MEPS

Methodology Report #29

Outpatient Prescription Drugs: Data Collection and Editing in the 2011 Medical Expenditure Panel Survey

Abstract

The Medical Expenditure Panel Survey (MEPS), sponsored by the Agency for Healthcare Research and Quality (AHRQ), is a nationally representative survey of the U.S. civilian noninstitutional population's medical care use and expenditures. Household respondents report prescription drugs obtained by members of the household and the number of times each drug was obtained, while a follow-back survey of pharmacies is the primary source of prices, payers, and drug attributes. This report describes the household and pharmacy data collection and editing processes, the editing and imputation techniques used to supply values for missing data in the pharmacy database, the procedure for linking the data reported by the pharmacy to each prescription drug reported by the household and recent improvements in these procedures. Statistics on these methods are presented for the 2011 MEPS.

Suggested Citation

Table of Contents

Introduction

The Medical Expenditure Panel Survey (MEPS) is a nationally representative survey of medical care use and expenditures sponsored by the Agency for Healthcare Research and Quality (AHRQ) covering the U.S. civilian noninstitutional population. MEPS comprises a Household Component (HC) and a Medical Provider Component (MPC). MEPS data can produce estimates for individuals, families, and selected population subgroups. The Household Component (HC) provides data on health status, demographic and socio-economic characteristics, employment, access to care, and experiences with health care. The Medical Provider Component (MPC) collects expenditures data from hospitals, physicians, home health care providers, and pharmacies identified by MEPS-HC respondents. Its purpose is to supplement the HC information.

Each year, a new panel of households is selected and participates in five rounds of HC interviews, covering two full calendar years. This set of households is a subsample of households that participated in the previous year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics (NCHS) of the Centers for Disease Control and Prevention (CDC). The NHIS sampling frame provides a nationally representative sample of the U.S. civilian noninstitutionalized population and reflects an oversample of blacks, Hispanics and, beginning in 2006, Asians. MEPS then oversamples additional policy-relevant subgroups such as low income households. Two overlapping panels create an annual sample size of about 15,000 households.

MEPS supports longitudinal analysis; each panel participates in five rounds of interviews to collect two full calendar years of data. It is possible to analyze even more long-term trends by linking MEPS to the previous year's NHIS. Cross-sectional analysis over a long period is possible as well; MEPS began in 1996, and the data collected are comparable to those from earlier medical expenditure surveys conducted in 1977 and 1987.

MEPS defines expenditures as payments from all sources (including individuals, private insurance, Medicare, Medicaid, and other sources) for health care services during the year. To construct estimates of expenditures, MEPS collects detailed information about medical events including ambulatory visits (outpatient and office based), inpatient stays, emergency care, home health care, dental care, other medical care, and prescription medicines. At each HC interview, the household respondent supplies the names of any prescribed medicine that family members obtained and identifies the pharmacies used. In addition, family members are asked for permission for MEPS to contact the pharmacies. With this permission, the Pharmacy Component (PC) collects detailed information from the pharmacies about the drugs obtained, including payments (the sum of which is the price), payers, date each prescription was filled, quantity dispensed, the National Drug Code (NDC), and precise drug attributes. Matching drugs reported by the pharmacies in the PC to those reported by the household in the HC is accomplished by supplemental data and specialized software.

The MEPS conducts the PC to collect information that pharmacies can more easily and accurately provide than household respondents. Pharmacies, since they receive payments for the drugs they

dispense, are generally likely to have more accurate payment information than households, which may be more familiar with their out-of-pocket payments than third party payments. Some HC respondents may lack documentation, such as explanations of benefits, with details about third-party payments. The PC is designed to collect information about payments for drugs, but data users should note that rebates between manufacturers and pharmacies, pharmacy benefit managers (PBMs), and government programs are not collected. A second motivation for conducting the PC is that some households do not have easy access to details about their medications, such as the number or strength of pills. In addition, some households have many fills for multiple medications during the reference period and asking them for information about each of these would be repetitive and overly burdensome.

As with all surveys, inconsistent and missing data occur in MEPS. Errors and failures to report information happen. Some pharmacies do not respond to the PC and some sample members deny permission to contact pharmacies, so PC data are not available for every sampled person. Even when pharmacies respond, the data provided are not always complete.

This report describes the methods used to (1) supply values for missing HC and PC data and (2) link the data reported by the pharmacies to each prescription drug name reported by the household. Table 1 summarizes the key variables edited or imputed. Details of edits and imputation are in the body of the report. The objective of these procedures is to maximize the amount and quality of data available for analysis and to reduce the risk of bias associated with reporting error and missing data. Statistics on these methods are presented for the 2011 MEPS.

This report supplements the overview of expenditure imputation for other types of events provided in an earlier publication (Machlin and Dougherty 2007) and updates an earlier Methodology Report on the 1996 MEPS prescription drugs (Moeller et al. 2001). Since the 1996 data were released, several enhancements have been made to improve the quality and analytical capability of the MEPS prescription drug microdata. Briefly, some improvements implemented include the following:

- Variables specifying the therapeutic classes of drugs were added to the public use data files beginning with the 2002 data.
- Imputed NDCs were added to the public use files beginning with the 2005 data.
- The rules for editing pharmacy-reported prices were improved to account for changes in the relationship between actual and list prices, especially for generics, beginning with the 2007 data and refined further for the 2009 data. These changes also greatly reduced the extent to which pharmacy-reported quantities are edited.
- Improvements in matching the PC to the HC were first implemented with the 2008 data to (1) reduce the need for imputation, and (2) shift the distribution of drugs by patent status closer to the distribution found in independent data sources.
- Improvements in the PC editing first implemented with the 2008 data remove bias toward higher out-of-pocket and higher Medicare Part D expenditures.
- Improvements in price imputation within the PC were implemented beginning with the 2010 data to increase accuracy and preserve price variation within NDCs.

- Additional improvements in PC-to-HC matching were implemented for the 2010 data to more accurately represent the variety of diabetic supplies and equipment that individuals obtain.
- Improvements to account for price discounts in the Medicare Part D "donut hole"[1] were implemented beginning with the 2011 data. Under the Affordable Care Act, after reaching a spending threshold on drugs, the price of the brand-name drug is discounted 50 percent.

The importance of these changes depends on the research focus. Most changes appear not to be particularly important for general users, especially if they are not examining trends over time. But some improvements may affect analyses focusing on, for example, the prices of brand-name drugs compared to generics (drugs that are available from multiple manufacturers and pharmaceutically equivalent to a brand name drug), out-of-pocket expenditures for Medicare or Medicaid beneficiaries, and studies that rely on quantities, such as calculations of drug possession ratios. Footnotes in this report identify when changes were implemented to assist users in determining which years might be appropriate for their studies. More specific details about reasons for changes in data processing are available in Hill, Zuvekas, and Zodet 2011 and Zodet, Hill, and Miller 2010.

This report is organized as follows: First, it describes data collection for the Household Component and the procedures used to edit the Household Component data. Second, the discussion turns to data collection for the Pharmacy Component, and the procedures employed for editing and imputing data in the Pharmacy Component. Finally, an overview and details of matching pharmacy data to household data and the methods of editing the matched data are presented.

Household Component

Data Collection

The MEPS Household Component (MEPS-HC) has a panel design with two overlapping cohorts of the U.S. noninstitutionalized civilian population combined to produce annual estimates (J. Cohen 1997; S. Cohen 1997, 2003). A new cohort of households begins each year and is interviewed five times to collect two calendar years of data. A single informant in each household typically provides data about each household member during each interview. The MEPS asks that this person be the family member most knowledgeable about health and health care use. In an interview, the average recall period (the "round") is five months. The respondent-informant is given a calendar and other materials to aid memory, and instructed to retain paperwork such as receipts and insurance benefit statements.

During each interview, the HC gathers information on health care services used, including prescribed medicines. The first opportunity to mention prescription drugs occurs when the HC asks about non-

---

[1] The "donut hole" is a coverage gap encountered by some beneficiaries with Medicare Part D coverage. The "donut hole" is reached when a person's prescription drug expenses exceed the initial coverage limit of their plan but have not yet reached the plan's catastrophic coverage level. While in the donut hole, individuals have paid the entire price of the drugs out of pocket, but the Affordable Care Act is slowly reducing cost sharing.

prescription drug health service events that took place during the round. The respondent supplies the names of medications prescribed as part of a visit to an inpatient stay, emergency room, hospital outpatient clinic, or dentist's or doctor's office and subsequently filled. Another opportunity to mention prescriptions occurs in the dedicated prescribed medicines section of the survey where the respondent can identify prescription medicines not already mentioned. In either of these sections of the HC interview, the respondent may consult records such as medicine bottles or receipts, so the reported medication name is often quite specific. However, the information can also be minimal, for example, "pain pills." When drugs are reported, the drug names are entered in a dynamic roster.

Besides the drug name, the HC collects the following information about each medicine:

- the number of times it was obtained
- the health condition it was prescribed for
- the year and month in which the person first used it (asked only in the first interview in which it is mentioned)
- whether free samples of the drug were received.

Payment information gathered in the HC consists of the following:

- the usual third-party payer for the person's prescriptions
- the amount the person paid out of pocket for the last acquisition of any drug in the round
- whether the person is a self-filer (filing insurance claims directly with the insurance company) or not a self-filer (the pharmacies file the claims automatically at the point of purchase, or the person is uninsured)
- for self-filers, charges, and payments by source for the last purchase of each medicine in the round (because pharmacies are unlikely to have information about private third-party payments in this situation).

