full-year weights, nor were they related to anyone with such weights. Remaining in the full-year population, however, were some non-key persons who had zero person weights but were members of families in which at least one individual on the full-year file had a positive-valued person weight.

For SFs, a preliminary set of edits based on similar edits applied to other nonprescription MEPS expenditure events was run on the household-reported data for each drug event. These edits mostly relied on household variables created to classify actual payment source variables or potential payment source variables. Actual payment source variables were coded as "known payer, known amount paid," "known payer, unknown amount paid," "not known to be a payer," or " source not available." Potential payment source variables were coded as "covered or source available in round" or "not covered or source available to all in round." "Source available to all in round" was the way the potential payment source variable for self-payments was always coded. In addition, round-specific household variables indicating whether the person was enrolled in a private health insurance plan covering prescription drugs and whether the person was enrolled in a Medicare HMO, a Medicaid or other public HMO, or a private HMO were used in combination with other variables from the HC charge and payment section to edit the data and/or correct for inconsistencies. The purposes of these edits are summarized below. The number of events affected by the edit is shown in parentheses.

- Make it impossible for elderly persons enrolled in a Medicare HMO plan to claim any source other than Medicare as a source of payment for prescriptions (95 events).
- Equate the total charge to the sum of payments if the two differed by $2 or less (126 events).
- Eliminate the inconsistency created from persons' mistakenly reporting private insurance when they actually had Medicare HMO or Medicaid HMO insurance coverage (6 events).
- Correct inconsistencies between Medicare and Medicaid reported as sources of payment and coverage (0 events).
- Correct inconsistencies between reported insurance coverage during the year and potential coverage from private insurance, Medicaid, Medicare, and CHAMPUS (Armed-Forces-related coverage) (68 events).
- Assign a missing payment amount to self-pay when no information was available to link the missing payment to a third-party payer (2 events).
- Eliminate Medicaid as a source of payment when sources other than out-of-pocket are present and when there is an out-of-pocket payment greater than $5 (118 events).

In a second set of preliminary edits on the household drug event data for SFs, five edit rules were imposed in the order shown below. Only one edit rule was allowed per event. For example, if an event was edited because of the first rule listed, that event would not be eligible for any of the editing rules that followed the first rule.

- Medicaid cases were processed so they would be correctly classified for later imputation (131 events).

- The problem of having persons confuse "other" source of insurance with private insurance coverage was resolved (0 events).
- If total charge was reported to be $5 or more, no payments were reported, and self-pay was the only missing payment source, then self-pay was set equal to the total charge for the drug product (24 events).
- If the payment sum exceeded $30 and no payment sum information was reported as missing, then the editing for the event was designated as completed (2,639 cases).
- If the payment sum exceeded $30 and was within $5 of the total charge, and one payment source was missing, then the missing payment amount was set equal to zero (43 events).

During each round, households were asked in the prescription drug section of the HC questionnaire whether they received any free samples of prescribed medicines from any medical or dental provider. If they responded in the affirmative, respondents were then asked to name the medicines they received as free samples in a given round. However, no information was reported on the number of drug events for a given medication in a given round that were free samples. Initially, 693 prescription drug events for SFs were deemed to be free samples. For a medicine to qualify as a free sample for an SF, the following two conditions had to be met: (1) all payment sources were reported as zero and (2) the household reported receiving free samples of the medication. After 41 free sample Round 3 prescriptions of SFs were allocated to 1997, the final number of 1996 free samples for SFs in the MEPS HC was 652. (See the "Allocation to 1996 and 1997" section, below.)

For NSFs, 4,252 drug events initially were designated as free samples. Only one drug event per round per drug product was allowed as a free sample if the household reported receiving free samples of the drug product during the round. If the person in the NSF household was in the Pharmacy Component and the designated "free sample" was later exactly matched to a prescription purchase on the PC file, then the free-sample designation was overridden. Of drugs originally designated free samples for NSF households, 462 were later matched to prescription purchases on the PC file. After 603 Round 3 prescriptions in MEPS were allocated to 1997, the final number of free samples for NSFs in 1996 was 3,187.

Preliminary PC Event Edits

Some preliminary edits were imposed on the PC data. They were patterned after similar edits applied to nonprescription data collected in the MEPS Medical Provider Component. If, based on information from the HC, a sampled person did not potentially have coverage from a public or private insurance source, then any missing code for the corresponding payment amount from the PC for a specific drug event was coded as a zero payment (4,912 events). After this, the sum of payments for a drug event on the PC was set to missing if any payment sources were coded as missing (11,486 events).

Additional edit rules to the PC drug events, patterned after similar edit rules applied to the nonprescription MPC events, were designed to do the following:

Rule P1: Correct information from pharmacies that mistakenly reported a private insurance payment source instead of a Medicare HMO (406 events) or Medicaid HMO (211 events) payment.

Rule P2: Set the total charge equal to the sum of payments if the two measures differed by no more than $2 and none of the payments or the total charge was missing (791 events).

Rule P3: Allocate the excess of total charge over a partially reported payment sum to a specific payment source either based on the pharmacy's identification of a single third-party source or based on potential third-party coverage of the person (95 events).

Rule P4: Eliminate an out-of-pocket payment when it appears that the pharmacy provider shows an unlikely amount for a patient copayment that is equal to the reported amount for Medicaid (18 events) or Worker's Compensation (1 event).

Matching Software

From the outset of the MEPS prescription drug editing work, it was clear that the hot-deck approach to imputing missing data was not going to be appropriate for this task. Donor hot-deck cells defined on the basis of a specific drug code or medication name were too small to stratify by other variables deemed to be correlated with the purchase of a specific drug product. Also, because the GPI codes encompassed numerous specific drug products with different NDC values, it was clear that software would have to be developed that

could "read" medication names from both donor and recipient files to improve the quality of the matches.

To meet these needs, the data processing contractor, Social and Scientific Systems, Inc. (SSS), developed software that imputed PC drug data to the household drug mentions by matching drug events from each file based on variables with both numeric characters (e.g., GPI codes, potential payment sources, age, sex, health status, and geographic location) and alpha characters such as the medication names supplied by the household and the pharmacy providers. The matching software that SSS developed had the following features:

- An overall score ranging from –1 to +1 was assigned to each donor drug that represented a potential match to a recipient drug, with +1 representing the highest score attainable when each match variable received the highest score possible. Separate weights were assigned to the match variables to reflect their importance relative to other match variables in determining the final overall score. Values between –1 and +1 for final scores were constructed as the weighted score for the match divided by the weighted score when all match variables receive the highest possible score. (An example is given in the next section.) Certain match variables could be required to match exactly or a potential match between drugs was not deemed possible.
- All numeric match variables had to match exactly or no positive contribution was made to the final score for the specific variable. A value of +1 was assigned in the event of any exact match. A value of –1 was assigned if there was no exact match.
- For words (alpha characters), the best match was found according to the following hierarchy:
  The words matched exactly.
  The words sounded the same using a Soundex function. (The Soundex system indexes names by how they sound rather than how they are spelled).
  A pair of characters was swapped in the third to last characters.
  Only one character was different in the third to last characters.
  The shorter word exactly matched the first characters in the other word.
  The words started with the same characters.
  None of the above.
  To resolve ties, except when the words started with the same characters, the longer word was used.
  When the words started with the same characters, the word that started with more of the same characters was used, and then the longer word was used if there were still ties.
- For words (alpha characters), once the best match was found, it was assigned a score measuring how closely the words matched, and it was also assigned a weight indicating how strongly a good or poor match should be considered. Except for an exact match or a non-exact match where the words started with the same characters, higher scores were assigned to matches between longer words. For a non-exact match where the words started with the same characters, the larger the portion of the words that matched, the higher the score that was assigned. An exact match was given the highest score regardless of the size of the word.
- When there were ties among the final scores, a random number between zero and one was generated to break the ties, and the highest number assigned determined the final match. After scrutinizing numerous examples of these ties, it was determined that in general the ties reflected the lack of enough information in the database to identify a uniquely best match. Under these circumstances, giving each of the donors tied for first place an equal chance for the final match seemed preferable to any further experimentation with weights, match variables, or scores for alpha variables to find the best match.
- The software allowed donor records to be matched to recipient records either with or without replacement of the donor records in the donor pool for subsequent matches.

Example of Matching

Suppose a household member reported a purchase of "AMOX" in a given round of MEPS, and an attempt is made to match a PC record of "AMOXICILLIN" to it. A total weight of 100 is assigned to the pharmacy name match. The group weight of 100 is multiplied by 0.8 if the word is 4 characters long and by 1.0 if the word is 6 or more characters long. In this example, comparing the HC name to the PC name counts 80 toward the final score. Comparing the PC name to the HC name has the potential to contribute 100 to the final score, but because there is no match in this case, it contributes –100 to the final score. In this example, if the medication name is the only match variable, then the final score equals 80–100, or –20, divided by 180, or –.111.

In the same example, if an exact match is required on the GPI code, then this case probably cannot even be in the running for the best match. If an NDC were on the PC record for the "AMOXICILLIN," then a full 14-digit GPI code would have been assigned. With only "AMOX" available on the HC record, at best only 6 of the 14 digits of the GPI could have been coded.

In this example, weighted match variables (match variables not requiring an exact match) for the 2-, 4-, and 8-digit GPI code (with weights of 20, 40, and 80, respectively) could be added to the match variable for the medication name. The total potential score for the match becomes 320 rather than 180 with the addition of the weighted match GPI variables. Assuming that only the first 4 digits of the GPI matched between the HC and PC drug purchase, then the final score for the attempted match becomes –20+20+40–80, or –40, divided by 320, which equals –.125.

