MEPS HC-110A:
2007 Prescribed Medicines

October 2009

Table of Contents

A.   Data Use Agreement

Individual identifiers have been removed from the micro-data contained in these files. Nevertheless, under sections 308 (d) and 903 (c) of the Public Health Service Act (42 U.S.C. 242m and 42 U.S.C. 299 a-1), data collected by the Agency for Healthcare Research and Quality (AHRQ) and/or the National Center for Health Statistics (NCHS) may not be used for any purpose other than for the purpose for which they were supplied; any effort to determine the identity of any reported cases is prohibited by law.

Therefore in accordance with the above referenced Federal Statute, it is understood that:

1. No one is to use the data in this data set in any way except for statistical reporting and analysis; and

2. If the identity of any person or establishment should be discovered inadvertently, then (a) no use will be made of this knowledge, (b) the Director Office of Management AHRQ will be advised of this incident, (c) the information that would identify any individual or establishment will be safeguarded or destroyed, as requested by AHRQ, and (d) no one else will be informed of the discovered identity; and

3. No one will attempt to link this data set with individually identifiable records from any data sets other than the Medical Expenditure Panel Survey or the National Health Interview Survey.

By using these data you signify your agreement to comply with the above stated statutorily based requirements with the knowledge that deliberately making a false statement in any matter within the jurisdiction of any department or agency of the Federal Government violates Title 18 part 1 Chapter 47 Section 1001 and is punishable by a fine of up to $10,000 or up to 5 years in prison.

The Agency for Healthcare Research and Quality requests that users cite AHRQ and the Medical Expenditure Panel Survey as the data source in any publications or research based upon these data.

<center>**B.   Background**</center>

1.0   Household Component (HC)

The Medical Expenditure Panel Survey (MEPS) provides nationally representative estimates of health care use, expenditures, sources of payment, and health insurance coverage for the U.S. civilian non-institutionalized population. The MEPS Household Component (HC) also provides estimates of respondents' health status, demographic and socio-economic characteristics, employment, access to care, and satisfaction with health care. Estimates can be produced for individuals, families, and selected population subgroups. The panel design of the survey, which includes 5 Rounds of interviews covering 2 full calendar years, provides data for examining person level changes in selected variables such as expenditures, health insurance coverage, and health status. Using computer assisted personal interviewing (CAPI) technology, information about each household member is collected, and the survey builds on this information from interview to interview. All data for a sampled household are reported by a single household respondent.

 The MEPS-HC was initiated in 1996. Each year a new panel of households is selected. Because the data collected are comparable to those from earlier medical expenditure surveys conducted in 1977 and 1987, it is possible to analyze long-term trends. Each annual MEPS-HC sample size is about 15,000 households. Data can be analyzed at either the person or event level. Data must be weighted to produce national estimates.

The set of households selected for each panel of the MEPS HC is a subsample of households participating in the previous year's National Health Interview Survey (NHIS) conducted by the National Center for Health Statistics. The NHIS sampling frame provides a nationally representative sample of the U.S. civilian non-institutionalized population and reflects an oversample of blacks and Hispanics. MEPS oversamples additional policy relevant sub-groups such as Asians and low income households. The linkage of the MEPS to the previous year's NHIS provides additional data for longitudinal analytic purposes.

2.0   Medical Provider Component (MPC)

Upon completion of the household CAPI interview and obtaining permission from the household survey respondents, a sample of medical providers are contacted by telephone to obtain information that household respondents can not accurately provide. This part of the MEPS is called the MEPS Provider Component (MPC) and information is collected on dates of visit, diagnosis and procedure codes, charges and payments.   The Pharmacy Component (PC), a subcomponent of the MPC, does not collect charges or diagnosis and procedure codes but does collect drug detail information, including National Drug Code (NDC) and medicine name, as well as date filled and sources and amounts of payment. The MPC is not designed to yield national estimates. It is primarily used as an imputation source to supplement/replace household reported expenditure information.

3.0   Survey Management and Data Collection

MEPS HC and MPC data are collected under the authority of the Public Health Service Act. Data are collected under contract with Westat, Inc. Data sets and summary statistics are edited and published in accordance with the confidentiality provisions of the Public Health Service Act and the Privacy Act. The National Center for Health statistics (NCHS) provides consultation and technical assistance.

As soon as data collection and editing are completed, the MEPS survey data are released to the public in staged releases of summary reports, micro data files, and tables via the MEPS web site: www.meps.ahrq.gov .   Selected data can be analyzed through MEPSnet, an on-line interactive tool designed to give data users the capability to statistically analyze MEPS data in a menu-driven environment.

Additional information on MEPS is available from the MEPS project manager or the MEPS public use data manager at the Center for Financing Access and Cost Trends, Agency for Healthcare Research and Quality, 540 Gaither Road, Rockville, MD 20850 (301-427-1406).

C.   Technical Information

1.0   General Information

This documentation describes one in a series of public use event files from the 2007 Medical Expenditure Panel Survey Household Component (HC) and Medical Provider Component (MPC). Released as an ASCII data file and SAS transport file, this public use file provides detailed information on household reported prescribed medicines for a nationally representative sample of the civilian noninstitutionalized population of the United States and can be used to make estimates of prescribed medicine utilization and expenditures for calendar year 2007. As illustrated below, this file consists of MEPS survey data obtained in the 2007 portion of Round 3 and Rounds 4 and 5 for Panel 11, as well as Rounds 1, 2 and the 2007 portion of Round 3 for Panel 12 (i.e., the rounds for the MEPS panels covering calendar year 2007).

2006			2007			2008
Jan			Jan		Dec	Jan

Panel 11 2006-2007: Round 1 | Round 2 | | Round 3 | Round 4 | Round 5

Panel 12 2007-2008: Round 1 | Round 2 | Round 3
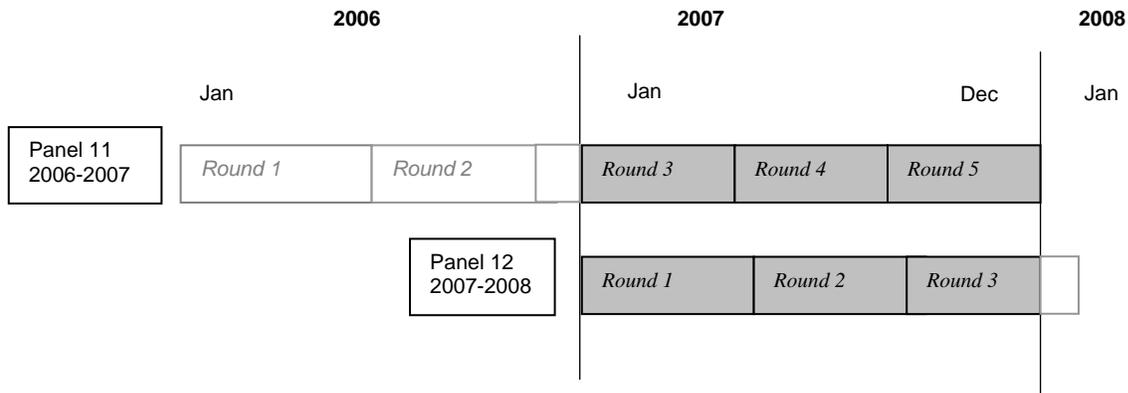
Each record on this event file represents a unique prescribed medicine event; that is, a prescribed medicine reported as being purchased by the household respondent. In addition to expenditures related to the prescribed medicine, each record contains household reported characteristics and medical conditions associated with the prescribed medicine.