The names, addresses, and types of pharmacies that filled each household member's prescriptions are also requested, along with permission for MEPS to acquire data from these pharmacies. Signed authorization forms allow pharmacies to respond to the Pharmacy Component (PC) of MEPS when it is conducted. For the interviews collecting information about medication obtained in 2011, 69.7 percent of pharmacy permission forms were signed.

If insulin and diabetic supplies, such as syringes and glucometers, are reported in the prescription drug section of the HC, the survey collects the same information as it does for drugs. But when insulin or diabetic supplies are mentioned in the section on other medical expenses of the HC, the HC instrument (1) creates entries in the drug roster with the names "Insulin" or "Other Diabetic Supplies and Equipment" and (2) treats these items like those of self-filers, asking for the number of acquisitions, and charges and payments by source for the last acquisition of each. Thus names of the diabetic supplies are

not collected. (Note that any diabetic supplies purchased without a prescription represent a nonprescription addition to the prescription drug expenditure and utilization data.)

In 2011, household respondents reported 132,874 unique drug names obtained by family members in a round (hereafter referred to as "person-round-drugs"), including insulin and diabetic supplies.[2] Of these, 1,618 were obtained by self-filers. Person-round-drugs reported in the other medical expenses section of the survey included 673 for insulin and 1,507 for diabetic supplies. After accounting for insulin and other diabetic supplies obtained by self-filers, there were 3,525 person-round-drugs with payment data.

The end result of HC data collection is a roster of drug names that each sample person obtained in the round, the person's health condition or conditions associated with each drug, the number of acquisitions, whether any free samples were received, a roster of the person's pharmacies, and limited information on expenditures.

Because the HC collects limited information on expenditures and no information on drug characteristics, the rest of the expenditure information and other details about each prescription medicine—quantities, form, strength, dates obtained, price, payments by source, etc.—come from the PC. A supplementary data source (Cerner Multum Lexicon) provides information on therapeutic class and pregnancy category.

Editing Household Component Data

Editing the HC drug data consists mainly of imputing a value for the number of times a drug was obtained in the round when this information is deficient as collected in the HC. Reported payments are also edited for insulin, diabetic supplies and equipment, and drugs purchased by self-filers.

GPI coding. A first step in preparing the HC data for editing is to code each drug name to a Generic Product Identifier (GPI). A proprietary data set, Wolters Kluwer's Master Drug Data Base (MDDB®) Version 2.5 provides the GPI, a 14-digit code which identifies groups of pharmaceutically equivalent drugs that have the same active ingredient(s), dosage form and strength. The GPI is the common variable linking the PC data to the HC. It allows information to be drawn from other sample members (imputed) in cases of missing information. The first pair of digits of the GPI represents the drug group, the second pair represents the therapeutic class, and the third pair represents the therapeutic subclass. The next four digits of the GPI (digits 7–10) represent the active ingredient(s). The first 8 digits of the GPI are sufficient to identify most active ingredients, but 10 digits are needed for compounds and salts. The remaining digits of the GPI represent dosage form and strength. For the HC, professional coders use the MDDB to identify as many digits of the GPI as possible based on the medication name and any other information appended to the name provided by the respondent. Typically, 8 to 10 digits can be coded for household-reported drug names, but in 2011, 1,125 (0.8 percent) of person-round-drugs could not be coded.

---

[2] Each person-round-drug could be acquired one or multiple times during the round.

Imputing the number of acquisitions. The number of times a drug identified by name in the HC was obtained by a person in a round needs to be imputed when it is missing or extreme. In the 2011 data, for 4.2 percent of household-reported drug names, the respondent did not know or remember the number of times the drug was obtained during the round. Outlier values for the number of times a household reported obtaining a drug in a round occur as well, and are determined by comparing the number of days in the round and the number of acquisitions of the drug reported in the round. Reports of very frequent purchases of a drug—6 days or less per acquisition—are reviewed by a pharmacist to assess their plausibility.[3] For missing and implausible values, a hot-deck procedure imputes a new number of acquisitions, drawing from the donor pool of drugs with valid values. Specifically, days per acquisition are calculated in the donor pool, imputed to recipients with replacement, and then the recipient's days in the round are divided by the days per acquisition to calculate the number of acquisitions in the round. Four attempts are made to match a person-round-drug to a recipient person-round-drug. The first attempt requires an exact match on drug, dosage form, and strength (14-digit GPI). The second attempt requires an exact match on drug and dosage form (12-digit GPI). The third attempt requires an exact match on drug (8-digit GPI). The fourth attempt requires an exact match on drug group (2-digit GPI), and therapeutic class and subclass are weighting variables in the match. Weighted match variables used in all four attempts are: whether the person used any mail-order pharmacies, insurance status, and potential payment sources for the person, private and public HMO enrollment, and the person's age, state of residence, census division, region, and urbanicity. The weight for each variable is based on its relative contribution to the variation in days per acquisition.[4]

Table 2 summarizes the results of the process of imputing acquisitions in the 2011 HC. Nearly all (99.1 percent) of the imputations were based on exact matches to pharmaceutically equivalent drugs (with the same active ingredients, dosage form, and strength). After this imputation, the 2011 data file contained 348,520 acquisitions.

Editing HC payment data. For self-filers' acquisitions, and for acquisitions of diabetic supplies, a preliminary set of edits applies to the HC payment data. These procedures are based on similar edits applied to nonprescription events. The edits correct inconsistencies between payments and insurance, using round-specific variables indicating the usual third-party payer and whether the person had insurance coverage for prescription drugs (private health insurance, Medicare Part D, Medicaid, other public program, or private HMO). Although the HC asks about payments for the last acquisition of the drug in the round, the payment information is assumed to be relevant for all acquisitions of the drug in the round. The editing is at the acquisition level, but it is the same for all acquisitions within a person-round-drug. The purpose of these edits is summarized below, with the number in parentheses indicating how many of the 348,520 total acquisitions the edit affected in 2011.

---

[3] From 1996 to 2008, the threshold was 3 or fewer days per fill. Based on validation data and the distribution of the days per fill in the Medicare Part D administrative data, the threshold was raised to 6 or fewer days per fill starting with the 2009 data.

[4] Prior to the 2011 data, the imputation method gave greater weight to donors with more acquisitions in the round and did not condition on insurance status, geography, and personal characteristics. See Zodet, Hill, and Zuvekas 2010 for validation results for prior years.

- Eliminate the inconsistency created from respondents' mistakenly reporting private insurance when they actually had Medicare Part D coverage (178 acquisitions).
- Eliminate Medicaid as a source of payment when sources other than Medicare Part D and out of pocket are present and when there is an out-of-pocket payment greater than $10 (188 acquisitions).
- Process acquisitions covered by Medicaid so they are correctly classified for later imputation. Specifically, when no other payers are reported, the out-of-pocket amount is zero, and Medicaid is a possible payer, then Medicaid is assumed to be the third-party payer. Household respondents are assumed to not know the Medicaid payment (373 acquisitions).
- If total charge is reported to be between $5 and $10,000, no payments are reported, and out of pocket is the only missing payment source, set the out-of-pocket amount equal to the total charge (46 acquisitions).

After these edits and imputation, the HC data are comprised of each drug named during every HC interview. These data include a GPI code, the number of acquisitions, and payment data corrected for consistency with the person's insurance information. The data are now ready to be linked to the payment and other drug details from the PC.

Pharmacy Component

Data Collection

The Pharmacy Component (PC) collects data from the pharmacies identified by HC respondents as places where household members obtained prescription medicines during the calendar year. The PC is a voluntary survey and is conducted each year by telephone, fax, and mail. During the second, third, fourth, and fifth HC interviews, sample members are asked to sign permission forms authorizing MEPS to contact their pharmacies and authorizing the pharmacies to release records to MEPS. Only those pharmacies for which an HC sample member signed this permission form and which voluntarily participated are included in the PC.

The pharmacy is asked to provide patient profiles or information about each prescription filled or refilled for each patient during the calendar year. For each acquisition of a drug, the following information is requested:

- the dates the prescription was filled or refilled
- the National Drug Code (NDC)
- the quantity dispensed
- the number of days supplied
- the amount of out-of-pocket payment
- whether there were any third-party payments, and, if so, the third parties and payment amounts.

If the pharmacy does not provide the NDC, the PC asks instead for the medication name, dosage form, strength, and strength unit.

The data collection protocol for pharmacies is as follows. For small retail pharmacies, that is, pharmacies not associated with a large chain, the data collection staff contacts the pharmacy to explain the study's purpose, verifies receipt of authorization forms, requests specific data elements, and determines if patient profiles are available. The data collection staff may collect information by telephone, or request that the data be sent by fax or mail so that the necessary data elements can be abstracted. Pharmacy data are received in any format including hardcopy patient profiles, electronic files with patient profile data, or reports via telephone.

For large retail pharmacy chains, individual pharmacies are grouped by chain. Negotiators follow up with these pharmacies in one of two basic ways:

- If the corporate office of the retail chain prefers that the local stores respond, data collection follows the small retail model.

- If the pharmacy prefers that the data request be handled with a regional or central contact, the negotiator facilitates the most efficient method for data collection.

In addition, pharmacy patient profiles are requested directly from HC family members who reported using one or more of several pharmacy chains that had repeatedly refused to participate in the PC. Families are mailed a request for profiles. To minimize burden on households, the mailing is limited to households that had completed their participation in all 5 rounds of the MEPS, and no follow-up attempts are made for those who do not reply to a mail request. For the 2011 data collection, patient profile requests were mailed to 1,586 families with 2,315 patient-pharmacy pairs. Completed profiles were received for 15.9 percent of the pairs. Despite the overall low rate of return, the effort does provide a number of profiles for patients associated with pharmacy chains that would not otherwise be represented in the pharmacy data.[5]

Table 3 presents the sample size and participation rates in the 2011 PC. Conditional on having obtained 69.7 percent of signed permission forms sought in the household interview, about three-quarters of eligible pharmacies responded. The responding pharmacies provided data on 252,176 unique acquisitions of drugs during 2011.