NDC Imputations

In order to identify potential outlier drug prices on the PC database, it was necessary to add the average wholesale unit price (AWUP) from the MDDB database to the PC file. This was achieved by performing a crosswalk from the PC file to the MDDB file through the NDC to retrieve the AWUP on the MDDB file. This, however, presented a problem for the 4,604 PC drug events that had missing or unclassified NDC values. In addition, another 8,266 PC drug events had reported NDC values that did not match any NDC on the MDDB database. For these reasons, it became necessary to impute NDC values to these 12,870 records.

All but 498 of the 12,870 PC drug events that could not be linked by NDC to the MDDB data had previously been assigned a GPI code. For the 12,372 cases with a GPI, the matching software was utilized to find the best match to a drug product on the MDDB based on the GPI and the medication name on the PC file. Because a single GPI may cover multiple drug products with differing NDC values, medication names from the PC file were also used as match variables against both the generic and trade/brand names of drugs on the MDDB file. Quantity units, strength, and strength units were also available from the MDDB file and were used as match variables.

For 10,462 of the 12,372 PC drugs with GPI codes, NDC values from reasonable matches to the MDDB file were imputed to the PC file by requiring an exact match to the full GPI and by also using the medication name as a match variable. Another 1,894 drug events were imputed by an exact match to the first 8 characters of the GPI and by again using the medication name as a match variable. The 16 PC drug events with a GPI that did not match in either of the imputation runs above were sent through a matching routine without requiring an exact match on any portion of the GPI. For 9 of these cases, the best match selected by the software was used. Three of the remaining 7 cases were hand-matched to one of the top 10 choices from the software output. Four cases remained. For 2 records the NDC was hand-coded to a specific NDC and for 2 others a matching program was run to find the NDC matching "orthonovum."

For the 498 PC drugs lacking a GPI code, the 71,159 PC drugs with a valid reported NDC were used as the donor base for the matching software. Match variables included the person's age, sex, geographic division, MSA status, health conditions, potential payment sources, and SF/NSF status. Most of the 498 PC drug events lacking a GPI code were also missing the drug name, as well as specific information regarding the quantity and strength of the drug. This explains why these drugs were never initially assigned a GPI code. Match donors from the PC file were restricted to those with non-missing drug product names. The medication names, along with the GPI and quantity and strength information from the donor record, were merged into the recipient drug record whenever this information, in addition to the NDC, was missing from the recipient record.

The quality of the matches was not as high for the 2,408 PC drugs for which there was not an exact GPI code match between donor and recipient. Therefore, these PC drugs were later removed from the donor pool for all matches between HC and PC drug events that were run with replacement. The initial matches, as discussed below, were run without replacement to identify exact matches between HC and PC drug events for the same person in the same round of the survey. The only drug events in the donor pool for these matches were whatever drug purchases were reported by the person's pharmacy provider(s) in the PC.

Other PC Imputation

In order to screen for drug price outliers, a retail unit price (RUP) was constructed from the Pharmacy Component data and compared against the AWUP taken

from the MDDB file. The RUP was constructed as the retail price for the drug product, defined as the sum of payments for the prescription divided by the dosage amount or quantity dispensed as reported by the pharmacy provider. For 75 of the 84,029 PC drug events, the amounts dispensed were missing. The matching software was used to impute these missing quantities. Match variables included the NDC and GPI for the drug product and the person's age, sex, health conditions, and health status. Exact matches for the first eight characters of the GPI were required and heavier weight was placed on the NDC, followed by the GPI.

For 211 drug products, reported dosage amounts contained more than one value. Values were typically separated by a slash or dash on the computer printouts supplied by the pharmacy providers. Pharmacy consultants to Westat, Mediquest Associates, provided technical assistance in establishing editing rules to determine the appropriate value to use for the dosage amount in these cases.

For 206 drug products, the round in which the prescription was purchased was missing from the PC file. This variable was initially prepared for the PC flat file, but because of missing month (168 events), day (173 events), and/or year (14 events) information for certain prescriptions on the file, the round variable was missing. The matching software was used to impute the round for 110 events, using the person identifier as an exact match and the NDC and GPI as weighted match variables. The remaining 96 events were imputed a value for the round by a second application of the matching software. During this second application, the variables for the year at the beginning and the end of the round were exact matches and the year at the beginning of the first round was a weighted match variable.

The AWUP variable was then merged onto the PC file from the MDDB database by NDC value. Up to six AWUP values, measured at different times, were provided on the MDDB file. Prices dated the closest to the middle of 1996 were selected, but in some cases only prices for years prior to or after 1996 were available. Distributions of the ratio of RUP to AWUP (called PRATIO) for cases in which the merged AWUP was dated before 1996, in 1996, and after 1996 were analyzed to determine whether adjustments were needed for the outlier analysis. Depending on the measure used for the price variable on the PC file (described below), AWUP prices before 1996 were inflated by either 10 or 40 percent to produce a PRATIO distribution comparable to one using only 1996 AWUP values. Similarly, certain AWUP prices after 1996 were deflated by either 10 or 35 percent to produce a PRATIO distribution comparable to one using only 1996 AWUP values.

For example, PRICE3 is defined below as the total charge reported by the pharmacy provider. It serves as the prescription price on the PC when payment information is missing for a given prescription purchase. The PRATIO distribution when only 1996 values of the AWUP are used indicates that about 17 percent of the prescriptions with the total charge as the drug price have a retail drug price below 80 percent of the AWUP. The PRATIO distribution when only AWUP prices before 1996 are used shows that 23 percent of the prescriptions with the total charge as the drug price have a retail drug price below 80 percent of the AWUP. Inflating each pre-1996 AWUP by 10 percent for these cases is equivalent to lowering a PRATIO value of .8 to approximately .7. The PRATIO distribution when only AWUP prices before 1996 are used shows that about 17 percent of the prescriptions with the total charge as the drug price have a retail drug price below 70 percent of the AWUP. The price deflation in this case produces a price distribution of total charge relative to standardized price (the AWUP) for AWUP values before 1996 that compares favorably to a similar relative price distribution using 1996 AWUP values.

Outlier Editing

After the preliminary editing described above was performed on the PC file, three different measures of the retail price of the drug product were constructed:[2]

PRICE1 = the sum of payments when no payments were coded as missing (68,664 events)

PRICE2 = the sum of payments when at least one payment is reported greater than zero and at least one payment is reported as missing (8,030 events)

PRICE3 = the reported total charge for the drug product when reported to be greater than zero and all payments are reported to be either missing or zero (2,141 events)

The remaining 5,194 drug products on the PC file (referred to as PRICE4 cases) lacked any positive-valued payments or total charge data.

[2] Recall that from rule P2 (Preliminary PC Event Edits section), the total charge was set equal to the sum of payments when the difference between the two was no greater than $2 and none of the payments or the total charge was missing.

Based on PRICEn values (n = 1,2,3), the 78,835 drug events with a positive-valued price were sorted into three groups for outlier analysis:

PRGRP1n = 1 if PRICEn = self-payment and PRICEn ≤ $30 (32,528 events)
= 0, otherwise
PRGRP2n = 1 if PRICEn = self-payment and PRICEn > $30 (7,788 events)
= 0, otherwise
PRGRP3n = 1 if PRICEn does not equal self-payment (38,519 events)
= 0, otherwise

The 78,835 drug products were next allocated to three outlier groups based on PRATIOn, the ratio of retail unit price (RUPn) to AWUP (using PRICEn, n = 1,2,3 and dosage amounts to construct RUPn), as follows:

OUT1n = 1 if PRATIOn < .8 (22,002 events)
= 0, otherwise
OUT2n = 1 if .8 ≤ PRATIOn < 20 for n = 1 and AWUP before 1996 *or* if .8 ≤ PRATIOn < 10 otherwise (56,293 events)
= 0, else
OUT3n = 1 if PRATIOn ≥ 20 for n = 1 and AWUP before 1996 or if PRATIOn ≥ 10 otherwise (540 events)
= 0, else

The thresholds for determining unit price outliers were established after consulting with Mediquest Associates, Inc., Westat's pharmacy consultant experts for the study.

An additional 14 edit rules for the prices and payment sources on the PC file were implemented based on the PRICE, PRGRP, and OUT classifiers. These edit rules, described below, are labeled Rule P5 to Rule P18. Outlier cases were edited by setting the outlier RUP equal to the AWUP. Tabulations of the PRATIO distribution for non-outlier prescriptions verified that the AWUP was the modal value for the distribution.

Rule P5: PRICE1/PRGRP11, PRGRP21, and PRGRP31/OUT21 (53,351 events).

This edit rule stated that when no payment sources were missing and when the drug price was equal to a positive-valued sum of payments and not an outlier, no edit was required. Payment shares were constructed for these cases for use in implementing editing rule P7.

Rule P6: PRICE2/PRGRPm2/OUT22 (m = 1,2,3) (1,261 events).

This edit rule stated that if at least one payment source was missing but the others summed to a positive amount, then no edit was required if the sum was not an outlier price, unless the sum was less than the total charge. In the latter case, the drug price was set equal to the total charge. The new price was checked to make sure it was not still an outlier price. If it was an outlier price, then it was reclassified and edited by another appropriate rule (P15 or P17). If it was not an outlier price, then a single missing payment source was set equal to the difference between the total charge and payment sum. If more than one source was missing, a hierarchy was established for deciding which missing source would be allocated the difference. In the hierarchy, Medicare was allocated as payer if the person was 65 or older and was in a Medicare HMO; otherwise, Medicaid was allocated as payer if the person had Medicaid coverage; otherwise, private insurance paid the difference if the person had private coverage; and so on. After editing, payment shares were constructed for each drug event for use in implementing editing rule P7.

Rule P7: PRICE3/PRGRP33/OUT23 (1,681 events).