Data from this event file can be merged with other 2007 MEPS-HC data files, for purposes of appending person characteristics such as demographic or health insurance coverage to each prescribed medicine record.

Counts of prescribed medicine utilization are based entirely on household reports. Information from the Pharmacy Component (PC) (within the MEPS-MPC, see section B.2.0 for more details on the MPC) was used to provide expenditure and payment data, as well as details of the medication (e.g., strength, quantity, etc.).

The file can be used to construct summary variables of expenditures, sources of payment, and other aspects of utilization of prescribed medicines. Aggregate annual person-level information on the use of prescribed medicines and other health services use is provided on the 2007 Full Year Consolidated Data File, where each record represents a MEPS sampled person.

The following documentation offers a brief overview of the types and levels of data provided and the content and structure of the files and the codebook. It contains the following sections:

> Data File Information
> Sample Weight
> General Data Editing and Imputation Methods
> Strategies for Estimation
> Merging/Linking MEPS Data Files
> References
> Variable to Source Crosswalk

For more information on MEPS-HC survey design see S. Cohen, 1997; J. Cohen, 1997; and S. Cohen, 1996. For information on the MEPS-MPC design, see S. Cohen, 1998. A copy of the survey instrument used to collect the information on this file is available on the MEPS Web site at the following address: <http://www.meps.ahrq.gov>.

2.0   Data File Information

This public use data set contains 300,099 prescribed medicine records. Each record represents one household reported prescribed medicine that was purchased during calendar year 2007. Of the 300,099 prescribed medicine records, 293,986 records are associated with persons having a positive person-level weight (PERWT07F). The persons represented on this file had to meet either criterion a or b below:

a) Be classified as a key inscope person who responded for his or her entire period of 2007 eligibility (i.e., persons with a positive 2007 full-year person-level sampling weight (PERWT07F > 0), or

b) Be classified as either an eligible non-key person or an eligible out-of-scope person who responded for his or her entire period of 2007 eligibility, and belonged to a family (i.e., all persons with the same value for a particular FAMID variable) in which all eligible family members responded for their entire period of 2007 eligibility, and at least one family member has a positive 2007 full-year person weight (i.e., eligible non-key or eligible out-of-scope persons who are members of a family, all of whose members have a positive 2007 full-year MEPS family-level weight (FAMWT07F >0).

Persons with no prescribed medicine use for 2007 are not included on this file (but are represented on MEPS person-level files). A codebook for the data file is provided (in file H110ACB.PDF).

This file includes prescribed medicine records for all household survey respondents who resided in eligible responding households and reported at least one prescribed medicine. Only prescribed medicines that were purchased in calendar year 2007 are represented on this file. This file includes prescribed medicines identified in the Prescribed Medicines section of the HC survey instrument, as well as those prescribed medicines identified in association with other medical

events. Each record on this file represents a single acquisition of a prescribed medicine reported by household respondents. Some household respondents may have multiple acquisitions of prescribed medicines and thus will be represented in multiple records on this file. Other household respondents may have reported no acquisitions of prescribed medicines and thus will have no records on this file.

When diabetic supplies, such as syringes and insulin, were mentioned in the Other Medical Equipment section of the MEPS-HC, the interviewer was directed to collect information on these items in the Prescription Medicines section of the MEPS questionnaire. The respondent was also asked the questions in the Charge and Payment section of the HC. To the extent that these items are purchased without a prescription, they represent a non-prescription addition to the MEPS prescription drug expenditure and utilization data. Although these items may be purchased without a prescription, a prescription purchase may be required to obtain third party payments. Analysts are free to code and define diabetic supply/equipment and insulin events utilizing their own coding mechanism. If desired, this would enable analysts to subset the Prescribed Medicines file to exclude these types of events.

It should also be noted that refills are included on this file. The HC obtains information on the name of the prescribed medicine and the number of refills, if any, associated with it. The data collection design for the HC does not allow separate records to be created for multiple acquisitions of the same prescribed medicine. However, in the PC, each original purchase, as well as any refill, is considered a unique prescribed medicine event. Therefore, for the purposes of editing, imputation and analysis, all records in the HC were "unfolded" to create separate records for each original purchase and each refill. Please note: MEPS did not collect information in the HC to distinguish multiple acquisitions of the same drug between the original purchase and refills. The survey only collected data on the number of times a prescribed medicine was acquired during a round. In some cases, all purchases may have been refills of an original purchase in a prior round or prior to the survey year. The file also includes a variable, (SAMPLE), which indicates whether or not the household reported receiving a free sample of that drug in that round. (To obtain more details on free samples, please see section 2.7.2.5.)

Each record on this file includes the following: an identifier for each unique prescribed medicine; detailed characteristics associated with the event (e.g., national drug code (NDC), medicine name, selected Multum Lexicon variables [see section 2.7.3 for more information on the Multum Lexicon variables included on this file], etc.); conditions, if any, associated with the medicine; the date on which the person first used the medicine; total expenditure and sources of payments; types of pharmacies that filled the household's prescriptions; whether the prescription is one in which the household received a free sample of it during the round; and a full-year person-level weight.

Data from this file can be merged with previously released MEPS-HC person-level data using the unique person identifier, DUPERSID, to append person characteristics such as demographic or health insurance coverage to each record. Data from this file can also be merged with the 2007 Full Year Consolidated Data File to estimate expenditures for persons with prescribed medicines. The Prescribed Medicines event file can also be linked to the MEPS 2007 Medical Conditions

File and additional MEPS 2007 event files. Please see the 2007 Appendix File for details on how to link MEPS data files.

2.1 Using MEPS Data for Trend and Longitudinal Analysis

MEPS began in 1996 and several annual data files have been released. As more years of data are produced, MEPS will become increasingly valuable for examining health care trends. However, it is important to consider a variety of factors when examining trends over time using MEPS. Statistical significance tests should be conducted to assess the likelihood that observed trends are attributable to sampling variation. MEPS expenditures estimates are especially sensitive to sampling variation due to the underlying skewed distribution of expenditures. For example, 1 percent of the population accounts for about one-quarter of all expenditures. The extent to which observations with extremely high expenditures are captured in the MEPS sample varies from year to year (especially for smaller population subgroups), which can produce substantial shifts in estimates of means or totals that are simply an artifact of the sample(s). The length of time being analyzed should also be considered. In particular, large shifts in survey estimates over short periods of time (e.g., from one year to the next) that are statistically significant should be interpreted with caution, unless they are attributable to known factors such as changes in public policy or MEPS survey methodology. In particular, beginning with the 2007 data, the rules used to identify outlier prices for prescription medications became much less stringent than in prior years. Starting with the 2007 Prescribed Medicine file, there is less editing of prices and quantities reported by pharmacies, more variation in prices for generics, lower mean prices for generics, higher mean prices for brand name drugs, greater differences in prices between generic and brand name drugs, and a somewhat lower proportion of spending on drugs is by families, as opposed to third-party payers. Therefore users should be cautious in the types of comparisons they make about prescription drug spending before and after 2007. For other time periods or other characteristics of prescription drugs, looking at changes over longer periods of time can provide a more complete picture of underlying trends. Analysts may wish to consider using techniques to smooth or stabilize trends analyses of MEPS data such as pooling time periods for comparison (e.g., 1996-97 versus 1998-99), working with moving averages, or using modeling techniques with several consecutive years of MEPS data to test the fit of specified patterns over time. Finally, researchers should be aware of the impact of multiple comparisons on Type I error because performing numerous statistical significance tests of trends increases the likelihood of inappropriately concluding a change is statistically significant.