To the extent allowed by survey response, the end result of PC data collection is a roster of each HC household member's drug acquisitions during a calendar year. Information about each of these acquisitions includes the date of the transaction, the NDC (or analogous identifying drug information), the quantity, the number of days of the drug that the acquisition supplied,[6] and payments to the pharmacy by source.

Supplemental Data

As mentioned above, matching the pharmacy data to the household data requires assigning a common set of codes to the household- and pharmacy-reported drugs. Wolters Kluwer's proprietary Master Drug Data Base (MDDB®), Version 2.5 is the source of such a code, the Generic Product Identifier (GPI), also described above. For the PC, automated and manual coding generally succeed in recording all 14 digits of the GPI based on the NDC (which is more specific than the GPI) or the medication name, dosage form, and strength. A negligible number of acquisitions in the PC cannot be coded to a GPI based on the information the pharmacy provides.

The MDDB has several other uses besides GPI coding. Information about the patent status and the average wholesale unit price (AWUP), used in price editing and imputation, come from the MDDB. The MDDB is also used to impute the NDC where needed. (Price editing, price imputation, and NDC

---

[5] Families were first asked to provide these profiles for the 2004 MEPS data.
[6] 2010 is the first year days supplied was provided on the public use file.

imputation are discussed in more detail below.) If the pharmacy reports an NDC found in the MDDB, the MDDB provides information on the drug name, dosage form, and strength.[7]

Two other data sets are used in editing the pharmacy data. A price list obtained from the U.S. Department of Veterans Affairs (VA) is used to determine prices of drugs dispensed by federal pharmacies when total payments are missing. The Multum Lexicon database from Cerner Multum, Inc. provides variables for therapeutic classes and FDA pregnancy categories. These variables are merged onto the PC data file by NDC.

Editing the Pharmacy-Reported Data

To maximize the donor pool for imputation, the entire PC data file (252,176 acquisitions in 2011) is edited before matching it to the HC. Key aspects of this process are: imputing missing NDCs and missing quantity dispensed, editing ambiguous third-party payer information, identifying acquisitions with deficient payment information, and imputing missing payment information.

Most editing and imputation is conducted at the acquisition level, but there are exceptions. For imputing NDCs and matching to the HC, acquisitions are aggregated to the person-round-drug level. That is, for each person and round, all pharmaceutically equivalent acquisitions (having the same 14-digit GPI, which represents the active ingredient, dosage form, and strength) are grouped together. The GPI can group multiple NDCs such as brand name drugs and their generic equivalents.

Imputing NDC

Among the 252,176 acquisitions reported by pharmacies in the 2011 data, 12,770 had a missing or invalid NDC. A missing NDC is imputed rather than left missing because it distinguishes between brand name drugs and generics, which is important for identifying price outliers and imputing missing payment data. Furthermore, many analyses of medication use depend on the NDC to identify medications of interest. The NDC can identify the manufacturer, the patent status, and the average wholesale price of a drug. The NDC, therefore, increases the accuracy of any price imputation and the analytic utility of the data.

MEPS uses three approaches to identify the NDCs of acquisitions missing this item in the PC. First, wherever feasible, the NDC is taken from one of the person's other acquisitions in the same round. For each person and round, all acquisitions of drugs with the same 14-digit GPI (representing the active ingredient, dosage form, and strength) are grouped together. Within these groups, if at least one

---

[7] Prior to 2009, regardless of whether an NDC was reported, the pharmacy was the primary source for the drug name, dosage form, and strength, and missing values were obtained from the Red Book[TM] Database. In 2009, if any 11-digit number was reported for the NDC, then the drug name, dosage form, and strength were primarily imputed from the Multum Lexicon database from Cerner Multum, Inc. and secondarily from the MDDB.

acquisition is missing NDC and at least one acquisition has an NDC, then the missing NDC is assigned the reported NDC. If more than one NDC in GPI-14 group exists, one is selected at random.

Second, for the residual acquisitions still missing NDC, the MDDB supplies it. The basic approach is to use matching software to find the best match to a drug product on the MDDB based on the medication name, quantity units (for example, milliliters), and the GPI. Medication name is used as a match variable because various brand name and generic products, each with a unique NDC, can be grouped together under a single GPI. When a pharmacy reports NDCs, it almost always reports the same NDC for all the person's acquisitions of a drug, so in order to preserve homogeneity among the recipient acquisitions, one NDC is imputed to all the acquisitions for the same person-round-drug.

Three attempts are made to impute an NDC by matching to the MDDB. In the first attempt, one NDC is taken from the MDDB for all acquisitions with the same person, round, active ingredients, dosage form, and strength. In other words, an exact match on 14-digit GPI is required. For the second attempt, one NDC is imputed from the MDDB to all acquisitions with the same person, round, and active ingredients (10-digit GPI). Here, the dosage form and strength may differ. An exact match on active ingredients is required, and again the medication name is a match variable. In the third attempt, medication name but not the GPI is used.[8]

A third approach applies to the residual acquisitions still missing NDC, those for which the coders are unable to determine at least a 10-digit GPI from the information the pharmacy reports. Most of these acquisitions lack the drug name as well as any information about the quantity dispensed and strength. In these cases, an NDC is imputed from the acquisitions of a person with similar characteristics reported in the HC and for whom the pharmacy reported both drug names and valid NDCs. Match variables include the person's age, sex, geographic division, urbanicity, health conditions, potential payment sources, and self-filer status. In addition to the NDC, the recipient acquisition adopts any missing medication name, GPI, quantity dispensed, or strength.

Table 4 summarizes the results of the NDC imputation process for the 2011 data. Some imputations were from other acquisitions of the same person in the same round. Most of the imputations relied on detailed information provided by the pharmacies.[9]

---

[8] For the data from 1996 through 2007, all NDC imputation was conducted separately for each acquisition. Imputation within person-round was added starting with the 2008 data. Also starting in 2008, the attempts using the MDDB were modified to impute one NDC for (1) all acquisitions with the same person, round, drug, dosage form, and strength, and (2) all acquisitions with the same person, round, and drug. The new procedure was implemented because of the homogeneity of NDCs found among the acquisitions that did not have missing NDCs in the 2007 data: the NDCs were uniform for 95 percent of these person-round-drug-dosage form-strength combinations.

[9] For the data from 1996 through 2004, AHRQ's licensing agreement with the proprietary database precluded releasing the imputed NDC values to the public, so for these prescriptions, the household-reported name of the prescription and the pharmacy-reported NDC and prescription name were provided on the public use file to allow users to do their own imputation. In addition, for the years 1996-2004, the imputed NDC values could be accessed through the AHRQ Data Center.

Imputing quantity dispensed

Pharmacies report quantity dispensed for nearly all acquisitions, but occasionally it must be imputed. If the NDC is imputed and the quantity is missing, then the quantity is taken from the same acquisition that donated the NDC. Otherwise, matching software imputes a quantity from another acquisition. Match variables include the NDC; active ingredients, dosage form, and strength (GPI); and characteristics of the person reported in the HC (age, sex, health conditions, and health status). Exact matching on the drug is required, and heavier weight is placed on the NDC, followed by the dosage form and strength. In the 2011 data, the quantity dispensed was imputed for 0.7 percent of the 252,176 acquisitions (table 1).

Editing third-party payers

Sometimes the pharmacy does not clearly identify a third-party payer. For example, a pharmacy might report as a payer a pharmacy benefit management company, which could be under contract with a private insurer, Medicare prescription drug plan, Medicaid, or other public program. For some acquisitions, pharmacies do not report the third-party payer at all. In this case, information from the HC about insurance coverage and usual third-party payer can indicate the type of payer. For example, if the respondent in the HC reports Medicare Part D as the usual third-party payer, then the unknown payer in the PC is set to Medicare.[10]

In the 2011 data, among the 252,176 acquisitions reported by pharmacies, the third-party payer was edited for 26.3 percent of acquisitions (table 1). Specifically, the category of third-party payer was assigned in three situations using the insurance status reported in the HC.

- The third-party payer was reportedly a pharmacy benefit management company or HMO (41,126 acquisitions). The payer in these cases was coded mainly as private insurance, Medicare, Medicaid, or out of pocket.
- The third-party payer is unknown (18,172 acquisitions). The payer in these cases was also coded mainly as private insurance, Medicare, Medicaid, or out of pocket.
- The payer was reportedly a public clinic or some other ambiguous public payer (1,689 acquisitions). The payer in these cases was coded mainly as Medicaid, Medicare, or Other State and Local.

Additional edits to the PC acquisitions (patterned after edits applied to the nonprescription provider data) include the following:

- Correcting the information from pharmacies that mistakenly report a payment from private insurance instead of a Medicare prescription drug plan (3,241 acquisitions) or Medicaid HMO (2,008 acquisitions).[11]
- Eliminating the out-of-pocket payment when the pharmacy reports a patient copayment that is equal to the reported amount for Workers' Compensation (75 acquisitions).

---

[10] Starting with the 2009 data, drugs excluded from Medicare Parts B or D coverage are not assigned to Medicare in these edits.
[11] From 2006 through 2008, all Medicaid payments for Medicare beneficiaries were recoded as Medicare payments.

Identifying acquisitions needing price imputation

The primary aim of MEPS is to collect information about health care expenditures, including the costs of prescription drugs. Payment information that is altogether or partly missing, incorrectly reported as zero, or otherwise inaccurate will cause distortion in estimated expenditures. Fortunately, responding pharmacies provided all the information requested for 73.3 percent of acquisitions in 2011 (table 5). However, third-party payment amounts are often missing. Federal pharmacies are especially likely to omit third-party payment amounts, because these pharmacies report little payment information generally.