This edit rule was applied to drug events for which the drug price equaled a positive-valued total charge variable. The payment sum was zero or missing because no payment sources were positive valued. When the drug price (total charge) was not an outlier, payment source amounts were assigned by payment shares merged onto the event record by matching to another non-outlier drug event record (from P5 and P6) containing clean payment amounts. The exact match variables were the potential payment sources, a variable for whether the person had private prescription drug coverage, and the person's sex. The weighted match variables were the drug product's GPI, NDC, and price, and the person's age, region, and MSA status. Before matching, the potential payment source flags were set to indicate coverage from a given source if the payment amount for the drug product on the PC record was greater than zero.

Rule P8: PRICE4 (cases lacking any positive-valued payments or total charge data) (5,194 events).

In cases in which the total charge was missing or zero and there were no positive-valued payment amounts, the drug price and payment source amounts were imputed to the PC drug event record by a match to a donor drug event record. The donor records consisted of PC drug events in which at least one payment amount was positive valued and the drug price was not an outlier

(P5 and P6 events). Initially 1,447 of these cases were matched by using exact match variables for the NDC and GPI of the pharmaceutical, the potential payment source flags, and an indicator of private prescription coverage for the person. Weighted match variables included the drug name, the pharmacy name, and the person's age, sex, geographic region and division, and MSA status. Because of the high quality of these matches, the donor record's dosage amount was also imputed to the P8 recipient record to avoid creating new price outliers unnecessarily. No replacement for the recipient dosage amount was imputed in the remaining matches because the quality of the matches was not as good. In a second match, the first six GPI characters replaced the full GPI and full NDC as exact match variables from the first match, and sex was used as an exact match variable. This produced 2,397 more matches. An additional 602 P8 records were matched in a third run, in which the first two GPI characters replaced the first six GPI characters from the second run as an exact match variable. All but 25 of the 748 remaining unmatched P8 cases were matched in a fourth run, in which the GPI was omitted altogether as an exact match variable. The remaining 25 cases were hand edited by assigning them a retail drug price equal to the product of the AWUP and the dosage amount from their own record. Based on the name of the pharmacy provider, which identified it as a Veterans Affairs (VA) or military pharmacy provider, and the potential coverage variables, the entire imputed purchase price of each of the 25 drug products was assigned to either VA or CHAMPUS coverage.

After all prices were imputed to P8 cases, 1,838 new outliers were created by combining the prescription price from the donor record and the dosage amount from the recipient record of the 3,747 drug events that failed to match in the initial run of the matching software. These 1,838 outlier events were further edited by one of the five remaining "PRICE1" edit rules (i.e., Rules P9-P13).

Further analysis of the names of the pharmacy providers for P7 and P8 drug events showed that a substantial majority of them appeared to be VA, CHAMPUS, or "other Federal/Indian Health Service" providers. These pharmacy providers generally do not know the charges and payments for specific prescriptions. At this point, any P7 or P8 drug event in which the words VETERAN, VETERANS, VETRANS, VA, or VAMC appear in the name of the pharmacy provider and the person reported VA as a potential coverage source was assigned VA as the sole payer of the purchase price of the drug product. Any P7 or P8 drug with ARMY, AFB, NAVY, NAVAL, NAVCARE, MARINE, USAF, or AIR FORCE in the pharmacy provider name and CHAMPUS as a potential coverage source was assigned CHAMPUS as the sole payer of the purchase price of the prescription product. If the person indicated "other Federal," and not CHAMPUS, as a potential payment source, then the full amount paid was assigned to other Federal rather than CHAMPUS. Finally, if INDIAN or TRIBAL (other than INDIAN RIVER or INDIAN TRAILS) appeared in the pharmacy name of a P7 or P8 prescription and other Federal was reported as a potential payment source, then the full purchase price of the drug product was assigned to other Federal.

Rule P9: PRICE1/PRGRP11/OUT11 (8,121 events).

This edit identified cases in which the RUP was under 80 percent of the AWUP, self-payment was the sole positive payment amount reported, no other source had a missing code, a corresponding potential source was indicated, and the self-payment was $30 or less. In the edit the RUP was increased to the AWUP. The size of the self-payment in these cases suggested a copayment situation. Therefore, the difference between the new and old retail price for the drug was allocated hierarchically to any potential third-party sources indicated for the person. If none was indicated, then the full retail price increase was allocated to self-payment.

Rule P10: PRICE1/PRGRP21/OUT11 (657 events).

This drug edit identified outlier drug prices in which the RUP of the drug was under 80 percent of the AWUP. The RUP of the drug was edited by increasing it to the AWUP. The resulting increase in the retail price of the drug was allocated entirely to self-payment because it was the only payment source indicated by the person purchasing the drug and because the reported amount was over $30 and not considered a potential copayment case.

Rule P11: PRICE1/PRGRP11 and PRGRP21/OUT31 (231 events).

In these cases, the reported RUP of the drug equaled or exceeded 10 times the AWUP (or 20 times the AWUP if the AWUP was measured before 1996). In this edit, the reported RUP was reduced to the AWUP. The subsequent reduction in the retail price of the drug

was taken entirely out of the self-payment amount because it was the sole payment source with a positive amount reported for the individual.

Rule P12: PRICE1/PRGRP31/OUT11 (7,575 events).

For cases in which the RUP was less than 80 percent of the AWUP, the reported RUP was increased to the AWUP. A hierarchy was established for allocating the increase in the retail price of the drug to a single positive-valued payment source because there might have been more than one payment source with a positive amount reported. In the hierarchy, any third-party payer had a higher priority than a self-payer.

Rule P13: PRICE1/PRGRP31/OUT31: (567 events).

For cases in which the RUP was greater than or equal to 10 times the AWUP (or 20 times the AWUP if the AWUP was measured before 1996), the RUP was decreased to the AWUP. A hierarchy was established for allocating the decrease in the retail price of the drug to a single payment source because multiple sources might have been reported with positive amounts. If the entire decline in the retail price was not used up in these cases by one payment source, the remainder was taken from the next source in the hierarchy. This allocation process continued until the entire price reduction was fully allocated.

Rule P14: PRICE2/PRGRP12 & PRGRP22/OUT12 (6,726 events).

In this case, at least one potential third-party payment source was missing and the self-payment amount, the only positive-valued source of payment reported, initially was set equal to the retail drug price. If the implied RUP was less than 80 percent of the AWUP, then the RUP was inflated to the AWUP. The difference between the new and old retail drug price was allocated entirely to the missing third-party payer if only one was missing. If more than one third-party payer was missing, then a hierarchy was established to allocate the entire difference to one of the missing payers.

Rule P15: PRICE2/PRGRP12 & PRGRP22/OUT32 (15 events).

In this case, the RUP equaled or exceeded 10 times the AWUP. The self-payment equaled the retail price of the drug, although at least one third-party payment was reported as missing. The RUP was reduced to the AWUP, and the reduction in the retail price of the drug was taken entirely out of the out-of-pocket payment reported by the pharmacy provider.

Rule P16: PRICE2/PRGRP32/OUT12 (25 events).

In this case, self-payment was not the sole positive-valued payment amount, but there was at least one missing payment amount and the RUP was less than 80 percent of the AWUP. The RUP was inflated to the AWUP, and the retail drug price increase was allocated to the missing payment source if only one source was missing. A hierarchy that included self-payment was used to allocate the drug price increase to missing payment sources when more than one source was missing.

Rule P17: PRICE2/PRGRP32/OUT32 (3 events).

In this edit, the RUP equaled or exceeded 10 times the AWUP. At least one potential payment source was missing but at least one was positive valued. Self-payment was not the only positive-valued source reported. The RUP was deflated to the AWUP, and the reduction in the retail drug price was removed from reported positive-valued payment sources in the same way as in editing rule P13.

Rule P18: PRICE3/PRGRP33/OUT13 and OUT33 (460 events).

In this outlier edit, the retail price was the reported total charge when it exceeded zero and no reported payment amount was positive valued. The RUP either was less than 80 percent of the AWUP, or it equaled or exceeded 10 times the AWUP. The RUP was set equal to the AWUP, and the new price was allocated to potential payment sources by payment shares merged into the event record by matching to a non-outlier drug event record containing clean payment data.

The sum of the records affected by PC edit rules P5 through P18 is 85,867, 1,838 more than the 84,029 PC drug events. The 1,838 difference between these totals represents the Rule P8 cases in which an imputed prescription price, combined with the dosage amount on the original record, produced an outlier requiring additional editing by one of the PRICE1 rules, P9 through P13.

After these edit rules were implemented, analysis of the results revealed that for 530 OUT1 cases, the edited retail price for the drug product was inflated to be inordinately high (in excess of $200) compared to its original retail price, and for 697 OUT3 cases, the edited retail price for the drug product was deflated to be inordinately low (under $2) in comparison to its original retail price. The majority of these cases occurred when NDC values were imputed for PC drug products and created mismatches among retail drug prices, dosage

amounts, and the AWUP to create the outlier PRATIO values. Because most of these cases were not going to be used for imputation to household events, they were edited by using the original retail price reported by the pharmacy provider in combination with an edited dosage amount for these 1,227 PC events.

Self-Filer Matching and Imputation

After cleaning the PC database, the next step was to match or impute PC data to the HC drug mentions of SFs by using the matching software in a series of applications. The first matching run was set up to identify exact matches between prescription events mentioned by the household respondent and reported by the individual's pharmacy provider. For this task, 6,356 PC potential donor events were tested for exact matches to 8,223 HC prescription mentions (not including 361 free samples) for SF households that were in the linked pharmacy followback survey. The exact match variables in the first run were the person's identification number (PERSID), the GPI, and the round in which the drug was purchased. The weighted match variable was the medication name supplied by the household and by the pharmacy provider. In the first run, 512 HC prescription mentions were exactly matched to prescriptions reported for the same person in the same round in the pharmacy database. Once a match was made, the PC donor was effectively removed from the donor pool and not matched with any other HC drug mentions by the individual (i.e., the matches were made without replacement).