The records on this file can be linked to all other 2007 MEPS-HC public use data sets by the sample person identifier (DUPERSID). Panel 11 cases (PANEL=11) can be linked back to the 2006 MEPS-HC public use files by DUPERSID and the panel indicator (PANEL).

2.2 Codebook Structure

For each variable on the file, both weighted and unweighted frequencies are provided. The codebook and data file sequence list variables in the following order:

    Unique person identifiers
    Unique prescribed medicine identifiers

Other survey administration variables
Prescribed medicine characteristics variables
ICD-9 codes
Clinical Classification Software codes
Multum Lexicon variables
Expenditure variables
Weight and variance estimation variables

2.3   Reserved Codes

The following reserved code values are used:

Value		Definition
-1	INAPPLICABLE	Question was not asked due to skip pattern.
-7	REFUSED	Question was asked and respondent refused to answer question.
-8	DK	Question was asked and respondent did not know answer.
-9	NOT ASCERTAINED	Interviewer did not record the data.
-14	NOT YET TAKEN/USED	Respondent answered that the medicine has not yet been used.

Generally, values of -1, -7, -8 and -9 have not been edited on this file. However, this is not true if the pharmacist determined a prescription drug name to be a confidentiality risk. In these instances, generally, the corresponding NDC was replaced with -9, and the Multum Lexicon therapeutic class was the replacement for the drug name determined a confidentiality risk. The values of -1 and -9 can be edited by analysts by following the skip patterns in the questionnaire. The value -14 was a valid value only for the variable representing the year the respondent reported having first used the medicine (RXBEGYRX). RXBEGYRX= -14 means that when the interviewer asked the respondent the year he/she first started using the medicine, he/she responded that he/she had not yet starting using the medicine.

A copy of the Household Component questionnaire can be found on the World Wide Web at http://www.meps.ahrq.gov/mepsweb/survey_comp/survey.jsp by selecting Prescribed Medicines (PM) from the questionnaire section.

2.4   Codebook Format

The codebook describes an ASCII data set (although the data are also being provided in a SAS transport file). The following codebook items are provided for each variable:

Identifier	Description
Name	Variable name (maximum of 8 characters)
Description	Variable descriptor (maximum of 40 characters)

Format	Number of bytes
Type	Type of data: numeric (indicated by NUM) or character (indicated by CHAR)
Start	Beginning column position of variable in record
End	Ending column position of variable in record

2.5 Variable Naming

In general, variable names reflect the content of the variable, with an 8 character limitation. Generally, imputed/edited variables end with an "X."

2.5.1 General

Variables contained on this file were derived from the HC questionnaire itself, the MPC data collection instrument, the CAPI, or from the Multum Lexicon database from Cerner Multum, Inc. The source of each variable is identified in section D, entitled "Variable-Source Crosswalk." Sources for each variable are indicated in one of four ways: (1) variables which are derived from CAPI or assigned in sampling are so indicated; (2) variables which come from one or more specific questions have those numbers and the questionnaire section indicated in the "Source" column; (3) variables constructed from multiple questions using complex algorithms are labeled "Constructed" in the "Source" column; (4) variables which have been imputed are so indicated; and (5) variables derived from the Multum Lexicon database from Cerner Multum, Inc. are so indicated.

2.5.2 Expenditure and Source of Payment Variables

Only imputed/edited versions of the expenditure variables are provided on the file. Expenditure variables on this event file follow a standard naming convention and are 7 characters in length. The 12 source of payment variables and one sum of payments variable are named consistently in the following way:

The first two characters indicate the type of event:

IP - inpatient stay	OB - office-based visit
ER - emergency room visit	OP - outpatient visit
HH - home health visit	DV - dental visit
OM - other medical equipment	RX - prescribed medicine

In the case of the source of payment variables, the third and fourth characters indicate:

SF - self or family	OF - other Federal Government	XP - sum of payments
MR - Medicare	SL - State/local government	
MD - Medicaid	WC - Worker's Compensation	
PV - private insurance	OT - other insurance	

VA - Veterans   OR - other private
TR - TRICARE   OU - other public

The fifth and sixth characters indicate the year (07). All imputed/edited expenditure variables end with an "X."

For example, RXSF07X is the edited/imputed amount paid by self or family for the 2007 prescribed medicine expenditure.

2.6   Data Collection

Data regarding prescription drugs were obtained through the HC questionnaire and a pharmacy follow-back component (within the Medical Provider Component).

2.6.1   Methodology for Collecting Household Reported Variables

During each round of the MEPS-HC, all respondents were asked to supply the name of any prescribed medicine they or their family members purchased or otherwise obtained during that round. For each medicine in each round, the following information was collected: whether any free samples of the medicine were received; the name(s) of any health problems the medicine was prescribed for; the number of times the prescription medicine was obtained or purchased; the year, month, and day on which the person first used the medicine; and a list of the names, addresses, and types of pharmacies that filled the household's prescriptions. In the HC, respondents were asked if they send in claim forms for their prescriptions or if their pharmacy providers do this automatically for them at the point of purchase. For those that said their pharmacy providers automatically send in claims for them at the point of purchase, charge and payment information was collected in the pharmacy follow-back component (unless the purchase was an insulin or diabetic supply/equipment event that was mentioned in the household component; see section 3.0 for details). However, charge and payment information was collected for those that said they send in their own prescription claim forms, because it was thought that payments by private third-party payers for those that filed their own claim forms for prescription purchases would not be available from pharmacies. Uninsured persons were treated in the same manner as those whose pharmacies filed their prescription claims at the point of purchase. Persons who said they did not know if they sent in their own prescription claim forms were treated as those who said they did send in their own prescription claim forms.

In consultation with an industry expert, outlier values for the number of times a household reported purchasing or otherwise obtaining a prescription drug in a particular round were determined by comparing the number of days a respondent was in the round and the number times the person reported obtaining the drug in the round. For these events, a new value for the number of times a drug was purchased or otherwise obtained by a person in a round was imputed. In addition, the prescribed medicine events in which a household respondent did not know/remember the number of times a certain prescribed medicine was purchased or otherwise obtained were imputed a value for that variable.

For those rounds that spanned two years, drugs mentioned in that round were allocated between the years based on the number of times the respondent said the drug was purchased in the respective year, the year the person started taking the drug, the length of the person's round, the dates of the person's round, and the number of drugs for that person in the round. In addition, a "folded" version of the PC on an event level, as opposed to an acquisition level, was used for these types of events to assist in determining how many acquisitions of the drug should be allocated between the years.