Even when payment data appear to be complete, MEPS attempts to detect inaccurate payment data by comparing an acquisition's price (the sum of payments) to a price provided in the MDDB. If the sum of payments falls outside a certain range relative to the benchmark, then MEPS imputes a price and distributes it among sources of payment.

The threshold used for defining what constitutes deficient payment information depends on the pattern of reported payments. Generally, deficiency comes under one of four categories:

- The acquisition lacks any payment information at all.
- The acquisition has missing third-party payments.
- The acquisition has zero third-party payments and the out-of-pocket amount is unrealistically low as a total price.
- The acquisition's payments sum to an unrealistically high or low price.

The primary method for identifying unrealistically high and low prices ("price outliers") is to assess prices per unit (tablet, capsule, ounces, etc.). This entails comparing the acquisition's retail unit price (RUP) reported in the PC to the NDC's average wholesale unit price (AWUP). Thresholds (described below) define a RUP as implausible—either too high a price or too low. The RUP is the sum of payments for the acquisition divided by the quantity dispensed. The AWUP for the NDC is taken from the MDDB. It is calculated as the average wholesale price (AWP) divided by the number of units in the package. When multiple AWPs are available, the one dated closest to the middle of the calendar year is selected, but in some cases only an AWP for a prior or subsequent year is available.

The AWUP is an imperfect benchmark, and recent research has led to changes in how the AWUP is used to identify prices that need editing. In particular, the AWP is a list price, not an average of transaction prices. The editing rules were first modified for the 2007 data based on a study benchmarking the distribution of prices in the MEPS to private claims data (Zodet, Hill, Miller 2010). A subsequent validation study of MEPS sample members in Medicare Part D led to further modifications first implemented in the 2009 data (Hill, Zuvekas, Zodet 2011).

Lower outliers. Table 6 summarizes the rules used to identify acquisitions with prices that are too low (lower outliers). The rules vary by patent status, because the relationship between RUP and AWUP varies by patent status (Zodet, Hill, Miller 2010). The three patent statuses are (1) single source brand name drugs, which are available only from one manufacturer; (2) originator drugs, which are brand

name drugs with therapeutically equivalent competitors; and (3) generic drugs, which are available from multiple manufacturers and pharmaceutically equivalent to a brand name drug. In the validation study, acquisitions with both out-of-pocket and third-party payment data were found to be mostly accurate, so only very low prices are flagged for imputation. In contrast, the thresholds are higher for acquisitions with zero third-party payments, because in the validation study acquisitions with low RUPs and zero third-party payments were seldom valid. Similarly, acquisitions with incomplete payment data were almost always missing positive payments in the validation data; therefore more of these prices are flagged for imputation. (Note that some free acquisitions, such as free antibiotics from pharmacies in grocery store chains offering these antibiotics to customers with prescriptions, are not treated as outliers.)

The thresholds are lower for brand name drugs obtained by Medicare Part D beneficiaries who have spent enough to enter the donut hole (table 6). The thresholds are lower because, starting in 2011, these acquisitions are discounted 50 percent.[12] These acquisitions are identified using the cumulative total out-of-pocket payments reported by the pharmacies for acquisitions covered by Part D. The lower threshold is not used for beneficiaries who participate in the Low Income Subsidy program and have therefore not entered the donut hole. (Participation in the Low Income Subsidy program is determined by HC reports of Medicaid coverage or paying no premium for Part D.) Acquisitions that are likely under the catastrophic coverage benefits of Medicare are compared with the usual, rather than reduced, thresholds.

Upper outliers. Prices that are too high (upper outliers) are defined as cases where the reported price per unit equals or exceeds 10 times the AWUP. This rule applies regardless of patent status or the completeness of the payment data.[13]

2011 Results. Table 7 presents the number of acquisitions identified as needing and not needing price imputation in the 2011 data according to these criteria. Few of the acquisitions with complete payment data, 4,628 of 184,874 or 2.5 percent, were flagged as lower outliers for imputation, and these were primarily cases with zero third-party payment amounts.[14] Two-thirds of the acquisitions with partial

---

[12] These lower thresholds were first implemented with the 2011 data.

[13] For the data from 1996 to 2006, all acquisitions, regardless of patent status or the completeness of payment data, were edited if the price was less than 0.8×AWUP or greater than 10×AWUP. For the data for 2007 and 2008, all single source drugs were edited if the price was less than 0.5×AWUP or greater than 2×AWUP, all originators were edited if the price was less than 0.3×AWUP or greater than 2×AWUP, and all generics were edited if the price was less than 0.15×AWUP or greater than 3×AWUP. Starting with the 2009 data, the upper outlier threshold returned to 10×AWUP. Starting with the 2009 data, third-party payment data were collected in a manner that improved the ability to distinguish between missing and zero payments. In 2009, the lower threshold for generics was 0.3, but it was changed back to 0.15 starting with the 2010 data. Prior to the 2009 data, a preliminary editing rule reset missing third-party payment amounts to zero if the household respondent did not report the person had that type of insurance. Validation studies have found, however, that sometimes insurance coverage is misreported in the MEPS (Hill 2007–08; Hill, Zuvekas and Zodet 2012; State Health Access Data Assistance Center 2010), so those missing payments are no longer set to zero.

[14] Prior to the 2010 data, the acquisitions with complete or partial data were flagged for editing, rather than imputation. For these acquisitions, from 1996 through 2006, the price or quantity was edited so that the edited RUP equaled the AWUP. For the 2007 through 2009 data, the price (and sometimes the quantity) was edited so that the edited RUP equaled the mean RUP found in the data that was not flagged for editing. While a clear improvement over prior years, that procedure concentrated

payment data, 28,689 of 42,301, were flagged as lower outliers for imputation. For acquisitions with partial payment data not identified as lower or upper outliers, missing payments were set to zero.

In 2011, a small number of acquisitions were identified as upper outliers. Very few of these needed imputation, for two reasons: First, some of these were reportedly acquisitions of one pill (capsules, tablets, lozenges, etc.); the quantity of one likely represents the number of containers instead of pills. In these cases, the price was likely correct and the quantity incorrect. Therefore, the quantity was increased to a multiple of 30 that is consistent with the price per acquisition and the AWUP. Second, the price per unit may be high, but the price per acquisition low (under $16). These low prices per acquisition were not flagged for imputation, because they were consistent with chain pharmacy prices for generics. The remaining upper outliers were flagged to receive imputed prices.

Imputing price and payments

Donors for price imputation are the acquisitions with complete price data not identified as outliers and the acquisitions with partial price data whose missing values are set to zero. Four attempts are made to match a donor acquisition to a recipient acquisition. The first attempt requires an exact match on NDC and the set of payers, and the second requires an exact match on NDC. The third attempt requires an exact match on 14-digit GPI and patent status and a weighted match on drug name. The fourth attempt requires an exact match on patent status and a weighted match on therapeutic class, active ingredients, and drug name. All four attempts include the following weighted match variables: third-party payer, the person's age and private and public HMO enrollment, pharmacy name, pharmacy chain name, quantity, state, census division, and region. The weight for each variable is based on its relative contribution to the variation in both price and the proportion of the price paid out of pocket.[15]

The imputed price is the product of the recipient acquisition's quantity dispensed and an imputed RUP. The calculation of the imputed RUP depends on the quality of the match. If there is an exact match on NDC, then the imputed RUP is the donor acquisition's RUP. In the remaining matches, the NDC differs between the donor and recipient acquisitions, so the imputed RUP is based on the donor's ratio of RUP to AWUP and the recipient's AWUP[16]:

$$\text{Imputed RUP} = (\text{donor's RUP} \div \text{AWUP}) \times (\text{recipient's AWUP})$$

the distribution of RUPs for each NDC at the mean. Imputing RUPs allows for more variation across payers and other factors and more price dispersion generally.

[15] For the years 1996 through 2009, the second, third, and fourth attempts all required exact matches on payers. An exact match on therapeutic subclass was required in the second attempt. An exact match on drug group was required in the third attempt, and a weighted match was used in the fourth attempt. Beginning with the 2011 data, in order to prevent inappropriate imputation of donut hole discounts, prices of drugs that appear to have been in or after the Medicare Part D donut hole cannot match to drugs that were not in or after the donut hole.

[16] For the years 1996 through 2008, the price for the fill, not the RUP, was imputed. Then a new RUP based on the imputed price was calculated and compared with the AWUP to determine whether additional editing was necessary. In particular, if the match was exactly on NDC and the RUP was an outlier, then the quantity was imputed from the donor. Starting with the 2009 data, the RUP was imputed for all exact matches on NDC, so that the pharmacy-reported quantity was preserved. For 1996 through 2009, outliers resulting from matches on other criteria (see footnote 15) went through the editing rules described in footnotes 13 and 14.

There are four exceptions to basing the imputed price on an imputed RUP. First, for the partial payment cases, the imputed price can be less than the recipient acquisition's sum of the reported payments. In these cases the final price is the recipient's sum of reported payments, and the missing values for payments are set to zero. Second, for upper outliers, when the imputed price is less than $2, then the sum of payments reported by the recipient's pharmacy is used. Third, sometimes the imputed price for lower outliers and acquisitions without any payment information is very high. In these cases, the price reported by the recipient's pharmacy is used, and any missing values for payments are set to zero, but if there is no payment data, the price is set to the imputed RUP. The imputed price thresholds (shown in table 8) for this third exception are higher for single source drugs than generics. Examination of the data has found that a better balance between over- and under-editing the data results from setting higher thresholds for pills than for medications sold in units that are less straightforward. Due to the lack of precision in price imputation for acquisitions with imputed NDCs or imputed quantities, a lower threshold is used in these cases.[17] Fourth, for recipient acquisitions of brand name drugs that appear to be in the Medicare Part D donut hole and the donor did not appear to be in the donut hole, then the donor's price is halved to reflect the donut hole discount.