In a second run, exact matches were identified for an additional 4,495 HC prescription mentions for SFs by requiring exact matches only for the PERSID and the round, and by using weighted match variables for the medicine name and the first 2 consecutive characters, first 4 consecutive characters, first 8 consecutive characters, and first 10 consecutive characters of the GPI code. After being reviewed, only 3,446 of these matches met the final criteria for an exact match (a match score greater than zero, an exact match on the first 4 consecutive characters of the GPI code, or an exact match on the medicine name). The quality of the 1,049 matches that did not meet these criteria was deemed too low to be classified as an exact match.

Another 2,028 HC drug mentions of SFs were "refills" associated with one of the 3,958 exactly matched HC drug mentions and were matched to the same PC drug donor event as the exactly matched event. (Although the term "refill" is used in this context, the MEPS HC did not collect sufficient information to determine which drug acquisitions were original prescriptions and which were refills in a given round of the survey.) Refill drug records for an individual on the HC file were linked through a common, round-specific event identification number (EVNTID). In case two or more HC prescription records with the same EVNTID were exactly matched to two or more separate PC records, any other unmatched refills with the same EVNTID were matched by random selection to one of the exactly matched PC events. This left 2,237 unmatched HC prescription mentions by SFs who participated in the MEPS PC, in addition to 6,888 prescription mentions (not including 332 free samples) for SFs not in the PC that remained unmatched.[3]

The remaining 9,125 unmatched HC-reported prescriptions were eventually imputed data from a PC prescription drug through a series of matching runs that included all 81,621 PC drug records in the donor base that either did not require any NDC imputation or had an NDC imputed with an exact GPI match. In these matches, if a PC drug was selected for an imputation, it went back into the donor pool and was made available for subsequent imputations (i.e., the remainder of the SF imputations were done with replacement).

The 9,125 unmatched HC drugs contained 4,066 unique EVNTID values. Each EVNTID contained at least one unmatched non-free prescription or prescription refill. Of these EVNTIDs, 447 were imputed PC records by using the GPI as an exact match variable. Weighted match variables, in descending order of the weights, included the medication name; potential payment sources and private prescription coverage indicators for the person; name of the person's pharmacy provider(s); and the person's age, sex, condition codes, geographic region and division, MSA status, employment status, and self-reported health status. PC records were imputed based on a match on the first 10 digits of the GPI code for 2,931 EVNTIDs, based on a match on the first 8 digits of the GPI for 19 EVNTIDs, based on a match on the first 4 digits for 323 EVNTIDs, based on a match on the first 2 digits for 110 EVNTIDs, and based on a match on no portion of the GPI for 236

[3] A total of 15,111 HC drug mentions of SFs were imputed PC prescription data in this part of the editing. Initially, there were 15,804 HC drug mentions of SFs. The difference of 693 between these two totals represents SF drug acquisitions identified as free samples.

EVNTIDs. All HC drug mentions with an EVNTID in common were imputed values based on the match to a single PC donor drug product.

Self-Filer Payment Reconciliation

For SF households, charge and payment information after matching or imputation was available from both the self-reported household data and from the data merged onto the prescription drug record from the PC prescription donor record. The financial data were reconciled as follows:

- If none of the household-reported payment amounts were missing and the HC payment sum equaled the PC payment sum, then the HC payment amounts were used along with the HC payment sum (63 events).
- If none of the household-reported payment amounts were missing and the HC payment sum was greater than zero but did not equal the PC payment sum, then the PC payment amount was used and the payment amounts were allocated according to the shares for each amount based on the HC data (10,984 events).
- If the HC payment sum was missing because at least one payment source was missing, then the PC payment sum/price was used along with the PC payment amounts if all of the HC payment amounts were missing. If at least one of the HC payment amounts was not missing and the sum of the non-missing HC payment amounts exceeded the PC payment sum, then the HC payment amounts were used after scaling by the ratio of the PC payment sum to the "partial" HC payment sum. If at least one of the HC payment amounts was not missing and the sum of the non-missing HC payment amounts was less than the PC payment sum, then the difference was allocated to the missing HC payment source. If more than one HC payment source was missing, then the difference was assigned hierarchically in the following descending order: Medicare HMO, Medicaid, private insurance, VA, CHAMPUS, other Federal coverage, State or local coverage, Worker's Compensation, other insurance, or self-payment (3,525 events).
- If the HC payment sum was zero with no missing payments indicated, then the PC payment sum and amounts were used (539 events).

In general, an attempt was made to retain as much information as possible regarding source-of-payment shares from the household-reported data. However, the retail drug price information, which had been edited for price outliers on the PC database, was provided from the PC data. Using the PC retail prices instead of the HC payment sums avoided doing additional editing of potential HC outlier prices and kept detailed drug identification and pricing information from the PC intact when it was imputed to the HC drug mentions.

Non-Self-Filer Matching and Imputation

Data on matching and imputation for NSF information are shown in Table 2. The procedure used to match HC prescription mentions for NSF households to PC prescriptions mirrors the procedure described above for SFs. As with SFs, the first set of matches was designed to find exact matches between the HC and PC drug events. The recipient HC database for these matches contained 108,353 prescription mentions, while the donor PC database consisted of 72,615 drug product purchases for NSFs. Drug mentions that had been identified as free samples also were included in the HC recipient group of drug mentions for NSF exact matches because of the rather arbitrary way in which a specific drug mention was identified as a free sample for the NSF population. (See Preliminary HC Event Edits section.)

Even when only one prescription was reported for a medication in a given round and the household reported receiving a free sample of the drug, it was not clear whether the free sample was reported as the single drug acquisition. This was not clarified in the interview.

In the first exact matching run, 5,950 of the HC drug mentions for NSF households in the linked PC followback were matched exactly and without replacement to PC drug mentions by using the PERSID, the complete GPI code, and the round of the drug as exact match variables, and the medication name as the weighted match variable. In the second run for exact matches, 50,566 additional HC drug mentions were matched without replacement to PC drug records; only PERSID and the round were required to match exactly, and the first 2-character, 4-character, 8-character, and 10-character GPI codes were used as weighted variables. As in the SF matching, not all second-run matches were retained. Only 37,405 of these NSF matches in the second run were deemed to be exact matches because

they either had a match score greater than zero, matched exactly on the first four characters of the GPI, or matched exactly on the medication name. Of the 2,279 HC drug mentions designated as free samples for NSF households in the PC followback, 462 were selected as exact matches and are included in the count of the 43,355 exact matches.

Next, an additional 23,838 unmatched drug refill mentions by NSF household respondents in the PC were matched to one of the 43,355 PC drugs that had been exactly matched to a purchase of the same drug by the same person in the same round. This produced 67,193 matches to the original 188,422 drug mentions by NSF household respondents. Of the remaining 121,229 unmatched NSF drug mentions, 3,790 were set aside as unmatched free samples, leaving 117,439 HC drug mentions unmatched for all NSF households (39,343 for households in the PC and 78,096 for households not in the PC), representing 47,705 unique EVNTID values. These 47,705 EVNTIDs were eventually imputed values from PC prescription drugs through the same general series of matching runs without replacement as for SFs, described above. The donor pool for these matches consisted of 81,621 PC prescriptions, the full 84,029 sample minus the 2,408 drugs with an imputed NDC that had no exact match to the GPI code.

In the first NSF imputation without replacement, which required an exact match with the full GPI, 5,624 of the 47,725 EVNTIDs were matched. In the subsequent imputations, 32,066 of these EVNTIDs matched on the first 10 digits of the GPI; 274 matched on the first 8 digits; 4,746 matched on the first 4 digits; 1,854 matched on the first 2 digits; and 3,161 did not match on any portion of the GPI. After the matches were linked to every unmatched refill with the same EVNTID, all 117,439 previously unmatched HC drug mentions of the NSF population were imputed a PC prescription record. Because no charge and payment data were collected in the HC for the NSF population, there was no need to reconcile charge and payment data from the two sources after the matching and imputation.

Unmatched PC Data

In the PC donor database of 81,621 drugs (all 84,029 PC drugs less the 2,408 excluded from the matches with replacement), over one-third, or 29,871 drugs, were never imputed or matched exactly to an HC

Table 2. Matching and imputation for non-self-filers in the 1996 Medical Expenditure Panel Survey

Number	Description
188,422	HC drug mentions by NSFs
80,069	HC drug mentions by NSFs without PC data
108,353	HC drug mentions by NSFs with PC data
43,355	Exact matches of HC drug mentions by NSFs to PC data (includes 462 free samples that were exactly matched to PC drugs)
23,838	"Refills" (additional acquisitions) of exact matches of HC drug mentions by NSFs
41,160	HC drug mentions by NSFs with PC data still unmatched
121,229	HC drug mentions by NSFs with and without PC data still unmatched (includes 3,790 free samples that remain unmatched)
117,439	HC drug mentions by NSFs that were imputed a PC drug in matches "with replacement" (excludes 3,790 unmatched free samples)

Note: HC is the Household Component of the Medical Expenditure Panel Survey (MEPS). NSF is non-self-filer (someone whose pharmacy automatically sends the insurance company claim form for prescriptions or handles third-party payments electronically at the point of sale). PC is the Pharmacy Component of MEPS.

Source: Center for Cost and Financing Studies, Agency for Healthcare Research and Quality: Medical Expenditure Panel Survey, 1996, public use file HC-010A.

drug mention. Because this number is close to the 25,666 refills of household drug mentions that were exactly matched to PC prescriptions, an exercise was undertaken to determine how many of the unimputed, unmatched PC drug purchases could be matched to the imputed, matched PC drug purchases by NDC, GPI, and medication name.