2.6.2   Methodology for Collecting Pharmacy Reported Variables

If the respondent with the prescription gave written permission to release his or her pharmacy records, pharmacy providers identified by the household were contacted by telephone for the pharmacy follow-back component. Following an initial telephone contact, the signed permission forms and materials explaining the study were faxed (or mailed) to cooperating pharmacy providers. The materials informed the providers of all persons participating in the survey who had prescriptions filled at their place of business and requested a computerized printout of all prescriptions filled for each person. For each medication listed, the following information was requested: date filled; national drug code (NDC); medication name; strength of medicine (amount and unit); quantity (package size/amount dispensed); and payments by source.

2.7   File Contents

2.7.1   Survey Administration Variables

2.7.1.1   Person Identifier Variables (DUID, PID, DUPERSID)

The dwelling unit ID (DUID) is a 5-digit random number assigned after the case was sampled for MEPS. The 3-digit person number (PID) uniquely identifies each person within the dwelling unit. The 8-character variable DUPERSID uniquely identifies each person represented on the file and is the combination of the variables DUID and PID. For detailed information on dwelling units and families, please refer to the documentation for the 2007 Full Year Population Characteristics File.

2.7.1.2   Record Identifier Variables (RXRECIDX, LINKIDX)

The variable RXRECIDX uniquely identifies each record on the file. This 15-character variable is comprised of the following components: prescribed medicine event generated through the HC (positions 1-12) + enumeration number (positions 13-15). The prescribed medicine event generated through the HC (positions 1-12) can be used to link a prescribed medicine event to the conditions file and to other event files, via link files, and is provided on this file as the variable LINKIDX. (For more details on linking, please refer to section 5.2 and to the 2007 Appendix File.)

The following hypothetical example illustrates the structure of these ID variables. This example illustrates a person in Round 1 of the household interview who reported having purchased Amoxicillin three times. The following example shows three acquisition-level records, all having

the same RXNDC (00093310905), for one person (DUPERSID=00002026) in one round. Only one NDC is associated with a prescribed medicine event because matching was performed at an event level, as opposed to an acquisition level. (For more details on matching, please see section 3.0). The LINKIDX (000020260083) remains the same for all three records, whereas the RXRECIDX (000020260083001, 000020260083002, 000020260083003) differs for all three records.

DUPERSID	RXRECIDX	LINKIDX	RXNDC
00002026	000020260083001	000020260083	00093310905
00002026	000020260083002	000020260083	00093310905
00002026	000020260083003	000020260083	00093310905

2.7.1.3 Panel Variable (PANEL)

PANEL is a constructed variable used to specify the panel number for the person. Panel will indicate either Panel 11 or Panel 12 for each person on the file. Panel 11 is the panel that started in 2004, and Panel 12 is the panel that started in 2007.

2.7.1.4 Round Variable (PURCHRD)

The variable PURCHRD indicates the round in which the prescribed medicine was purchased and takes on the value of 1, 2, 3, 4, or 5. Rounds 3, 4, and 5 are associated with MEPS survey data collection from Panel 11. Similarly, Rounds 1, 2, and 3 are associated with data collected from Panel 12.

2.7.2 Characteristics of Prescribed Medicine Events

2.7.2.1 Date When Prescribed Medicine Was First Taken (RXBEGDD-RXBEGYRX)

There are three variables which indicate when a prescribed medicine was first taken (used), as reported by the household. They are the following: RXBEGDD indicates the day a person first started taking a medicine, RXBEGMM denotes the month in which a person first started taking a medication, and RXBEGYRX reflects the year in which a person first started taking a medicine. These "first taken" questions are only asked the first time a prescription is mentioned by the household. These questions are not asked of refills of the prescription for a person in subsequent rounds and result in a value of -1 being assigned to those types of events for these variables. For purposes of confidentiality, RXBEGYRX was bottom-coded at 1922 which makes RXBEGYRX consistent with the top-coding of the age variables on the 2007 Full Year Population Characteristics Public Use File (HC-107).

2.7.2.2 Prescribed Medicine Attributes (RXNAME-RXSTRUNT)

For each prescribed medicine included on this file, several data items collected describe in detail the medication obtained or purchased. These data items are the following:

a. Medication name - pharmacy reported (RXNAME)
b. National drug code (RXNDC)
c. Quantity of the prescribed medicine dispensed (RXQUANTY); e.g., number of tablets in the prescription
d. Form of the prescribed medicine (RXFORM); e.g., powder
e. Unit of measurement for form of Rx/prescribed medicine (RXFRMUNT); e.g., oz
f. Strength of the dose of the prescribed medicine (RXSTRENG); e.g., 10
g. Unit of measurement for the strength of the dose of the prescribed medicine (RXSTRUNT); e.g., gm

Please refer to Attachments 1, 2, and 3 for definitions for RXFORM, RXFRMUNT, and RXSTRUNT abbreviations, codes and symbols.

The national drug code (NDC) generally, and on this file, is an 11-digit code. The first 5 digits indicate the manufacturer of the prescribed medicine. The next 4 digits indicate the form and strength of the prescription, and the last 2 digits indicate the package size from which the prescription was dispensed. NDC values were imputed from a proprietary database to certain PC prescriptions because the NDC reported by the pharmacy provider was not valid. These records are identified by RXFLG=3.

For the years 1996-2004, AHRQ's licensing agreement with the proprietary database precluded the release of the imputed NDC values to the public, so for these prescriptions, the household reported name of the prescription (RXHHNAME) and the original NDC (RXNDC) and prescription name (RXNAME) reported by the pharmacy were provided on the file to allow users to do their own imputation. In addition, for the years 1996-2004, the imputed NDC values for the RXFLG=3 cases could be accessed through the MEPS Data Center. For those events not falling in the RXFLG=3 category, the reserve code (-13) was assigned to the household reported medication name (RXHHNAME). The household reported name of the prescription (RXHHNAME) is no longer provided on this file; however, this variable may be accessed through the MEPS Data Center as can the original pharmacy reported name and NDC. For information on accessing data through the MEPS Data Center, see the Data Center section of the MEPS Web site at: http://www.meps.ahrq.gov/mepsweb/data_stats/onsite_datacenter.jsp.

Imputed data on this event file, unlike other MEPS event files, may still have missing data. This is because imputed data on this file are imputed from the PC or from a proprietary database. These sources did not always include complete information for each variable but did include an NDC, which would typically enable an analyst to obtain any missing data items. For example, although there are a substantial number of missing values for the strength of the prescription that were not supplied by the pharmacist, these missing values were not imputed because this information is embedded in the NDC.

2.7.2.3 Type of Pharmacy (PHARTP1-PHARTP12)

Household respondents were asked to list the type of pharmacy from which their medications were purchased. A household could list multiple pharmacies associated with their prescriptions

in a given round or over the course of all rounds combined covering the survey year. All household reported pharmacies are provided on this file, but there was no link in the survey or in the data file enabling users to know the type of pharmacy from which a specific prescription was obtained if multiple pharmacies are listed. The set of variables (PHARTP1-PHARTP12) identify the types of pharmacy providers from which the person's prescribed medicines were purchased. The possible types of pharmacies include the following: (1) mail-order, (2) another store, (3) HMO/clinic/hospital, (4) drug store, and (5) on-line. A -1 value for PHARTPn indicates that the household did not report an "n^{th}" pharmacy.