How the price is distributed by payer depends on whether the recipient acquisition has any payment information. For recipient acquisitions with partial payment information, the additional price above the known payments is assigned to the payer with missing payment amounts. For recipient acquisitions with no payment information, the donor acquisition supplies both the unit price and a distribution of out-of-pocket and third-party payment amounts. The donor's distribution is sometimes modified using the following hierarchy of rules. The number of acquisitions in 2011 edited using each rule is shown in parentheses.

- For the perfect matches on NDC and payers, the donor's proportions by payer are applied to the imputed price (4,695 acquisitions).
- For acquisitions from federal pharmacies, the out-of-pocket amounts are set using program rules and the amount paid out of pocket for the last acquisition of any drug in the round as reported in the HC (333 acquisitions).
- If the recipient's pharmacy did not report a third-party payer, then the out-of-pocket amount is set to the entire imputed price (827 acquisitions).
- If the recipient's acquisition appears to be in the Medicare Part D donut hole, then the price is allocated between out-of-pocket and Medicare based on program rules. For 2011, consistent with the rules in place in that year, the out-of-pocket amount was set to the imputed price (56 acquisitions).

---

[17] For the 1996 through 2006 data, the threshold was $200 and often the reported quantity was decreased to obtain reasonable RUPs for the recipient's reported price. For 2007 and 2008 the thresholds were: $8,000 for single source brand name pills with pharmacy-reported quantities and NDCs, $2,000 for other single source drugs, $3,000 for originator pills with pharmacy-reported quantities and NDCs, $1,400 for generic pills with pharmacy-reported quantities and NDCs, and $1,000 for other acquisitions.

- If Medicare Part D catastrophic coverage (beyond the donut hole) appears to apply to the recipient's acquisition, then 95 percent of the price is allocated to Medicare and the rest to out-of-pocket (57 acquisitions).[18]
- In the remaining matches, the donor and recipient have different payers. The recipient acquisition's out-of-pocket amount is set equal to the donor's proportion paid out of pocket multiplied by the imputed price. The third-party amount is set to the imputed price less the out-of-pocket price (1,090 acquisitions).

For lower outlier acquisitions with complete payment data, the increase in price due to imputation is an allocation based on recipient's pharmacy-reported third-party payments. If the pharmacy reported a third-party payment, then the increase is allocated to that payer. If the pharmacy did not report any third-party payments, the increase is allocated to a third-party payer based on household-reported insurance coverage and usual third-party payer, and based on the pharmacy-reported payers for the person's other acquisitions.[19]

For upper outlier acquisitions, the reduction in price is taken first from the recipient's third-party payment and, if necessary, from the out-of-pocket amount.

Results of Preparing the Pharmacy-Reported Data

The information collected in the PC about an acquisition—the date of the transaction, the NDC (or analogous identifying drug information), the quantity, the number of days of the drug that the acquisition supplied, and payments to the pharmacy by source—now also includes the GPI. Imputation has supplied values for any missing NDC or quantity dispensed. Logical edits have determined third-party payers if necessary, and missing or unrealistic prices have been replaced with imputed values. With these enhancements the PC data are now ready to be linked to the drug names reported in the HC.

Matching Pharmacy Data to Household Data

Overview of Matching

Two general approaches accomplish matching the data reported by pharmacies to the data reported by households. First, for each of a person's acquisitions in the HC, an attempt is made to find the same or

---

[18] Prior to 2011, no effort was made to identify acquisitions that might be in the donut hole or in the range for catastrophic coverage; the donut hole and catastrophic coverage edits were first used in the 2011 data.

[19] When there are multiple third-party payers, a hierarchy is used to determine which is chosen. If the recipient is a federal pharmacy, then the third-party payer is selected based on the pharmacy name. Otherwise, the order is: Medicare, Medicaid, private, CHAMPVA, TRICARE©, other federal, other state or local, Workers' Compensation, other. Prior to the 2007 data, federal pharmacy names were not used, and instead all acquisitions by honorably discharged veterans were treated as if they were eligible for VA payments, regardless of where the drug was purchased. Prior to the 2008 data, CHAMPVA was included with TRICARE.

similar acquisition obtained by that person in the PC. If this approach fails, the second approach matches pharmacy data from some other person. Because data from some other person supplies values for the missing pharmacy data, the second approach constitutes imputation.

The matching procedures are conducted at the level of the person-round-drug, rather than the acquisition. In the HC data, a "drug" is a unique drug name reported for the person in the round. In the PC data, a "drug" is a set of acquisitions of pharmaceutically equivalent drug products identical in the active ingredients, dosage form, and strength (14-digit GPI), whether brand name or generic, by the person in the round. That is, the acquisitions in the PC data are aggregated to the drug level in order to mirror the structure of the HC data.[20] In 2011, 132,874 person-round-drugs in the HC represented 348,520 acquisitions; in the PC, 252,176 acquisitions aggregated to 124,322 person-round-drugs.

After matching a PC person-round-drug to an HC person-round-drug to create a pair, the PC acquisitions within the aggregated set are unrolled, the HC drug is fanned out into a set of acquisitions, and each HC acquisition is paired with a PC acquisition in the drug set. If the number of acquisitions differs between the HC and PC, then the number of acquisitions is determined by the HC and some randomization is used to allocate the PC acquisitions to the HC acquisitions.

Matching diabetic supplies and equipment is similar. Pharmacy-reported acquisitions of diabetic supplies and equipment are mostly aggregated into one "drug" record for each person-round, so that the PC data structure parallels the HC. Aggregating in this way before matching generates a more accurate representation of the variety of diabetic supplies purchased.[21]

Details of Matching

The first approach to matching—within the person—entails three attempts. First, the procedure seeks PC data exactly matching on person, round, and pharmaceutically equivalent drug (14-digit GPI) for each household-reported drug name. Once a match is made, the PC donor person-round-drug is removed from the donor pool and not matched with any other household-reported drugs for the individual (i.e., the matches are made without replacement). In the second attempt, the residual household- and pharmacy-reported drugs must match exactly on person and round. Weighted match variables, which are not required to match exactly, are the drug group, therapeutic class, active ingredients, and medication name. An HC-PC pair is accepted only if exactly matching on therapeutic class or the reported medication name, or it otherwise meets the quality standard established for the match attempt. For the HC residuals of the second attempt, the requirement of matching within the round is removed, and all the person's pharmacy-reported drugs are candidate donors. This third attempt

---

[20] Households very rarely report multiple drug names in the same 14-digit GPI, so HC drug names do not need to be aggregated for matching. For the data prior to 2008, all PC acquisitions were aggregated to the person-round-NDC level, and the change to person-round-14 digit GPI increased the number of PC drugs matched to the HCs.

[21] From 1996 through 2009, PC diabetic supplies and equipment were aggregated separately by type of supply (for example, types of syringes, glucometers, lancets, and test strips). Due to the way household-reported diabetic supplies and equipment were coded for matching, this resulted in homogenous and sometimes unlikely purchases within person-rounds for the data prior to 2010.

requires exact matches on person and active ingredients; dosage form, strength, and medication name are weighted match variables.[22] In the third attempt, matches are made with replacement.

One variable in the weighted match is medication names, and matching medication names requires specialized software. The data processing subcontractor, Social & Scientific Systems, Inc., developed software to match names allowing for misspelling and names that sound similar but are spelled differently. The best match is found according to the following hierarchy:

- The names are exactly the same.
- The names sound the same using a Soundex function, which indexes names by how they sound rather than how they are spelled.
- A pair of characters is switched in the third to last characters (typographical errors occur more often near the end of words).
- Only one character is different in the third to last characters.
- The shorter name is exactly the same as the first characters in the longer name.
- The names begin with the same characters.
- None of the above.

To resolve ties, except when the names start with the same characters, the longer name from the PC is used. When the names start with the same characters, the name with the greatest number of matching sequential characters is used, and then the longer name is used if there are still ties. Once the best match is found, it is assigned a score measuring how closely the names matched. An exact match is given the highest score regardless of the size of the name. For a non-exact match where the names start with the same characters, the score is higher if a larger proportion of the words match. For non-exact matches where the names do not start with the same characters, higher scores are assigned to matches between longer names.[23]

Matching PC and HC drugs within person leaves some HC drugs unlinked to PC data. Some sample persons lack pharmacy data altogether, and some HC drug names remain unmatched for other reasons (for example, the household respondent failed to mention a pharmacy). A second approach is needed for this reason.

The second approach to matching—imputation from a different person—entails five attempts. It draws on a donor pool of all pharmacy-reported drugs regardless of person (excluding specific free drugs, which are reconciled later). PC drugs with imputed NDCs are part of the donor pool as well. The first attempt requires an exact match on active ingredient, dosage form, and strength (14-digit GPI).[24] The

---

[22] The third attempt, matching on person and drug, was added starting with the 2008 data.

[23] This matching algorithm is also used in matching the PC to the MDDB in order to impute NDC (when missing).