Results of this exercise confirmed the hypothesis that the vast majority of the unmatched PC drug purchases were duplicates, or refills, of PC drug purchases that were matched to HC drug mentions. Using NDC and GPI as exact match variables, and the medication name, prescription price, and potential payment sources (in descending order) as weighted match variables, 25,908 of the 29,871 unmatched PC drug purchases matched to one of the 51,750 PC donor drug purchases. Of the remaining 3,963 unmatched PC drug purchases, 3,699 matched to a PC donor by using only the GPI as an exact match variable and the same partial match variables as before. All but 2 of the remaining 264 unmatched PC drug purchases were ultimately matched to one of the PC donor purchases by successively reducing from 8 to 2 the number of digits in the GPI that were required to match exactly. The final two PC drug purchases were matched to a PC donor by not requiring any exact match variables and using only partial match variables for the drug name, prescription price, and potential payment sources.

The average prescription price for PC drug purchases that were not matched or imputed to any HC drug mentions was $32.87. This was close to $5.00 less than the average price of PC drug purchases that were matched or imputed to HC drug mentions one or more times ($36.53). Of the latter 51,750 PC drug purchases, 2,794 were matched or imputed five times or more to HC drug mentions. These drugs had an average prescription price of $31.51. The 48,956 PC drugs that were matched or imputed to HC drug mentions from one to four times had an average prescription price of $36.82. The PC drug purchases that were not matched or imputed to HC drugs had slightly lower average out-of-pocket and private insurance payment shares (38.3 and 39.9 percent, respectively) than those of PC drug purchases that were matched or imputed at least once (40.6 and 41.7 percent, respectively). PC drug purchases that were not matched or imputed to HC drug mentions also had slightly higher Medicaid and other payment shares (14.8 and 7.0 percent, respectively) than those of PC donor drug purchases (12.0 and 5.7 percent, respectively). This suggests that if all of the PC prescription drugs had been matched or imputed to HC drug mentions, the average HC prescription price and aggregate prescription expenditures, out-of-pocket payments, and private insurance payments would have been slightly lower, and aggregate Medicaid and other public prescription expenditures would have been slightly higher.

Allocation to 1996 and 1997

After all SF and NSF prescription mentions in the MEPS HC were matched to PC prescriptions, Round 3 drug mentions that were not exactly matched to PC drug products were allocated to either 1996 or 1997. As mentioned earlier, the Round 3 survey spanned both years for respondents, and reported prescription purchases in the HC were not dated within the round unless they were matched exactly to a PC drug record. Each Round 3 drug mention that was not exactly matched was allocated to 1996 or 1997 by using the beginning and ending dates of Round 3 for each person and the percentage of Round 3 covering 1996 and covering 1997, in combination with a random draw between 0 and 1.

For the HC population, 12,923 of the 77,110 Round 3 prescriptions were exactly matched to PC drugs and classified as 1996 drug purchases. Of the remaining 64,187 prescriptions, 52.0 percent (33,394 prescriptions) were classified as 1996 drug purchases using the random allocation method, while the other 48.0 percent (30,793 prescriptions) were classified as 1997 drug purchases.

Sensitivity Testing

The edit rules imposed on the PC donor database described above, on balance, were likely to increase the average retail price of prescriptions because substantially more PC drugs were outliers at the low end of the distribution (22,002 PRATIO values below .8) than at the high end (540 PRATIO values above 10 or 20). Of particular concern were the outlier cases at the low end of the distribution that were classified as not missing any payment sources (PRICE1 cases). Furthermore, there was some evidence to suggest that, in certain situations, the outlier PRATIO value may have been caused by a misreported dosage amount rather than a misreported price in the PC. To analyze the sensitivity of national estimates of prescription expenditures to the outlier edits,

alternative estimates were made by successively undoing various outlier unit price edits that had been performed on the PC data.

The national estimate of prescription expenditures in 1996 for the MEPS population before any changes were made to the edit rules described above was $75.4 billion. The edits reversed were all edits to drug prices reported without any missing payment sources (PRICE1 cases). A hierarchy was established for undoing the edits such that the most likely candidates for a quantity edit rather than a price edit had their price edits reversed first. Reversing the 7,575 lower end edits under rule P12, in which no payments were missing and the original retail price/payment sum did not equal the out-of-pocket payment, lowered the national prescription expenditure estimate by $3.7 billion to $71.7 billion. Reversing the 657 lower end edits under rule P10, in which no payments were missing and the original retail price/payment sum equaled the self-payment but was greater than $30, lowered national prescription expenditures by another half billion dollars to $71.2 billion.

The final lower end edit that was reversed when no payment sources were reported missing involved the 8,121 rule P9 cases in which the retail price/payment sum equaled the self-payment amount and was less than or equal to $30. These are more likely to be outlier retail price cases because the low price and 100 percent self-payment suggest a copayment coupled with a unit retail price less than 80 percent of the AWUP. Undoing this edit reduced the national prescription expenditure estimate by another $4.6 billion to $66.6 billion.

Finally, undoing the two higher end price edits in which no payment sources were missing for the 798 events edited under rules P11 and P13 increased the national prescription expenditure estimate by $.1 billion to $66.7 billion.

As suspected, the PC pricing edit rules on the linked pharmacy followback data for cases in which no payment source amounts were reported missing had an impact on the national prescription expenditure estimate from the MEPS data. For the public use data (the 1996 Prescribed Medicines File, HC-010A), the middle estimate of $71.2 billion was selected. For this estimate, pricing edit rules P9, P11, and P13 were left intact, but pricing edits P10 and P12 were replaced by quantity/dosage amount edits. New dosage amounts were imputed by dividing the original retail price of the drug by the AWUP. Price changes for PC prescriptions that were matched to SFs required that the SF HC-reported charge and payment data be reconciled to the new PC price and payment data before finalizing expenditure and payment amounts for the public use file.

Consistency Edits

As discussed above, imputations of PC data to HC drug mentions that were not exactly matched to PC drug purchases, or refills (additional acquisitions) thereof, were primarily based on match variables for the GPI code and the medication name. Additional weighted match variables included potential third-party payment sources, but these were not required to be exact-match variables in the imputations because of small cell sizes. As a result, 16,829 HC drug purchases for 1996 had at least one inconsistent third-party payment source. These were defined as imputed payments from a given third-party source that was not indicated by the individual in the HC as a potential payment source. For SFs only, this source also had to be coded as "not a known payer" in a second set of variables to be classified as "inconsistent."

For the 16,829 drug purchases with inconsistent imputed payment sources, a hot deck was run in which the 148,644 purchases in 1996 with consistent payments were used as the donor group for the hot deck. The class variables selected for the hot deck were the primary payer and the price category for the drug purchase. The primary payer was determined for the recipient group according to potential payment sources in the following hierarchical order: Medicare; private insurance; CHAMPUS; Medicaid; VA; other Federal; State and local; other; Worker's Compensation; out of pocket or self-payment. The same order was applied to donors, although in their case, the order referred to actual payments from a source rather than variables indicating potential payment sources. Then six categories of the primary payer variable were constructed as the class variable for the hot deck: private; Medicaid; Medicare; other public (CHAMPUS, VA, other Federal, State and local, and Worker's Compensation); other; and self or family. The price category was divided into four categories: greater than zero and up to $15; $15.01 up to $30.00; $30.01 up to $100.00; and over $100.00. Sort variables for the hot deck consisted of six variables indicating whether each

of the six categories of the primary payer variable was a consistent payment source.

The hot deck imputed percentage shares to each recipient drug purchase. The percentage shares came from each of the six payment source categories from the donor record. They were used to allocate the drug price of the recipient drug to each source. If the "other public" share from a donor was non-zero and only one of the five sources in this group was a consistent source, then the full share was allocated to that source. If more than one of the five sources in the "other public" category was a consistent source, then the full share was allocated according to the hierarchy above for choosing among the five sources in this category.

The March 2001 Revision

In the late fall of 2000, records not included on the original public use version of file HC-010A were identified with missing values of the variable MEDCYCLE, which indicates the number of times a drug was purchased during a round. In addition, it was discovered that the value of MEDCYCLE had been misreported in the survey for certain medications on the original HC-010A file. A revised public use version of the 1996 prescription drug event file, which corrected for the missing and misreported MEDCYCLE values, was released in March 2001. This section of the report briefly describes the modifications made to the previously released public use version of the HC-010A file in May 2000 and the consequent impacts on the size of the file and on national estimates of prescription drug utilization and expenditures.

To identify misreported MEDCYCLE values, a variable called MEDRATE was constructed as the days in a round divided by MEDCYCLE. MEDRATE indicates the maximum value of the average number of days between refills of a prescription drug within a round for an individual. This estimate is considered to be a maximum value because it assumes that the person purchased the drug on the first day of the round. Identifying values of MEDRATE of 3 or less enabled AHRQ staff, in consultation with a pharmacy expert, to eliminate all of the implausible MEDCYCLE values. Certain values of MEDCYCLE with values of MEDRATE at or below this threshold were left intact because a low MEDCYCLE value was combined with a small number of days in the MEPS round; these were plausible cases, according to the pharmacy expert. Cases identified as misreported MEDCYCLE values were imputed new values drawn at random from the distribution of valid MEDCYCLE values for drugs with the same GPI code.

Moreover, as part of the revision to the HC-010A file, all previously identified free samples were reclassified as purchases because respondents were not specifically instructed to include counts of free samples in reporting the survey data that produced the MEDCYCLE variable. Also, a revision was made to the allocation of Round 3 HC prescription mentions between 1996 and 1997 to incorporate information regarding the year in which a sampled individual first started taking the prescription drug (RXBEGYR). With the exception of exactly matched HC drug mentions, all Round 3 HC drug mentions with MEDCYCLE = 1 and RXBEGYR = 1996 were allocated to 1996, and all Round 3 HC drug mentions with RXBEGYR = 1997 were allocated to 1997 regardless of the MEDCYCLE value. All other Round 3 HC drug acquisitions that were not exact matches but had been assigned new MEDCYCLE values were allocated to 1996 and 1997 based on the proportion of time the sampled person was in Round 3 in each year.