2.7.2.4 Analytic Flag Variables (RXFLG-INPCFLG)

There are four flag variables included on this file (RXFLG, PCIMPFLG, CLMOMFLG, and INPCFLG).

The variable RXFLG indicates how the NDC for a specific prescribed medicine event was imputed. This variable indicates whether or not there was any imputation performed on this record for the NDC variable, and if imputed, from what source the NDC was imputed. If no imputation was performed, RXFLG=1. If the imputation source was another PC record, RXFLG=2. Similarly, if the imputation source was a secondary, proprietary database and not the PC database, RXFLG=3.

PCIMPFLG indicates the type of match between a household reported event and a PC reported event. There are only two possible values for this variable (PCIMPFLG =1 or =2). These values indicate the possible "match-types" and are the following: =1 is an exact match for a specific event for a person between the PC and the HC and =2 is not an exact match between the PC and HC for a specific person (not an exact match means that a person's household reported event did not have a matched counterpart in their corresponding PC records). PCIMPFLG assists analysts in determining which records have the strongest link to data reported by a pharmacy. It should be noted that whenever there are multiple purchases of a unique prescribed medication in a given round, MEPS did not collect information that would enable designating any single purchase as the "original" purchase at the time the prescription was first filled, and then designating other purchases as "refills." The user needs to keep this in mind when the purchases of a medication are referred to as "refills" in the documentation. Because matching was performed at an event level as opposed to an acquisition level, the values for PCIMPFLG are either =1 or =2. Additionally, matching on an event- versus acquisition-level results in only one NDC being associated with a prescribed medicine event. (For more details on general data editing/imputation methodology, please see section 3.0).

CLMOMFLG indicates if a prescription medicine event went through the charge and payment section of the HC. Prescription medicine events that went through the charge and payment section of the HC include: (1) events where the person filed their own prescription claim forms with their insurance company, (2) events for persons who responded they did not know if they filed their own prescription claim forms with their insurance company, and (3) insulin and diabetic supply/equipment events (OMTYPE=2 or =3) that were mentioned in the Other Medical section of the HC. For these types of events information on payment sources was retained to the

extent that these data were reported by the household in the charge and payment section of the HC.

INPCFLG denotes whether or not a household respondent had at least one prescription drug purchase in the PC (0=no, 1=yes).

2.7.2.5   The Sample Variable (SAMPLE)

SAMPLE indicates if a respondent reported receiving a free sample of the prescription medicine in the round (0=no, 1=yes). Each household respondent was asked in each round whether or not they received any free samples of a reported prescribed medicine during the round. However, respondents were not asked to report the number of free samples received, nor was it made clear that free samples were included in the count of the number of times that the respondent reported purchasing or otherwise obtaining the prescribed medicine during the round. It is important for analysts to note, SAMPLE is *not* a count variable of free samples; SAMPLE =1 for all acquisitions of a prescribed medicine that a respondent reported getting a free sample of during the round. This flag variable simply allows individual analysts to determine for themselves how free samples should be handled in their analysis.

2.7.2.6   Condition Codes (RXICD1X-RXICD3X) and Clinical Classification Codes (RXCCC1X-RXCCC3X)

Information on household reported medical conditions associated with each prescribed medicine event are provided on this file. There are up to three condition and clinical classification codes listed for each prescribed medicine event (99.7 percent of prescribed medicine events have 0-3 condition records linked). To obtain complete information associated with an event, the analyst must link to the 2007 Medical Conditions File. Details on how to link to the MEPS 2007 Medical Conditions File are provided in the 2007 Appendix File. The user should note that due to confidentiality restrictions, provider reported condition information (for non-prescription medicines events) is not publicly available. Provider reported condition data (again, for non-prescription medicines events) can be accessed through the MEPS Data Center only.

The medical conditions reported by the HC respondent were recorded by the interviewer as verbatim text, which were then coded to fully-specified 2007 ICD-9-CM codes, including medical condition, V codes, and a small number of E codes, by professional coders. Although codes were verified and error rates did not exceed 2.5 percent for any coder, analysts should not presume this level of precision in the data; the ability of household respondents to report condition data that can be coded accurately should not be assumed. For detailed information on conditions, please refer to the documentation on the 2007 Medical Conditions File. For frequencies of conditions by event type, please see the 2007 Appendix File, HC-110I.

The ICD-9-CM condition codes were aggregated into clinically meaningful categories. These categories, included on the file as RXCCC1X-RXCCC3X, were generated using Clinical Classification Software (CCS) (formerly known as Clinical Classifications for Health Care

Policy Research (CCHPR), which aggregates conditions and V-codes into 263 mutually exclusive categories, most of which are clinically homogeneous.

In order to preserve respondent confidentiality, nearly all of the condition codes provided on this file have been collapsed from fully-specified codes to 3-digit code categories. The reported ICD-9-CM code values were mapped to the appropriate clinical classification category prior to being collapsed to the 3-digit categories. Because of this collapsing, it is possible for there to be duplicate ICD-9-CM condition codes linked to a single prescribed medicine event when different fully-specified codes are collapsed into the same code. This would result in two or more of the code variables on this file being set to the same value on a single record. For more information on ICD-9-CM codes, see the HC-112 documentation.

The condition codes (and clinical classification codes) linked to each prescribed medicine event are sequenced in the order in which the conditions were reported by the household respondent, which was in chronological order of reporting and not in order of importance or severity. Analysts who use the 2007 Medical Conditions file in conjunction with this prescribed medicines event file should note that the conditions on this file are sorted differently than they appear on the Medical Conditions file.

2.7.3   Multum Lexicon Variables from Cerner Multum, Inc.

Each record on this file contains the following Multum Lexicon variables:

> PREGCAT - pregnancy category variable - identifies the FDA pregnancy category to which a particular drug has been assigned

> TCn - therapeutic classification variable - assigns a drug to one or more therapeutic/chemical categories; can have up to three categories per drug

> TCnSn - therapeutic sub-classification variable - assigns one or more sub-categories to a more general therapeutic class category given to a drug

> TCnSn_n - therapeutic sub sub-classification variable - assigns one or more sub sub-categories to a more general therapeutic class category and sub-category given to a drug

For additional information on these and other Multum Lexicon variables, as well as the Multum Lexicon database itself, please refer to the following Web site: multum.com/Lexicon.htm

Researchers using the Multum Lexicon variables are requested to cite Multum Lexicon as the data source.

2.7.4   Expenditure Variables (RXSF07X-RXXP07X)

2.7.4.1   Definition of Expenditures

Expenditures on this file refer to what is paid for health care services. More specifically, expenditures in MEPS are defined as the sum of payments for care received, including out of pocket payments and payments made by private insurance, Medicaid, Medicare and other sources. The definition of expenditures used in MEPS differs slightly from its predecessors, the 1987 NMES and 1977 NMCES surveys, where "charges" rather than "sum of payments" were used to measure expenditures. This change was adopted because charges became a less appropriate proxy for medical expenditures during the 1990s due to the increasingly common practice of discounting charges. Although measuring expenditures as the sum of payments incorporates discounts in the MEPS expenditure estimates, the estimates do not incorporate any manufacturer or other rebates associated with Medicaid or other purchases. Another general change from the two prior surveys is that charges associated with uncollected liability, bad debt, and charitable care (unless provided by a public clinic or hospital) are not counted as expenditures, because there are no payments associated with those classifications. For details on expenditure definitions, please reference the following, "Informing American Health Care Policy" (Monheit, Wilson, Arnett, 1999).