[24] Starting with the 2008 data, less weight was put on medication name in the first attempted match in order to better reduce matching too many originators from the PC due to households reporting brand names for generics, for example "Tylenol" instead of "acetaminophen." Mail-order pharmacy variables and additional insurance status variables were added to the match procedure in the same year. More weight was put on insurance status and potential payers to improve match quality and reduce subsequent source of payment editing.

second attempt requires an exact match on active ingredient and dosage form (12-digit GPI).[25] The third requires an exact match on active ingredient (8-digit GPI), and the fourth requires an exact match on drug group (4-digit GPI). The fifth attempt does not require any exact matching. Weighted match variables used in all five attempts are: the medication name, number of months per acquisition in the round, insurance status and potential payment sources for the person, name of the pharmacy, whether the HC respondent reported the person used any mail-order pharmacies, and the person's age, sex, medical conditions, geographic region and division, urbanicity, employment status, and self-reported health status. The type of health insurance coverage is used because it affects prices and rates of generic substitution.

In 2011, a match variable was added to all five attempts to account for variation in the level of benefits, which depends on the person's intensity of health care utilization. For example, insurance benefits vary depending on whether someone has met their deductible for the year. A proxy for the person's utilization is the cumulative number of HC-reported acquisitions of all drugs in the prior and current rounds of the calendar year. This variable accounts somewhat for differences in benefit generosity between high and low users such as the donut hole and out-of-pocket maximums.

After matching drug name to drug name, the HC and PC drug records are expanded into acquisitions, as mentioned above. Each drug name reported in the HC interview is fanned out to the number of acquisitions the household reported; each PC drug, a set of aggregated PC acquisitions, is unrolled back into distinct acquisitions as originally reported in the PC. Then each HC acquisition is paired with a PC acquisition within the drug-to-drug matched set. Because the household actually reports only the number of acquisitions, they have no natural order. Therefore, the acquisitions take the date order of the PC acquisitions; the order is preserved only in the unique record identifiers.[26] When the household and pharmacy report the same number of acquisitions, then the pairing is one-to-one by the date order in the PC. Otherwise, if the number of acquisitions differs between the HC and PC, then the PC acquisitions are paired with the HC acquisitions in a randomization process that follows the date order of the PC and the unique identifiers of the acquisitions in the HC. Note that when the drug match is imputed, the order of acquisitions represented by the record identifier may not be analytically useful because the date order is from a different person or round.

Table 9 summarizes the results of matching pharmacy data to household data in 2011. More than two-thirds (68.5 percent) of drugs were acquired by those sample members who had any pharmacy data. Matching was highly successful for these drugs: 83.7 percent (line 5) of their acquisitions were from their own pharmacies. The imputed pharmacy matches—acquisitions not matched within the person—made good use of the information given by household respondents: 96.4 percent (line 10) were matches to the active ingredient of the drug reported by the household.

In the 2011 PC donor database, 42 percent of acquisitions remained unmatched to the HC. A validation study found that household respondents reported as many acquisitions as were found in claims data,

---

[25] This attempt was added starting with the 2008 data.
[26] Prior to the 2008 data, PC acquisitions were selected at random with replacement. The change in 2008 resulted in more PC data being matched to the HC.

but households reported fewer drugs and more acquisitions per drug. The drugs acquired but not reported by the household tended to be for short-term use. They entailed fewer acquisitions and included many anti-infectives, topical agents, and pain medications (Hill, Zuvekas, Zodet 2011).

Editing Matched Data

This section describes six sets of edits that apply to the matched HC-PC data, including supplemental data merged onto the file to enhance analytic utility. The six sets of edits are:

- Free antibiotics, anti-diabetics, and prenatal vitamins. This edit allows a price of zero for certain free acquisitions, such as antibiotics, from chain pharmacies offering free drug programs.
- Federal pharmacy prices. This edit improves the accuracy of prices of acquisitions from federal pharmacies.
- Reconciling payments for self-filers and diabetic products. This edit reconciles the household- and pharmacy-reported payment information.
- Editing year in Round 3. This edit allocates purchases reported in the third interview, which covers parts of two calendar years, to each year.
- Medicare Part D and private insurance. This edit changes the payer from private insurance to Medicare Part D for certain people's acquisitions.
- Resolving payer inconsistencies arising from imputation. This edit reconciles differences in payers related to having matched across persons.

Free antibiotics, anti-diabetics, and prenatal vitamins

Under certain conditions, payment data are edited for the free antibiotics and prenatal vitamins that pharmacies report. Among the household-reported acquisitions matched to a different person's pharmacy data, all the price and payment variables are set to zero for selected free antibiotics, anti-diabetics, and vitamins. This editing rule applies only if the household respondent reported the person used only one pharmacy, that pharmacy chain had a free antibiotic, antidiabetic, or prenatal vitamin program, and the antibiotic, anti-diabetic, or prenatal vitamin was included in the chain's program. In addition, to maintain the pharmacy chain's confidentiality for prices of other purchases, the person must have resided in a state with two or more pharmacy chains with free programs.[27] In the 2011 data, 46 acquisitions were edited to a price of zero.

Federal pharmacy prices

A federal price list is used to determine prices for acquisitions from federal pharmacies.[28] For each NDC, the lowest price on the list is used. These prices reflect the government's cost of acquiring medicines, so an estimate of dispensing costs is added.

---

[27] This editing rule was first implemented for the 2008 data, the year free antibiotic programs began to spread.

[28] Prior to the 2008 data, prices were imputed from other PC acquisitions.

Federal prices are assigned to acquisitions in three situations. These situations are defined by information from both the HC and PC.

- The PC data are from a federal pharmacy that did not report payment information and the HC and PC acquisitions match on person and round.
- The household respondent reported the person used only federal pharmacies in that round.
- The household respondent reported the person used at least one federal pharmacy in that round and the PC data are from a federal pharmacy that did not report payment information.

In these three situations, the federal price is merged onto the matched HC-PC acquisition by the pharmacy-reported NDC. However, sometimes the NDC from the PC does not exactly match a drug on the federal price list—because the NDC is imputed, or because there is no negotiated price with the manufacturer of the drug. In these cases, instead of NDC, the match is by active ingredient, dosage form, and strength (14-digit GPI) with drug name as a weighted match variable. The pharmacy-reported NDC, whether reported or imputed, is replaced with the one selected from the federal price list. The remaining cases retain the price from the PC (whether reported or imputed), because the federal price schedule does not include every drug.

For the acquisitions with prices taken from the federal price list, the out-of-pocket and federal payments are set using federal program rules and the amount the household respondent reported as the out-of-pocket payment for the last acquisition of any drug in the round.

In the 2011 data, 1,717 acquisitions were assigned federal prices.

Reconciling payments for self-filers and diabetic products

For most HC family members, only the out-of-pocket payment for the last acquisition of any drug in the round is collected in the HC interview. For self-filers, however, complete payment information by source for the last acquisition of each drug is requested, as it is for diabetic supplies and equipment. Therefore, two sets of information about payments are available in these cases—one from the household and the other from the pharmacy. MEPS reconciles the two sets of payment information based on their relative reliability.

Prior to reconciliation, the HC payment data are reviewed to ascertain whether there were reporting errors in payment amounts. The process is an abbreviated version of the process used for the PC. The household-reported price (sum of reported payments) is compared to the AWUP calculated from the MDDB using the quantity and patent status from the pharmacy. Specifically, the HC RUP is calculated as the sum of household-reported payments divided by the pharmacy-reported quantity.[29] The thresholds are the same as those applied to the PC (table 6).

---

[29] This RUP is different from the RUP based entirely on PC data, which was used to edit the PC data.

- If no household payment information is missing and the RUP ≥ AWUP × 0.65, 0.20, or 0.03, depending on the product's patent status, then editing for the acquisition is complete (506 acquisitions).
- If one or more household-reported payment sources is missing; the sum of payments is within $5 of the total charge; and RUP ≥ AWUP× 0.75, 0.70, or 0.15, depending on the product's patent status (see table 6); then the missing payment amounts are set equal to zero (30 acquisitions).[30]

The following hierarchy of rules is applied to reconcile the payments reported by the household and pharmacy.

- If all the HC payment data are missing or zero, then the PC payment data are used (5,282 acquisitions).
- If at least one HC payment source is missing and the sum of HC payments equals the PC, then the HC payments are used and the missing value is set to zero (61 acquisitions).
- If at least one HC payment source is missing and the sum of HC payments exceeds the PC, then the PC price is used. The price is allocated to payers using the proportions reported in the HC (370 acquisitions).
- If at least one HC payment source is missing and the sum of HC payments is less than the PC price, then the PC price and HC payment amounts are used. The missing HC amount is set equal to the difference between the price and the sum of the HC payments (908 acquisitions). [31]

Editing year in Round 3

In the third interview, in which the reference period spans the later part of one year and the early part of the next, for each drug name, the household respondent is asked to report both the number of times the drug was obtained since the last interview and the number of those times the drug was obtained in the previous year.[32] When this information is missing for a drug, the number of acquisitions is allocated to each year. The allocation is based on the date the person started taking the medication and, for drugs with PC data exactly matched to the HC on person and round, the number of acquisitions reported by pharmacies in each year.[33] Otherwise, acquisitions are distributed in proportion to the duration of

---

[30] Prior to the 2008 data, these two edits did not vary with patent status, quantity dispensed, or AWUP; a minimum $30 sum of payments was used instead. For the 2008 data, the thresholds for both edits were 0.5, 0.3, 0.15 (see footnote 13).

[31] If more than one HC payment source is missing, then the difference is assigned to the first missing payment source in the following descending order: Medicare, Medicaid, private insurance, CHAMPVA, TRICARE, other Federal coverage, State or local coverage, Workers' Compensation, other insurance, or out-of-pocket.

[32] The second question, to determine the number of fills in the last year, was added to the 2000 MEPS HC.