Any 1996 HC drug acquisitions added to the revised 1996 HC-010A file that had been omitted from the original file were imputed prescription drug information from the PC by hot decking to HC drug purchases on both the revised and original files that had been previously matched to PC drugs. Class variables for the hot deck included the GPI code for the drug, filer status (SF or NSF), age, sex, and Medicaid and private drug insurance status. The sort variables included the remaining insurance status variables, HMO status, and region.

The results of the March 2001 revisions to the 1996 file are reported in Table 3. Of the original 167,784 drug purchases on the file, 26,368 acquisitions (15.7 percent) were deleted because of misreported MEDCYCLE values. In addition, 4,021 acquisitions with valid MEDCYCLE values (2.4 percent) were reassigned to 1997 because of the revised allocation rules. Another 6,712 acquisitions (4.0 percent) were added to the original file from the newly identified missing MEDCYCLE cases or were reassigned to 1996 from 1997 because of the new allocation rules. Finally, another 3,201 previous free samples on the original file (1.9 percent) became drug purchases in 1996 on the revised file. The March 2001 HC-010A file contains

147,308 drug purchases, representing a net decline of 20,476 purchases, or 12.2 percent, from the original drug file.

National estimates of drug purchases declined from 2.116 billion purchases to 1.865 billion purchases, representing a net decline of .251 billion purchases, or an 11.9-percent reduction. National estimates of prescription drug expenditures for the civilian noninstitutionalized population in the 1996 MEPS HC declined from $71.208 billion to $65.291 billion, representing a net decline of $5.917 billion, or an 8.3-percent decline in national expenditures because of the modifications made to the file.

Reference

Cohen JW, Monheit AC, Beauregard KM, et al. The Medical Expenditure Panel Survey: a national health information resource. Inquiry 1996; 33:379-89.

Table 3. Impact of March 2001 file revision on prescription utilization and expenditures in the 1996 Medical Expenditure Panel Survey

Category	Acquisitions (unweighted)	Weighted acquisitions (billions)	Expenditures (billions of dollars)
Original HC-010A file	167,784	2.116	$71.208
Changes:			
Misreported refills	–26,368	–.328	–8.874
Reallocated to 1997	–4,021	–.049	–1.433
Missing acquisitions	+5,565	+.069	+2.381
Reallocated to 1996	+1,147	+.014	+.350
Free samples	+3,201	+.042	+1.659
Total net change	–20,476	–.251	–5.971
Revised HC-010-A file	**147,308**	**1.865**	**65.291**

Note: Any differences between components and totals are a result of rounding.
Source: Center for Cost and Financing Studies, Agency for Healthcare Research and Quality: Medical Expenditure Panel Survey, 1996, public use files HC-010A.

Appendix A. List of Abbreviations

AHRQ	Agency for Healthcare Research and Quality
AWUP	Average wholesale unit price
CCFS	Center for Cost and Financing Studies
CHAMPUS	Civilian Health and Medical Program for the Uniformed Services
EVNTID	Event identification number
GPI	Generic Product Identifier
HC	Household Component (of MEPS)
HCFA	Health Care Financing Administration
HMO	Health maintenance organization
MDDB	Master Drug Data Base
MEDCYCLE	The number of times a prescription drug was purchased during a round
MEDRATE	The days in a round divided by MEDCYCLE
MEPS	Medical Expenditure Panel Survey
MPC	Medical Provider Component (of MEPS)
MSA	Metropolitan statistical area
NCHS	National Center for Health Statistics
NDC	National Drug Code
NSF	Non-self-filer
OUT1n	Outlier group based on PRATIOn (= 1 if PRATIOn < .8; = 0 otherwise).
OUT2n	Outlier group based on PRATIOn (= 1 if .8 ≤ PRATIOn < 20 for n = 1 and AWUP before 1996 *or* if .8 ≤ PRATIOn < 10 otherwise; = 0 else).
OUT3n	Outlier group based on PRATIOn (= 1 if PRATIOn ≥ 20 for n = 1 and AWUP before 1996 *or* if PRATIOn ≥ 10 otherwise; = 0 else).
PC	Pharmacy Component (of MEPS)
PERSID	Person identification number
PRATIOn	The ratio of RUPn to AWUP, using PRICEn, n = 1,2,3 and dosage amounts to construct RUPn.
PRGRP1	One of three groups used to sort drug events with a positive-valued price for outlier analysis (= 1 if PRICEn = self-payment and PRICEn ≤ $30; = 0 otherwise).
PRGRP2	One of three groups used to sort drug events with a positive-valued price for outlier analysis (= 1 if PRICEn = self-payment and PRICEn > $30; = 0 otherwise).
PRGRP3	One of three groups used to sort drug events with a positive-valued price for outlier analysis (= 1 if PRICEn does not equal self-payment; = 0 otherwise).
PRICE1	Measure of the retail price of the drug product (= the sum of payments when no payments were coded as missing)
PRICE2	Measure of the retail price of the drug product (= the sum of payments when at least one payment is reported greater than zero and at least one payment is reported as missing)
PRICE3	Measure of the retail price of the drug product (= the reported total charge for the drug product when reported to be greater than zero and all payments are reported to be either missing or zero)
PRICE4	Measure of the retail price of the drug product (lacked any positive-valued payments or total charge data)
RU	Reporting unit
RUP	Retail unit price
SF	Self-filer
SSS	Social and Scientific Systems, Inc.
VA	Veterans Affairs

Appendix B. Pharmacy Component Survey Booklet

Medical Expenditure Panel Survey

PHARMACY COMPONENT

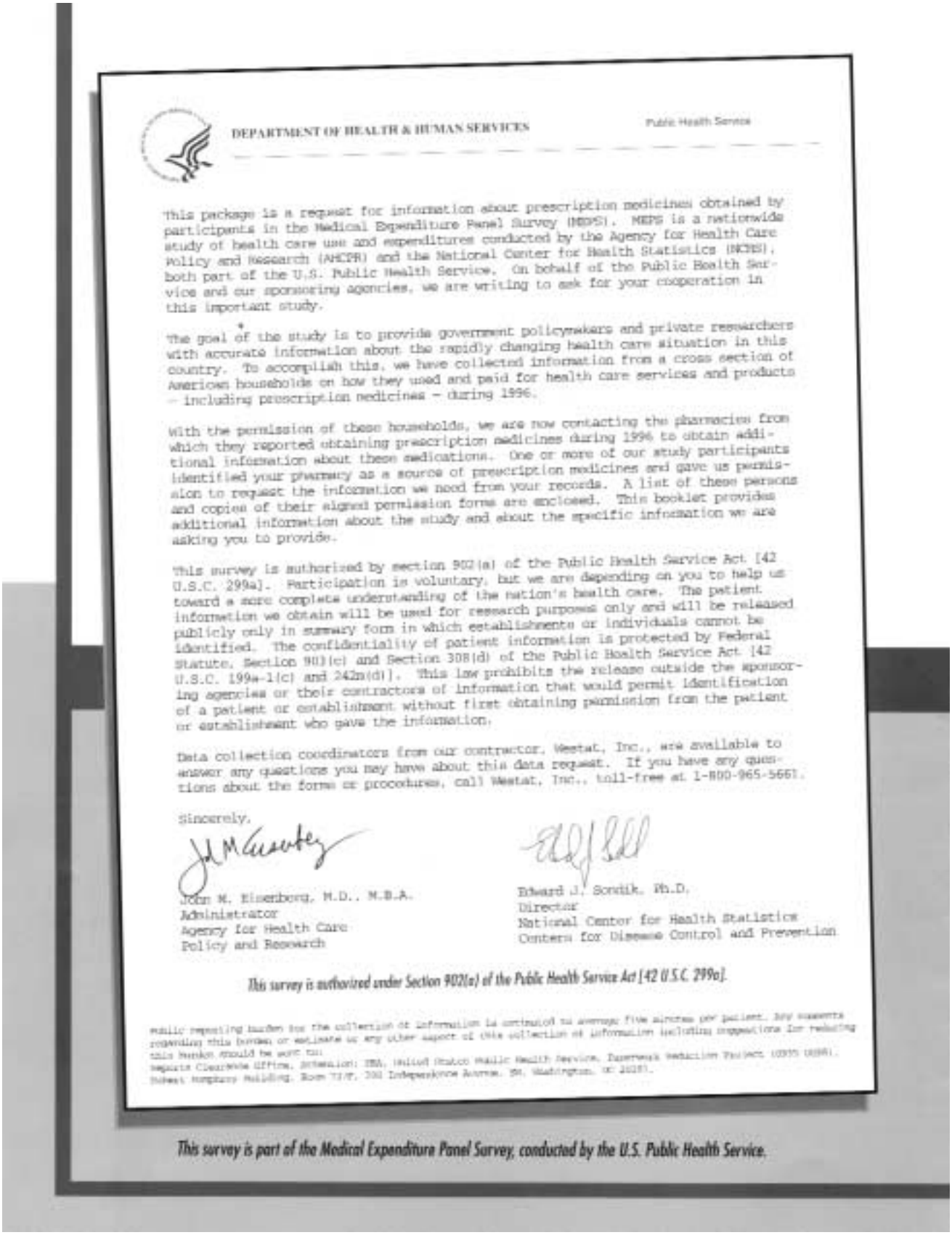
PHARID: _____

Pharmacy Name: _____

Street Address 1: _____

City: _____ State: _____ Zip: _____

Agency for Health Care Policy and Research
National Center for Health Statistics



DEPARTMENT OF HEALTH & HUMAN SERVICES — Public Health Service

This package is a request for information about prescription medicines obtained by participants in the Medical Expenditure Panel Survey (MEPS). MEPS is a nationwide study of health care use and expenditures conducted by the Agency for Health Care Policy and Research (AHCPR) and the National Center for Health Statistics (NCHS), both part of the U.S. Public Health Service. On behalf of the Public Health Service and our sponsoring agencies, we are writing to ask for your cooperation in this important study.