If examining trends in MEPS expenditures or performing longitudinal analysis on MEPS expenditures, please refer to section C, sub-section 2.1 for more information.

2.7.4.2 Sources of Payment

In addition to total expenditures, variables are provided which itemize expenditures according to major source of payment categories. These categories are:

1. Out of pocket by user or family
2. Medicare
3. Medicaid
4. Private Insurance
5. Veteran's Administration
6. TRICARE
7. Other Federal sources - includes Indian Health Service, Military Treatment Facilities, and other care by the Federal government
8. Other State and Local Source - includes community and neighborhood clinics, State and local health departments, and State programs other than Medicaid
9. Worker's Compensation
10. Other Unclassified Sources - includes sources such as automobile, homeowner's, liability, and other miscellaneous or unknown sources

Two additional source of payment variables were created to classify payments for particular persons that appear inconsistent due to differences between survey questions on health insurance coverage and sources of payment for medical events. These variables include:

11. Other Private - any type of private insurance payments reported for persons not reported to have any private health insurance coverage during the year as defined in MEPS; and

12. Other Public - Medicaid payments reported for persons who were not reported to be enrolled in the Medicaid program at any time during the year

Though relatively small in magnitude, users should exercise caution when interpreting the expenditures associated with these two additional sources of payment. While these payments stem from apparent inconsistent responses to health insurance and source of payment questions in the survey, some of these inconsistencies may have logical explanations. For example, private insurance coverage in MEPS is defined as having a major medical plan covering hospital and physician services. If a MEPS sampled person did not have such coverage but had a single service type insurance plan (e.g., dental insurance) that paid for a particular episode of care, those payments may be classified as "other private." Some of the "other public" payments may stem from confusion between Medicaid and other state and local programs or may be from persons who were not enrolled in Medicaid, but were presumed eligible by a provider who ultimately received payments from the program.

2.7.5   Sample Weight (PERWT07F)

2.7.5.1   Overview

There is a single full year person-level weight (PERWT07F) assigned to each record for each key, in-scope person who responded to MEPS for the full period of time that he or she was in-scope during 2007. A key person either was a member of an NHIS household at the time of the NHIS interview, or became a member of such a household after being out-of-scope at the time of the NHIS (examples of the latter situation include newborns and persons returning from military service, an institution, or living outside the United States). A person is in-scope whenever he or she is a member of the civilian noninstitutionalized portion of the U.S. population.

2.7.5.2   Details on Person Weights Construction

The person-level weight PERWT07F was developed in several stages. Person-level weights for Panels 11 and 12 were created separately. The weighting process for each panel included an adjustment for nonresponse over time and poststratification. Poststratification was achieved by controlling to Current Population Survey (CPS) population estimates based on five variables. Variables used in the establishment of person-level poststratification control figures included: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, and other); sex; and age. A 2007 composite weight was then formed by multiplying each panel weight by 0.5 and then poststratifying the resulting weight to the same set of CPS-based control totals. When poverty status information derived from income variables became available, a final poststratification was done on the resulting weight variable, including poverty status (below poverty, from 100 to 125 percent of poverty, from 125 to 200 percent of poverty, from 200 to 400 percent of poverty, at least 400 percent of poverty) as well as the original five poststratification variables in the establishment of control totals.

2.7.5.3   MEPS Panel 11 Weight

The person-level weight for MEPS Panel 11 was developed using the 2006 full year weight for

an individual as a "base" weight for survey participants present in 2006. For key, in-scope respondents who joined an RU some time in 2007 after being out-of-scope in 2006, the 2006 family weight associated with the family the person joined served as a "base" weight. The weighting process included an adjustment for nonresponse over Rounds 4 and 5 as well as poststratification to population control figures for December 2007. These control figures were derived by scaling back the population totals obtained from the March 2007 CPS to reflect the December 2007 CPS estimated population distribution across age and sex categories as of December 2007. Variables used in the establishment of person-level poststratification control figures included: census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, and other); sex; and age. Overall, the weighted population estimate for the civilian noninstitutionalized population on December 31, 2007 is 301,309,149. Key, responding persons not in-scope on December 31, 2007 but in-scope earlier in the year retained, as their final Panel 11 weight, the weight after the nonresponse adjustment.

2.7.5.4 MEPS Panel 12 Weight

The person-level weight for MEPS Panel 12 was developed using the MEPS Round 1 person-level weight as a "base" weight. For key, in-scope respondents who joined an RU after Round 1, the Round 1 family weight served as a "base" weight. The weighting process included an adjustment for nonresponse over Round 2 and the 2007 portion of Round 3 as well as poststratification to the same population control figures for December 2007 used for the MEPS Panel 11 weights. The same five variables employed for Panel 11 poststratification (census region, MSA status, race/ethnicity, sex, and age) were used for Panel 12 poststratification. Similarly, for Panel 12, key, responding persons not in-scope on December 31, 2007 but in-scope earlier in the year retained, as their final Panel 12 weight, the weight after the nonresponse adjustment.

Note that the MEPS Round 1 weights (for both panels with one exception as noted below) incorporated the following components: the original household probability of selection for the NHIS; ratio-adjustment to NHIS-based national population estimates at the household (occupied dwelling unit) level; adjustment for nonresponse at the dwelling unit level for Round 1; and poststratification to figures at the family and person level obtained from the March 2007 CPS data base.

2.7.5.5 The Final Weight for 2007

Variables used in the establishment of person-level poststratification control figures included: poverty status (below poverty, from 100 to 125 percent of poverty, from 125 to 200 percent of poverty, from 200 to 400 percent of poverty, at least 400 percent of poverty); census region (Northeast, Midwest, South, West); MSA status (MSA, non-MSA); race/ethnicity (Hispanic, black but non-Hispanic, and other); sex; and age. Overall, the weighted population estimate for the civilian noninstitutionalized population for December 31, 2007 is 301,309,149 (PERWT07F>0 and INSC1231=1). The weights of some persons out-of-scope on December 31, 2007 were also poststratified. Specifically, the weights of persons out-of-scope on December 31, 2007 who were in-scope some time during the year and also entered a nursing home during the

year were poststratified to a corresponding control total obtained from the 1996 MEPS Nursing Home Component. The weights of persons who died while in-scope during 2007 were poststratified to corresponding estimates derived using data obtained from the Medicare Current Beneficiary Survey (MCBS) and Vital Statistics information provided by the National Center for Health Statistics (NCHS). Separate control totals were developed for the "65 and older" and "under 65" civilian noninstitutionalized populations.

2.7.5.6 Coverage

The target population for MEPS in this file is the 2007 U.S. civilian noninstitutionalized population. However, the MEPS sampled households are a subsample of the NHIS households interviewed in 2005 (Panel 11) and 2006 (Panel 12). New households created after the NHIS interviews for the respective Panels and consisting exclusively of persons who entered the target population after 2005 (Panel 11) or after 2006 (Panel 12) are not covered by MEPS. Neither are previously out-of-scope persons who join an existing household but are unrelated to the current household residents. Persons not covered by a given MEPS panel thus include some members of the following groups: immigrants; persons leaving the military; U.S. citizens returning from residence in another country; and persons leaving institutions. The set of uncovered persons constitutes only a small segment of the MEPS target population.