[33] When the household reports no more than the number of fills reported by the pharmacies for the prior year, all those fills are allocated to the prior year. Prior to 2008, the household-reported fills in excess of those reported by the pharmacies for the prior year were randomly assigned to either year in proportion to the duration of Round 3 in each year. Starting with the 2008 data, these fills were assigned to the current year. The previous method accounted for persons for whom not all pharmacies provided data, but most persons reportedly used only one pharmacy. The change for the 2008 data has no impact on aggregate analyses using annual data. The documentation for the Panel 12 longitudinal file (HC-122) describes the types of analyses for which the change may be important and contains detailed suggestions for data users.

Round 3 in each year. After this allocation, there were 313,747 acquisitions in 2011, and this is the number of acquisitions (records) on the public use file.

Medicare Part D and private insurance

For Medicare beneficiaries, some private insurance payments are assumed to be Medicare Part D. This edit only applies to acquisitions for which the pharmacy data are matched from the beneficiary's pharmacies, not imputed from another person. For a beneficiary who has an acquisition with a private payment and, according to the HC, has Part D coverage, the private payment is assumed to be from a Medicare Part D plan. In addition, for an elderly person who has an acquisition with a private payment and, according to the HC, lacks private drug coverage, the private payment is assumed to be from a Medicare Part D plan. This edit applied to 3,017 acquisitions in 2011.

An error occurred for the years 2006 and 2007. Private payments were also assumed to be Medicare Part D for elderly persons with private insurance, again only for acquisitions with pharmacy data matched from the beneficiary's pharmacies, not imputed from another person. This error caused a substantial over-representation of Medicare expenditures and under-representation of private expenditures in the 2007 data.

Resolving payer inconsistencies arising from imputation

As addressed above, matching pharmacy data to drug names reported in the HC is based primarily on the drug group, therapeutic class, active ingredients, dosage form, and strength. Additional criteria include insurance status and potential third-party payers (although, due to small cell sizes, these variables are not required to match exactly in the second approach to matching, imputation from another person). Despite high-quality drug-to-drug matching, inconsistency in payment patterns can result. The donor acquisition and recipient acquisition may have different payers. For example, the pharmacy- (donor-) reported third-party payments can be from an insurer that, according to the HC, did not provide coverage to the sample person (recipient). This section describes the methods used to identify and resolve these inconsistencies

Acquisitions with inconsistencies are defined as those for which the HC-reported insurance coverage differs between the donor drug record and recipient drug record.[34] Both the donor and recipient records are classified using a hierarchy of combinations of insurance (see table 10). The classification reflects common situations that relate to the types of third-party payers and amount paid out of pocket. The most common situation is having only one type of insurance for the entire round, but the

---

[34] Between 1996 and 2007, inconsistencies were defined directly as differences between types of insurance and specific payments (Moeller et al. 2001). However, because acquisitions paid entirely out of pocket were always consistent, this method was biased toward out-of-pocket spending. Improvements in the resolution of inconsistencies, which were first implemented for the 2008 data, also better account for the complexity of individuals' insurance coverage, increase the prevalence of copayments relative to coinsurance, and allow for payments by insurers not reported by the household, because households do not perfectly report insurance status (Hill 2007–08; Hill, Zuvekas, Zodet 2012; State Health Access Data Assistance Center 2010).

classification accounts for other situations as well.[35] When the donor and recipient have different insurance classifications, they are considered inconsistent.

Most situations of inconsistency between a donor's sources of payment and the recipient's insurance require editing or additional imputation for resolution. However, resolving inconsistency between insurance and payers is unnecessary in several situations. First, acquisitions with prices taken from the federal price list are defined as consistent, because third-party payments have already been edited. Second, acquisitions edited using HC payment data are also consistent, because only the HC data are used. Third, free acquisitions are consistent, because there are no payments.

In three other rare situations, editing makes imputation unnecessary. First, if the person has only one acquisition in the round then the information collected in the HC about the usual third-party payer and last out-of-pocket amount is used and the PC source-of-payment data are discarded (1,337 acquisitions in 2011). Second, payments for Medicare Part B drugs for Medicare beneficiaries are allocated to Medicare (63 acquisitions). For acquisitions that appear to occur after the person has entered the Medicare Part D donut hole, 95 percent of the price is allocated to Medicare and 5 percent to out-of-pocket (6 acquisitions).

The method for correcting the remaining inconsistencies (39,285 acquisitions) is to impute the distribution of payments from an acquisition with consistent payments using a hot-deck procedure. The donor pool is composed of acquisitions with consistent payments: those where the PC data were matched within the person to the HC as well as those for which consistency was imposed by editing, with the exception of free acquisitions, because they lack payment data. A donor is drawn from this donor pool with replacement. Class variables are, in order of importance: insurance classification, price categories, patent status, whether any mail-order pharmacies were used by the person in the round, drug group, therapeutic class, active ingredients, dosage form, and strength. Except for insurance status, if there are fewer donors than recipients in a cell, then cells are collapsed until the ratio of donors to recipients is at least 1:1.

Payment information from the new donor acquisition is used to impute third-party amounts and either copayments or out-of-pocket coinsurance payments. Copayments, including values of zero, are always imputed for Medicaid, the VA, and TRICARE. For example, if the donor has a Medicaid payment, then the recipient acquisition's out-of-pocket amount is set equal to the donor acquisition's, and the recipient Medicaid payment is set equal to the recipient price less the donor out-of-pocket amount, preserving the total. For the remaining third-party payers, either a copayment or coinsurance is imputed, depending on the whether the donor's out-of-pocket amount is a whole number. If the donor acquisition is generic and the out-of-pocket amount is a whole number, or the out-of-pocket amount is a round multiple of $5, then the copayment is imputed, and the third-party amount is set as the difference between the price and copayment, preserving the total. Otherwise, the donor acquisition is treated as a coinsurance case. The proportion of the donor acquisition's price paid out of pocket and by

---

[35] Starting with the 2011 data, two insurance categories were added to improve consistency related to the Medicare Part D donut hole.

each third-party payer is calculated. The recipient acquisition's price is distributed in the same proportions to calculate the recipient out-of-pocket amount and payments by each third-party payer.

Additional variables

Two sets of variables are merged onto the dataset for release to the public: therapeutic classes and FDA pregnancy categories. The source is Cerner Multum, Inc.'s Multum Lexicon database.[36] These variables are merged onto the file by NDC.

Editing for confidentiality

Prior to releasing the data to the public, the data are reviewed by a pharmacist consultant to ensure the confidentiality of the sample members. Drugs that are rarely used or are associated with very rare conditions, particularly orphan drugs, are censored. In these cases, the drug name is replaced with a more general therapeutic class name and the NDC is set to missing. Confidentiality protection affected less than 1 percent of acquisitions in 2011.

---

[36] These variables were first added to the prescription drug PUF for the 2002 data. For previous years, they are available on a supplementary data file (HC-068).

References

Cohen, J. *Design and Methods of the Medical Expenditure Panel Survey Household Component.* MEPS Methodology Report No. 1. AHCPR Pub. No. 97-0026. Rockville, MD: Agency for Health Care Policy and Research, 1997. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr1/mr1.pdf

Cohen, J., Taylor, A. The provider system and the changing locus of expenditure data: survey strategies from fee-for-service to managed care. *Informing American Health Care Policy: The Dynamics of Medical Expenditure and Insurance Surveys, 1977–1996 (Chapter 4),* 57–62. The Jossey Bass Health Series—First Edition.

Cohen, S. *Sample Design of the 1996 Medical Expenditure Panel Survey Household Component.* MEPS Methodology Report No. 2. AHCPR Pub. No. 97-0027. Rockville, MD: Agency for Health Care Policy and Research, 1997. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr2/mr2.pdf

Cohen, S. Design Strategies and Innovations in the Medical Expenditure Panel Survey. *Medical Care,* July 2003: 41(7) Supplement: III-5–III-12.

Cohen S., Carlson, B. A Comparison of Household and Medical Provider Reported Expenditures in the 1987 NMES. *Journal of Official Statistics,* Vol. 10, No. 1, 1994, p 3–29.

Ezzati-Rice, T.M., Rohde, F., Greenblatt, J. *Sample Design of the Medical Expenditure Panel Survey Household Component, 1998–2007.* Methodology Report No. 22. Rockville, MD: Agency for Healthcare Research and Quality, 2008. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr22/mr22.pdf

Hill S.C. 2007–08. "The Accuracy of Reported Insurance Status in the MEPS." Inquiry 44(4): 443-468.

Hill, S.C., Zuvekas, S.H., Zodet, M.W. "Implications of the Accuracy of MEPS Prescription Drug Data for Health Services Research." *Inquiry*, Vol. 48, No. 3, 2011, p. 242-259.

Hill, S.C., Zuvekas, S.H., Zodet, M.W. "The Validity of Reported Part D Enrollment in the MEPS." *Medical Care Research and Review*, Vol. 69, No. 6, 2012, p. 737-750.

Machlin, S.R., Dougherty, D.D. *Overview of Methodology for Imputing Missing Expenditure Data in the Medical Expenditure Panel Survey.* Methodology Report No. 19. Rockville, MD: Agency for Healthcare Research and Quality, 2007. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr19/mr19.pdf

Machlin, S.R., Taylor A.K. *Design, Methods, and Field Results of the 1996 Medical Expenditure Panel Survey Medical Provider Component.* MEPS Methodology Report No. 9. AHRQ Pub. No. 00-0028. Rockville, MD: Agency for Healthcare Research and Quality, 2000. http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr9/mr9.pdf

Master Drug Data Base (MDDB®), Version 2.5. *Documentation Manual*. Indianapolis, IN: Wolters Kluwer Health, Inc., 2005.