The goal of the study is to provide government policymakers and private researchers with accurate information about the rapidly changing health care situation in this country. To accomplish this, we have collected information from a cross section of American households on how they used and paid for health care services and products – including prescription medicines – during 1996.

With the permission of these households, we are now contacting the pharmacies from which they reported obtaining prescription medicines during 1996 to obtain additional information about these medications. One or more of our study participants identified your pharmacy as a source of prescription medicines and gave us permission to request the information we need from your records. A list of these persons and copies of their signed permission forms are enclosed. This booklet provides additional information about the study and about the specific information we are asking you to provide.

This survey is authorized by section 902(a) of the Public Health Service Act [42 U.S.C. 299a]. Participation is voluntary, but we are depending on you to help us toward a more complete understanding of the nation's health care. The patient information we obtain will be used for research purposes only and will be released publicly only in summary form in which establishments or individuals cannot be identified. The confidentiality of patient information is protected by Federal Statute, Section 903(c) and Section 308(d) of the Public Health Service Act [42 U.S.C. 299a-1(c) and 242m(d)]. This law prohibits the release outside the sponsoring agencies or their contractors of information that would permit identification of a patient or establishment without first obtaining permission from the patient or establishment who gave the information.

Data collection coordinators from our contractor, Westat, Inc., are available to answer any questions you may have about this data request. If you have any questions about the forms or procedures, call Westat, Inc., toll-free at 1-800-965-5661.

Sincerely,

John M. Eisenberg, M.D., M.B.A.
Administrator
Agency for Health Care
Policy and Research

Edward J. Sondik, Ph.D.
Director
National Center for Health Statistics
Centers for Disease Control and Prevention

This survey is authorized under Section 902(a) of the Public Health Service Act [42 U.S.C. 299a].

Public reporting burden for the collection of information is estimated to average five minutes per patient. Any comments regarding this burden or estimate or any other aspect of this collection of information including suggestions for reducing this burden should be sent to: Reports Clearance Office, Attention: PRA, United States Public Health Service, Paperwork Reduction Project (0935-0098), Hubert Humphrey Building, Room 721B, 200 Independence Avenue, SW, Washington, DC 20201.

This survey is part of the Medical Expenditure Panel Survey, conducted by the U.S. Public Health Service.

MEDICAL EXPENDITURE PANEL SURVEY
PHARMACY COMPONENT

In order to get a complete picture of health care expenses for the Medical Expenditure Panel Survey (MEPS), we need information from pharmacies and other suppliers of prescription medicines about the prescribed medicines the study respondents have received.

The individuals listed on the Patient List (included in the back pocket of this booklet) have given permission to contact you for information about prescription medicines they obtained at this location during 1996. Each patient's name, date of birth, and gender are provided to help you locate them in your records. Copies of the signed permission forms are also included in the back pocket of this booklet.

For each patient, we are requesting information about each prescription filled or refilled by your pharmacy, at this location only, between January 1, 1996 and December 31, 1996. For each prescribed medicine, we need:

- the date on which the prescription was filled;
- the National Drug Code (NDC) of the medicine;
- the generic or trade name of the drug;
- the strength of the medicine and quantity dispensed;
- the total charge for the prescription; and
- the identity of each source that paid and the amount paid by each source.

HOW TO RESPOND

Use the enclosed postage paid return envelope to mail in the requested information to Westat.

Or fax the requested information to the attention of Marolyn Alpert at 1-800-867-7801

If the requested information is available to you in computerized records:

Please send us a printout of the patient's 1996 records. The printout should show the name of the person for whom the prescription was filled, and for each prescription filled, the requested data items.

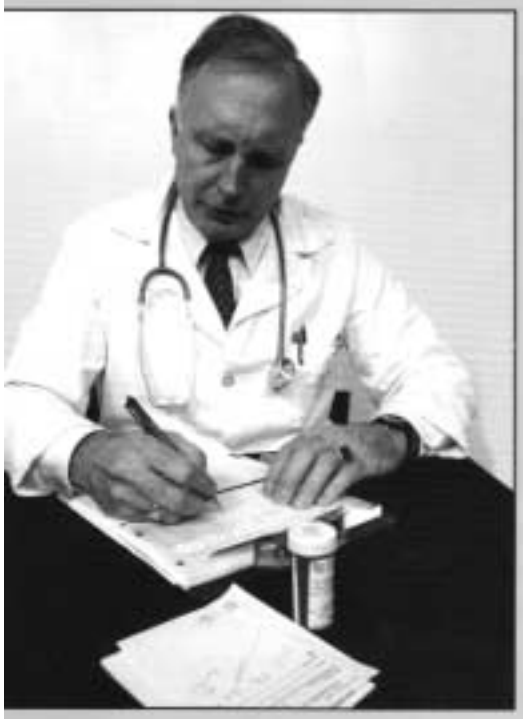
If the requested information is not available to you in computerized records:

Please copy the information from your records onto the enclosed "Prescription Information List." One is attached to each patient's permission form. Instructions for completing the Prescription Information List are included at the end of this booklet.

If you prefer to provide the information by telephone:

Please call 1-800-965-5661, and a data collection coordinator will be happy to assist you.

COMMON QUESTIONS AND ANSWERS ABOUT THE PHARMACY COMPONENT



Q. What is the Medical Expenditure Panel Survey?

A. The Medical Expenditure Panel Survey (MEPS) is a nationwide study conducted to learn more about the health care services people use, the charges for those services, and the sources that pay for them. The study is conducted by the U.S. Public Health Service through the Agency for Health Care Policy and Research and the National Center for Health Statistics. Data for the study are collected from a variety of sources, including:

- a nationally representative sample of households;
- hospitals, physicians, and other medical providers;
- pharmacies;
- employers and other sources of health insurance; and
- nursing homes.

Because of its scope, MEPS is the most complete source of data available on health care use and expenses in the United States.

Q. How are pharmacies chosen for this study?

A. Pharmacies and other providers of health care services and products were named by respondents in the household survey as places that filled or refilled prescriptions for them during 1996. These household respondents signed forms authorizing and requesting their pharmacies to release the information sought by the study. Records are being requested for a particular person at your location.

Q. Why is the participation of pharmacies so important?

A. We understand that pharmacies are very busy places, balancing the needs of serving patients and running a business. However, your participation is crucial to the success of this study and to the usability of the data collected. The information that you provide will aid in the evaluation of national policies affecting health care use and expenses in the United States.

Billions of dollars are spent each year for prescription medicines, much of it paid fully or in part by third-party sources. Patients who are not responsible for paying the entire cost of their prescriptions often have difficulty providing complete information about the total cost and amounts paid for their prescriptions. We contact pharmacies to collect details the patients cannot provide. The information you supply will supplement that given by the patient and help us build a more complete picture of health care expenditures for the patients in our study.

Q. How do I know my answers will be kept confidential?

A. All data provided by the participating pharmacies and their patients will be kept in strict confidence. Your rights to confidentiality are protected by Federal law under Sections 903(c) and 308(d) of the Public Health Service Act and the Privacy Act of 1974.

All personal identifying information such as names or addresses will be removed before information from the study is made available to researchers. The information you give will be published only in statistical summaries and tabular format.

Q. Why is a MEPS study needed now?

A. The U.S. Public Health Service is committed to improving the nation's health care system. Since this study was last conducted in 1987, many important changes have taken place in:

- the way people choose their providers of medical care,
- the ways in which health care is paid for, and
- the kinds of health insurance plans available and the services covered by those plans.

These and other changes have created a critical need for more up-to-date information on the types of health care services and products people obtain and how these services and products are paid for. The MEPS study results will inform the public, the health care community, and leaders in government and the private sector.

Q. Who is conducting the data collection?

A. The U.S. Public Health Service has chosen Westat, a national research company, to collect the pharmacy data.

Q. What questions will the study answer?

A. The study will provide answers to many important questions. For example:

- How much of health care costs, including prescriptions and medical services, are covered by insurance?
- How much do people pay out of pocket for health care services?
- What kinds of medications are not covered by most insurance plans?
- How many people have no health insurance at all?
- In what ways does the health care received by people in cities differ from the care received by people in rural areas?
- What health care costs do families and individuals face?
- What does a serious illness cost?

Q. Any further questions?

A. If you have any other questions about this study, please call toll-free at 1-800-965-5661, and a MEPS survey representative will be happy to assist you.

Many books, reports, and articles based on the information from this survey series have been published and are available at no cost. If you are interested in seeing a list of the most popular publications, please call the 800 number listed above.

PHARMACY COMPONENT
PRESCRIPTION INFORMATION LIST INSTRUCTIONS

Please follow these instructions as you complete the Prescription Information List.

The boxes on the left refer to the data items of the sample Prescription Information List on the facing page.

DATE FILLED Record the date the prescription was filled or refilled.

NDC Record the National Drug Code (NDC) for the medicine received. This is usually an 11-digit number (5-4-2) that is assigned to the specific medicine. For example: 00173-0428-00 for Zantac 150 GELdose Capsules in a bottle of 60.

GENERIC/ TRADE NAME(S) Record the generic (also, nonproprietary or product) name and/or the trade (also, manufacturer, proprietary, or brand) name of the medicine.
(*Generic Name:* The "common" or chemical name a pharmaceutical product is manufactured and sold under. Many but not all medications have a generic equivalent.)
(*Brand Name:* The name applied by a manufacturer to a particular medication. Sometimes brand names are more familiar than the generic name.)
(*Manufacturer Name:* Refers to the company that produces or sells the medicine.)