3.0 General Data Editing and Imputation Methodology

The general approach to preparing the household prescription data for this file was to utilize the PC prescription data to impute information collected from pharmacy providers to the household drug mentions. For events that went through the charge and payment section of the HC (events where the person filed their own prescription claim forms with their insurance company, events for persons who responded they did not know if they filed their own prescription claim forms with their insurance company, and insulin and diabetic supply/equipment events (OMTYPE=2 or =3) that were mentioned in the Other Medical section of the HC), information on payment sources was retained to the extent that these data were reported by the household in the charge and payment section of the HC. A matching program was adopted to link PC drugs and the corresponding drug information to household drug mentions. To improve the quality of these matches, all drugs on the household and pharmacy files were coded using a proprietary database on the basis of the medication names provided by the household and pharmacy, and, when available, the NDC provided in the pharmacy follow-back component. The matching process was done at an event level, as opposed to an acquisition level. Considerable editing was done prior to the matching to correct data inconsistencies in both data sets and to fill in missing data and correct outliers on the pharmacy file.

Drug price-per-unit outliers were analyzed on the pharmacy file by first identifying the average wholesale unit price (AWUP) of the drug by linkage through the NDC to a secondary data file. In general, prescription drug unit prices were deemed to be outliers by comparing unit prices reported in the pharmacy database to the AWUP reported in the secondary data file and were edited, as necessary.

Beginning with the 2007 data, the rules used to identify outlier prices for prescription medications in the PC changed. New outlier thresholds were established based on the distribution of the ratio of retail unit prices relative to the AWUP in the 2006 MarketScan Outpatient Pharmaceutical Claims data base. The new thresholds vary by patent status, whereas in prior years they did not. These changes improve data quality in three ways: (1) the distribution of prices in the 2007 MEPS better benchmarks to MarketScan, overall and by patent status, (2) fewer pharmacy-reported payments and quantities (for example, number of pills) are edited, and (3) imputed prices reflect prices paid, rather than AWUPs. As a result, compared with earlier years of the MEPS, starting with 2007 there is more variation in prices for generics, lower mean prices for generics, higher mean prices for brand name drugs, greater differences in prices between generic and brand name drugs, and a somewhat lower proportion of spending on drugs is by families, as opposed to third-party payers.

Drug matches between household drug mentions and pharmacy drug events for a person in the PC were based on drug code, medication name, and the round in which the drug was reported. The matching of household drug mentions to pharmacy drugs was performed so that the most detailed and accurate information for each prescribed medicine event was obtained. Exact dates of purchase were only available from the follow-back component. The matching program assigned scores to potential matches. Numeric variables required exact matches to receive a high score, while partial scores could be assigned to matches between character variables, such as prescription name, depending on the degree of similarity in the spelling and sound of the medication names. Household drug mentions that were deemed exact matches to PC drugs for the same person in the same round required sufficiently high scores to reflect a high quality match. Exact matches were used only once and were taken out of the donor pool from that point on (i.e., these matches were made without replacement). Any refill of a household drug mention that had been matched to a pharmacy drug event was also matched to the same pharmacy drug event. All remaining unmatched household drug mentions for persons either in or out of the PC were statistically matched to the entire pharmacy donor base with replacement by medication name, drug code, type of third party coverage, health conditions, age, sex, and other characteristics of the individual. Potential PC donor records were omitted from these matches whenever a NDC was imputed to the PC record and was not an exact match on a generic product code applied to all records in the HC and PC.

For more information on the MEPS Prescribed Medicines editing and imputation procedures, please see J. Moeller, 2001.

3.1   Rounding

Expenditure variables on the 2007 Prescribed Medicines file have been rounded to the nearest penny. Person-level expenditure variables released on the 2007 Full Year Consolidated Data File were rounded to the nearest dollar. It should be noted that using the 2007 MEPS event files to create person-level totals will yield slightly different totals than those found on the 2007 Full Year Consolidated Data File. These differences are due to rounding only. Moreover, in some instances, the number of persons having expenditures on the 2007 event files for a particular source of payment may differ from the number of persons with expenditures on the 2007 Full

Year Consolidated Data File for that source of payment. This difference is also an artifact of rounding only. Please see the 2007 Appendix File for details on such rounding differences.

3.2   Edited/Imputed Expenditure Variables (RXSF07X-RXXP07X)

There are 13 expenditure variables included on this event file. All of these expenditures have gone through an editing and imputation process and have been rounded to the second decimal place. There is a sum of payments variable (RXXP07X) which for each prescribed medicine event sums all the expenditures from the various sources of payment. The 12 sources of payment expenditure variables for each prescribed medicine event are the following: amount paid by self or family (RXSF07X), amount paid by Medicare (RXMR07X), amount paid by Medicaid (RXMD07X), amount paid by private insurance (RXPV07X), amount paid by the Veterans Administration (RXVA07X), amount paid by TRICARE (RXTR07X), amount paid by other federal sources (RXOF07X), amount paid by state and local (non-federal) government sources (RXSL07X), amount paid by Worker's Compensation (RXWC07X), and amount paid by some other source of insurance (RXOT07X). As mentioned previously, there are two additional expenditure variables called RXOR07X and RXOU07X (other private and other public, respectively). These two expenditure variables were created to maintain consistency between what the household reported as their private and public insurance status for hospitalization and physician coverage and third party prescription payments from other private and public sources (such as a separate private prescription policy or prescription coverage from the Veterans Administration, the Indian Health Service, or a State assistance program other than Medicaid). Users should exercise caution when interpreting the expenditures associated with these two additional sources of payment. While these payments stem from apparent inconsistent responses to health insurance and source of payment questions in the survey, some of these inconsistencies may have logical explanations. Please see section 2.7.4 for details on these and all other source of payment variables.

4.0   Strategies for Estimation

4.1   Developing Event-Level Estimates

The data in this file can be used to develop national 2007 event level estimates for the U.S. civilian noninstitutionalized population on prescribed medicine purchases (events) as well as expenditures, and sources of payment for these purchases. Estimates of total purchases are the sum of the weight variable (PERWT07F) across relevant event records while estimates of other variables must be weighted by PERWT07F to be nationally representative. The tables below contain event-level estimates for selected variables.