Moeller J.F., et. al. *Outpatient Prescription Drugs: Data Collection and Editing in the 1996 Medical Expenditure Survey (HC-010A)*. MEPS Methodology Report No. 12. AHRQ Pub. No. 01-0002. Rockville, MD: Agency for Healthcare Research and Quality, 2001.
http://www.meps.ahrq.gov/mepsweb/data_files/publications/mr12/mr12.pdf

State Health Access Data Assistance Center, National Center for Health Statistics, Department of Health and Human Services Assistant Secretary for Planning and Evaluation, Agency for Healthcare Research and Quality, Centers for Medicare and Medicaid Services, and U.S. Census Bureau. 2010. *Phase VI Research Results: Estimating the Medicaid Undercount in the Medical Expenditure Panel Survey Household Component (MEPS-HC).* Suitland, MD: Census Bureau [accessed on March 12, 2014]. Available online: http://www.census.gov/did/www/snacc/docs/SNACC_Phase_VI_Full_Report.pdf

Zodet, M.W., Hill, S.C., Zuvekas, S.H. "Assessing the accuracy of prescription drug purchase data for Medicare beneficiaries in the Medical Expenditure Panel Survey." *American Statistical Association Proceedings*, 2010.

Zodet, M., Hill, S.C., Miller, G.E. *Comparison of Retail Drug Prices in the MEPS and Market Scan: Implications for MEPS Editing Rules*. Agency for Healthcare Research and Quality Working Paper No. 10001, February 2010.

Table 1. Summary of editing and imputation rates for key variables, 2011

Data set and unit (number)	Variable to be imputed	Percentage with editing or imputation
HC data		
Unique drug names reported for a person in a round (132,874)	Number of acquisitions	6.9%
PC data		
Unique acquisitions of a drug in a year (252,176)	NDC	5.1%
	Quantity	0.7%
	Third-party payer	26.3%
	Price	28.1%
Matched HC-PC data		
Unique drug names reported for a person in a round (132,874)	Drug details	45.2%

Table 2. Method used to impute the number of acquisitions in the Medical Expenditure Panel Survey Household Component, 2011

	Number	Percentage
Person-round-drugs requiring imputation	5,939	100.0%
Exact match to HC donor pool on:		
Active ingredients, dosage form, and strength	5,884	99.1%
Active ingredients and dosage form	22	0.4%
Active ingredients	20	0.3%
Drug group	13	0.2%

Note: A person-round-drug is a unique drug name within a person and round. Matching used Wolters Kluwer's Generic Product Identifier to identify drug names with the same active ingredients.

MEPS

Table 3. Sample size and participation in the Medical Expenditure Panel Survey Pharmacy Component, 2011

	Pharmacies		Person-pharmacy pairs	
	Number	Percentage	Number	Percentage
Eligible sample	7,420	100.0%	17,414	100.0%
Response	5,555	74.9%	12,720	73.0%
Refusal	110	1.5%	386	2.2%
Other nonresponse	1,755	23.7%	4,318	24.8%

Note: The sample for the Pharmacy Component is derived from all persons in the Household Component who signed permission forms to contact the pharmacies from which they reported obtaining drugs during the rounds that include 2011.
Person-pharmacy pairs uniquely count the number of HC sample members reported in the HC as using each pharmacy.
Other nonresponse includes unlocatable pharmacies and those that have records for the patient but not in the correct data collection year.

Table 4. Method used to impute NDC in the Medical Expenditure Panel Survey Pharmacy Component, 2011

	Number	Percentage
Acquisitions	252,176	100.0%
Acquisitions with valid NDCs	239,406	94.9%
Acquisitions missing or invalid NDCs	12,770	5.1%
Among acquisitions missing NDCs	12,770	100.0%
Imputed from same person and round in PC	1,202	9.4%
Imputed from the Master Drug Data Base	11,568	90.6%
Exact match on active ingredients, dosage form, and strength	10,949	85.7%
Exact match on active ingredients	545	4.3%
Other	65	0.5%
Imputed from an acquisition of a person with similar characteristics	9	0.1%

Note: Matching used Wolters Kluwer's Generic Product Identifier to identify drugs with the same active ingredients.

Table 5. Completeness of payment data in the Medical Expenditure Panel Survey Pharmacy Component, 2011

	Number	Percentage
Acquisitions	252,176	100.0%
Complete payment data	184,874	73.3%
Partial payment data	42,301	16.7%
Has out-of-pocket amount, missing third-party amount	41,935	16.6%
Has third-party amount, missing out-of-pocket amount	366	0.1%
All payment data missing or zero	25,001	9.9%

Table 6. Lower thresholds for identifying prices needing imputation, by payment data quality, patent status, and potential for being in the Medicare Part D donut hole, 2011

| | Complete payment data | | | | Partial payment data |
| | Third party amount is zero | Third party amount is positive (must be below all three thresholds) | | | |
	RUP as percentage of AWUP	RUP as percentage of AWUP	Price	Out of pocket	RUP as percentage of AWUP
Brand name					
Single source					
Not in donut hole	65%	65%	$10	$0.25	75%
Potentially in donut hole	40%	40%	$10	$0.25	50%
Originator					
Not in donut hole	20%	20%	$2	$0.25	70%
Potentially in donut hole	20%	20%	$2	$0.25	50%
Generic	3%	3%	$2	$0.25	15%

Note: RUP = retail unit price; AWUP = average wholesale unit price.

Table 7. Editing categories by patent status in the Medical Expenditure Panel Survey Pharmacy Component, 2011

| | | Patent status | | |
| | | Brand name | | Generic |
	Total	Single source	Originator	
Total acquisitions	252,176	55,081	7,947	189,148
Complete payment data	184,874	39,794	5,854	139,226
No editing	179,820	37,579	5,650	136,591
Lower outlier (impute RUP)	4,628	2,146	202	2,280
No third-party payment	4,297	2,009	199	2,089
Positive third-party payment	331	137	3	191
Upper outlier	426	69	2	355
One pill (increase quantity)	70	25	0	45
Price < $16 (no change)	248	13	0	235
Other (impute RUP)	108	31	2	75
Partial payment data	42,301	9,429	1,243	31,629
Missing values set to zero	13,559	815	209	12,535
Lower outlier (impute RUP)	28,689	8,607	1,034	19,048
Upper outlier	53	7	0	46
One pill (increase quantity)	15	3	0	12
Price < $16 (no change)	30	1	0	29
Other (impute RUP)	8	3	0	5
Free antibiotics, prenatal vitamins, antidiabetics, and glucometers from some pharmacies (no change)	1,182	176	8	998
Missing all payment data (impute RUP)	23,819	5,682	842	17,295

Note: RUP = retail unit price.

Table 8. Upper thresholds on imputed acquisition prices, by patent status, dosage form, and data quality, 2011

	Patent status	
Dosage form and data quality	Brand name	Generic
Pills with pharmacy-reported NDCs and quantities	$11,000	$1,400
Not pills or imputed NDCs or imputed quantities	$5,000	$1,000

Note: NDC = National Drug Code.

Table 9. Matching and imputation of pharmacy-reported drugs and acquisitions to household-reported drug names in the Medical Expenditure Panel Survey, 2011

	Description	Drugs Number	Drugs Percentage	Acquisitions Number	Acquisitions Percentage
1	Household-reported totals	132,874	100.0%	348,520	100.0%
2	Person had any pharmacy data	91,060	68.5%	241,640	69.3%
3	Person had no pharmacy data	41,814	31.5%	106,880	30.7%
4	Person had any pharmacy data (line 2)	91,060	100.0%	241,640	100.0%
5	Total matched within person (sum of lines 6, 7)	72,817	80.0%	202,197	83.7%
6	Matched within person-round	63,606	69.9%	181,979	75.3%
7	Matched within person	9,211	10.1%	20,218	8.4%
8	Not matched within person	18,243	20.0%	39,443	16.3%
9	Total not matched within person (sum of lines 3, 8)	60,057	100.0%	146,323	100.0%
	Imputation with exact match on				
10	Active ingredient (sum of lines 11, 12, 13)	57,774	96.1%	141,007	96.4%
11	Active ingredient, dosage form, strength	36,537	60.8%	92,614	63.3%
12	Active ingredient, dosage form	7,945	13.2%	18,696	12.8%
13	Active ingredient	13,262	22.1%	29,697	20.3%
14	Therapeutic class	1,188	2.0%	2,702	1.8%
15	Drug group	1,628	2.7%	3,770	2.6%
16	Imputation without exact drug match	1,125	1.9%	2,614	1.8%

Notes: In the HC, a "drug" is a unique drug name reported for the person in the round. In the PC, a "drug" is a set of acquisitions of pharmaceutically equivalent drug products identical in the active ingredients, dosage form and strength, whether brand name or generic, by the person in the round.
Matching used Wolters Kluwer's Generic Product Identifier to identify drugs with the same active ingredients. Matches are weighted and take into account drug name, types of insurance, health status and chronic conditions, census division and urbanicity, and demographics.

Table 10. Major categories of hierarchical insurance classification for payer consistency editing

Medicare Part D
  And Medicaid
  Other Low Income Subsidy (no premium)
  Likely in the donut hole based on total expenditures on drugs covered by Part D in prior rounds in the calendar year
  And private drug coverage (no Medicaid)
  Other

Medicaid (no Medicare Part D)
  The whole round, no private insurance
  Part of the round, no private insurance
  And private insurance

Private drug coverage (no Medicare Part D or Medicaid)
  Part of the round, no TRICARE
  The whole round, no TRICARE

Private insurance (no private drug coverage, Medicare Part D, or Medicaid)
  No TRICARE
  And TRICARE

TRICARE (no private insurance, Medicare Part D, or Medicaid)
Other public
Private insurance only and not elderly
Private insurance only and elderly
State program with limited benefits
Medicare only
Uninsured