STRENGTH For the strength of the medicine, record both the amount and the unit (e.g., mg, gm, gr, mEq, mcg, %, mg, ml).

QUANTITY (PACKAGE SIZE/ AMOUNT DISPENSED) Record the amount of medicine dispensed to the patient (e.g., 20 pills, 4 fl. oz., etc.).

TOTAL CHARGE Record the dollar amount of the total charge. This is the price of the prescription: the cash price if it is a cash transaction, or the contract price if it is a third-party transaction. The prescription price is calculated based on the specific prescription and/or the specific payer's contracted reimbursement rate.

PAYMENTS

- *Patient Payment:* Record the total dollar amount of the payment made by the patient. Include any payments made to meet the patient's copayment, coinsurance, or deductible.
- *Private Insurance Payment:* Record the total dollar amount of all payments, if any, made by **private health insurance** sources. Private insurance sources do include insurance that is paid by an employer or an individual. Do not include payments made by public insurance sources, such as Medicaid, Medicare, etc. These should be included under "Medicaid," "Veterans' Administration," "CHAMPUS/CHAMPVA," "Other Federal," "Other State," "Workers' Compensation," or "Other," as appropriate.
- *Medicaid Payment:* Record the total dollar amount of the payment made by Medicaid, if any.
- *Veterans' Administration Payment:* Record the total dollar amount of the payment made by the Veterans' Administration, if any.
- *CHAMPUS/CHAMPVA Payment:* Record the total dollar amount of the payment made by CHAMPUS/CHAMPVA, if any.
- *Other Federal Payment:* Record the total dollar amount of the payment made by other Federal sources, if any.
- *Other State Payment:* Record the total dollar amount of the payment made by other State sources, if any.
- *Workers' Compensation Payment:* Record the total dollar amount of the payment made by workers' compensation, if any
- *Other Payment:* Record the total dollar amount of all payments made by other sources, if any.

If you have questions about how to complete these forms or all the requested data items are not available at your location, please call a data collection coordinator at this toll-free number: 1-800-965-5661.

SAMPLE PRESCRIPTION INFORMATION LIST
MEDICAL EXPENDITURE PANEL SURVEY - U.S. PUBLIC HEALTH SERVICE

SANDIE M KING
DOB: 07/10/65 • FEMALE
42816659-666123B5253

WALGREEN'S PHARMACY
63335109

For the patient listed to the left, please provide information about each prescription filled or refilled between January 1, 1996 and December 31, 1996.

1.
Date Filled: 03 / 26 / 96
NDC: 00003 - 0109 - 60
Generic/Trade Name(s): Amoxicillin (BMS)
Strength: 500 unit: mg
Quantity (Pkg. Size/Amt. Dispensed): 21
Total Charge[1]: $ 7.00

Payments:
Patient[2]: $ 6.00 Other Federal: $
Private Ins.[3]: $ 1.00 Other State: $
Medicaid: $ Workers' Comp: $
VA: $ Other: $
CHAMPUS/CHAMPVA: $

2.
Date Filled: 03 / 24 / 96
NDC: 00085 - 0647 - 03
Generic/Trade Name(s): Intro - A Inj
Strength: 3 unit: mmu/vial
Quantity (Pkg. Size/Amt. Dispensed): 12
Total Charge: $ 370.09

Payments:
Patient: $ 18.00 Other Federal: $
Private Ins.: $ 352.09 Other State: $
Medicaid: $ Workers' Comp: $
VA: $ Other: $
CHAMPUS/CHAMPVA: $

3.
Date Filled: ___ / ___ / ___
NDC:
Generic/Trade Name(s):
Strength: unit:
Quantity (Pkg. Size/Amt. Dispensed):
Total Charge: $

Payments:
Patient: $ Other Federal: $
Private Ins.: $ Other State: $
Medicaid: $ Workers' Comp: $
VA: $ Other: $
CHAMPUS/CHAMPVA: $

[1] *Total Charge:* Record the dollar amount of the total charge. This is the price of the prescription: the cash price if it is a cash transaction, or the contract price if it is a third party transaction. The prescription price is calculated based on the specific prescription and/or the specific payor's contracted reimbursement rate.

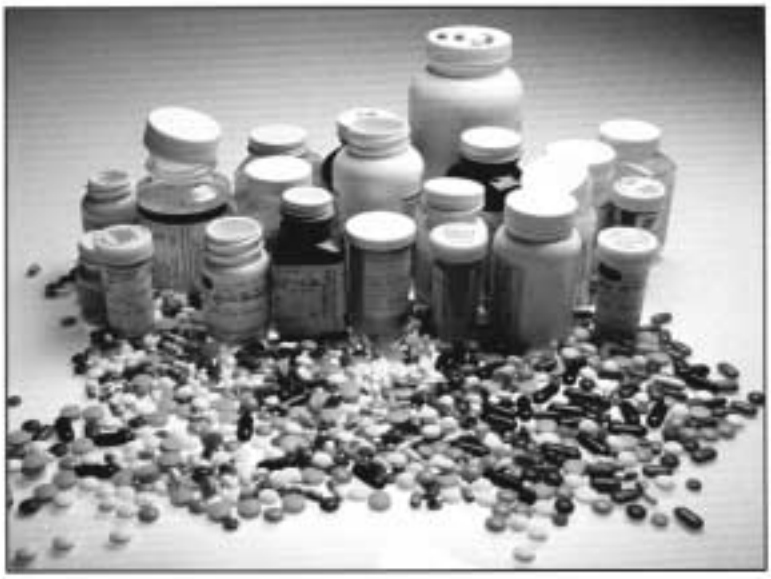
[2] *Patient Payment:* Record the total dollar amount of the payment made by the patient. Include any payments made to meet the patient's copayment, coinsurance, or deductible.

[3] *Private Insurance Payment:* Record the total dollar amount of all payments, if any, made by private health insurance sources. Private insurance sources do include insurance that is paid by an employer or an individual. Do not include payments made by public insurance sources, such as Medicaid, Medicare, VA, etc. These should be included under "Medicaid," "Veterans' Administration," "CHAMPUS/ CHAMPVA," "Other Federal," "Other State," "Workers' Compensation," or "Other," as appropriate.

You may respond to this request by any of the following methods:

- Mail the requested information in the postage paid envelope.
- Fax the requested information to the attention of Maralyn Alpert at 1-800-867-7801.
- Provide the requested information over the phone by calling 1-800-965-5661.

Thank you for your participation!



If you have any questions or concerns after you have returned the requested information, please call Maralyn Alpert at 1-800-965-5661.

Westat 1650 Research Blvd., Rockville, MD 20850

Medical Expenditure Panel Survey

Appendix C. Permission Form for Pharmacy Component

OMB #: 0935-0098

PERMISSION FORM

MEDICAL EXPENDITURE PANEL SURVEY - U.S. Public Health Service

AUTHORIZATION TO OBTAIN INFORMATION FROM PHARMACIES AND PHARMACY RECORDS

A

TO: Pharmacy: _____

Street Address 1: _____

Street Address 2: _____

City: _____ State: _____ Zip: _____

Telephone: (_____)_____-_____
Area Code

B I am voluntarily participating in this survey of health care use and expenses in the USA, a part of the Medical Expenditure Panel Survey being conducted by the United States Public Health Service. By this statement or a photocopy of it, I hereby authorize and request you to supply any needed medical or billing information about prescribed medicines filled or refilled for my use during the period January 1, 1996 to December 31, 1996. This request applies to any and all prescribed medications received by me during this period.

I understand that the Public Health Service will use the information for statistical purposes in health research, and that no information which identifies me or my pharmaceutical providers will ever be released or published. Information about me from the survey interview may be used to identify my records as necessary. This request expires 18 months from the date of signature, unless I inform you otherwise.

C

1. Patient's Name _____

2. Date of Birth ____/____/____
Month Day Year

3. Other Names Under Which Records May be Filed
_____

3A. Social Security Number** |___|___|___| - |___|___| - |___|___|___|___|

D

4. _____
Patient's Signature - 14 and over sign

5. Date Signed_____

E

6. _____
Parent, Guardian, Witness or Proxy's Signature

7. Date Signed _____

8. _____
Signer's Relationship to Patient

9. Reason for Parent, Guardian, Witness or Proxy's Signature:
- Patient 13 or Younger
- Patient 14-17 Years Old
- Patient Deceased
- Patient Disabled
- Patient in Health Care Institution

*NOTICE: Information contained on this form that would permit identification of any individual or establishment has been collected with a promise that it will be held in strict confidence by the sponsoring agencies or their contractors, will be used only for purposes stated in this study, and will not be disclosed or released to anyone other than authorized staff of the sponsoring agencies. AHCPR and NCHS, without the consent of the individual or the establishment in accordance with Section 903(c) and Section 308(d) of the Public Health Service Act (42 U.S.C. 299a-1(c) and 242m(d)). No information will be disclosed where prohibited by federal law and regulations governing the confidentiality of alcohol and drug abuse patient records, 42 USC 290dd-3 and 290cc-3, 42 CFR Part 2.

K:\RSGRP\MEPS\PFS\P1-5PHRM.DOC

**NOTICE: Your Social Security Number is requested to allow the addressee to accurately identify and locate your records. This information is voluntary and is collected under the authority of Title IX, Section 902(a) of the Public Health Services Act (42 U.S.C. 299a). There will be no effect on your benefits and no information will be given to any other government or nongovernment agency.

_____ SCAN: YES * NO *
CODE

U.S. Department of Health and Human Services
Public Health Service
Agency for Healthcare Research and Quality
2101 East Jefferson Street, Suite 501
Rockville, MD 20852

Official Business
Penalty for Private Use $300

AHRQ Pub. No. 01-0002
June 2001

ISBN 1-58763-024-9
ISSN 1531-5673