Selected Event (Purchase) Level Estimates

All Prescribed Medicine Purchases

Estimate of Interest	Variable Name	Estimate (SE)
Number of purchases (in millions)	PERWT07F	3070.8 (68.45)
Mean total payments per purchase	RXXP07X	$76 (1.2)

Estimate of Interest	Variable Name	Estimate (SE)
Mean out-of-pocket payment per purchase	RXSF07X	$21 (0.3)
Mean proportion of expenditures paid by private insurance per purchase	RXPV07X /RXXP07X	0.208 (0.0045)

Example by Drug Type: Statins (TC1S1_1=173 or TC1S1_2=173 or TC1S2_1=173 or TC1S3_1 =173 or TC2S1_1=173 or TC2S1_2=173)

Estimate of Interest	Variable Name	Estimate (SE)
Number of purchases (in millions)	PERWT07F	185.9 (6.07)
Mean total payments per purchase	RXXP07X	$100 (1.9)
Mean annual total payments per person	RXXP07X (aggregated across purchases within person)	$594 (12.3)

Example by Associated Condition: Hypertension (RXICD1X="401" or RXICD2X="401" or RXICD3X="401")

Estimate of Interest	Variable Name	Estimate (SE)
Number of purchases (in millions)	PERWT07F	472.1 (14.88)
Mean total payments per purchase	RXXP07X	$43 (0.8)
Mean annual total payments per person	RXXP07X (aggregated across purchases within person)	$422 (9.6)

4.2   Person-Based Estimates for Prescribed Medicine Purchases

To enhance analyses of prescribed medicine purchases, analysts may link information about prescribed medicine purchases by sample persons in this file to the annual full year consolidated file (which has data for all MEPS sample persons), or conversely, link person-level information from the full year consolidated file to this event level file (see section 5 below for more details). Both this file and the full year consolidated file may be used to derive estimates for persons with prescribed medicine purchases and annual estimates of total expenditures for these purchases. However, if the estimate relates to the entire population, this file cannot be used to calculate the denominator, as only those persons with at least one prescribed medicine purchase are

represented on this data file. Therefore, the full year consolidated file must be used for person-level analyses that include both persons with and without prescribed medicine events.

4.3   Variables with Missing Values

It is essential that the analyst examine all variables for the presence of negative values used to represent missing values. For continuous or discrete variables, where means or totals may be taken, it may be necessary to set minus values to values appropriate to the analytic needs. That is, the analyst should either impute a value or set the value to one that will be interpreted as missing by the computing language used. For categorical and dichotomous variables, the analyst may want to consider whether to recode or impute a value for cases with negative values or whether to exclude or include such cases in the numerator and/or denominator when calculating proportions.

Methodologies used for the editing/imputation of expenditure variables (e.g., total expenditures and sources of payment) are described in Section 3.2.

4.4   Variance Estimation (VARSTR, VARPSU)

MEPS has a complex sample design. To obtain estimates of variability (such as the standard error of sample estimates or corresponding confidence intervals) for MEPS estimates, analysts need to take into account the complex sample design of MEPS for both person-level and family-level analyses. Several methodologies have been developed for estimating standard errors for surveys with a complex sample design, including the Taylor-series linearization method, balanced repeated replication, and jackknife replication. Various software packages provide analysts with the capability of implementing these methodologies. Replicate weights have not been developed for the MEPS data. Instead, the variables needed to calculate appropriate standard errors based on the Taylor-series linearization method are included on this file as well as all other MEPS public use files. Software packages that permit the use of the Taylor-series linearization method include SUDAAN, Stata, SAS (version 8.2 and higher), and SPSS (version 12.0 and higher). For complete information on the capabilities of each package, analysts should refer to the corresponding software user documentation.

Using the Taylor-series linearization method, variance estimation strata and the variance estimation PSUs within these strata must be specified. The variance strata variable is named VARSTR, while the variance PSU variable is named VARPSU. Specifying a "with replacement" design in a computer software package, such as SUDAAN, provides standard errors appropriate for assessing the variability of MEPS survey estimates. It should be noted that the number of degrees of freedom associated with estimates of variability indicated by such a package may not appropriately reflect the actual number available. For MEPS sample estimates for characteristics generally distributed throughout the country (and thus the sample PSUs), one can expect at least 100 degrees of freedom for the 2007 full year data associated with the corresponding estimates of variance.

Prior to 2002, MEPS variance strata and PSUs were developed independently from year to year, and the last two characters of the strata and PSU variable names denoted the year. However, beginning with the 2002 Point-in-Time PUF, the variance strata and PSUs were developed to be

compatible with MEPS data associated with the NHIS sample design used through 2007. Such data can be pooled and the variance strata and PSU variables provided can be used without modification for variance estimation purposes for estimates covering multiple years of data. There are 368 variance estimation strata, each stratum with either two or three variance estimation PSUs.

Note: A new NHIS sample design is being implemented beginning in 2006. As a result, the MEPS variance estimation structure was modified for MEPS data collected in 2007 and beyond.

5.0   Merging/Linking MEPS Data Files

Data from this file can be used alone or in conjunction with other files for different analytic purposes. This section summarizes various scenarios for merging/linking MEPS event files. Each MEPS panel can also be linked back to the previous years National Health Interview Survey public use data files. For information on obtaining MEPS/NHIS link files please see www.meps.ahrq.gov/data_stats/more_info_download_data_files.jsp.

5.1   Linking to the Person-Level File

Merging characteristics of interest from the person-level file (e.g., MEPS 2007 Full Year Consolidated File) expands the scope of potential estimates. For example, to estimate the total number of prescribed medicine purchases of persons with specific demographic characteristics (such as age, race, sex, and education), population characteristics from a person-level file need to be merged onto the prescribed medicines file. This procedure is illustrated below. The MEPS 2007 Appendix File, HC-110I, provides additional detail on how to merge MEPS data files.

> 1) Create data set PERSX by sorting the 2007 Full Year Consolidated File by the person identifier, DUPERSID. Keep only variables to be merged onto the prescribed medicines file and DUPERSID.
>
> 2) Create data set PMEDS by sorting the 2007 Prescribed Medicines File by person identifier, DUPERSID.
>
> 3) Create final data set NEWPMEDS by merging these two files by DUPERSID, keeping only records on the prescribed medicines file.

The following is an example of SAS code, which completes these steps:

```
PROC SORT DATA=IN.HC112 (KEEP=DUPERSID AGE31X AGE42X AGE53X SEX
RACEX EDUCYR)
OUT=PERSX;
  BY DUPERSID;
RUN;

PROC SORT DATA=IN.HC110A
OUT=PMEDS;
```

```
    BY DUPERSID;
  RUN;

  DATA NEWPMEDS;
  MERGE PMEDS (IN=A) PERSX (IN=B);
    BY DUPERSID;
        IF A;
  RUN;
```

5.2   Linking to the Medical Conditions File

The CLNK provides a link from MEPS event files to the 2007 Medical Conditions File. When using the CLNK, data users/analysts should keep in mind that (1) conditions are self-reported, (2) there may be multiple conditions associated with a prescribed medicine purchase, and (3) a condition may link to more than one prescribed medicine purchase or any other type of purchase. Users should also note that not all prescribed medicine purchases link to the condition file.

5.3   Pooling Annual Files

To facilitate analysis of subpopulations and/or low prevalence events, it may be desirable to pool together more than one year of data to yield sample sizes large enough to generate reliable estimates. For more details on pooling MEPS data files see www.meps.ahrq.gov/data_stats/download_data_files_detail.jsp?cboPufNumber=HC-036.

Starting in Panel 9, values for DUPERSID from previous panels will occasionally be re-used. Therefore, it is necessary to use the panel variable (PANEL) in combination with DUPERSID to ensure unique person-level identifiers across panels. Creating unique records in this manner is advised when pooling MEPS data across multiple annual files that have one or more identical values for DUPERSID.

5.4   Longitudinal Analysis

Panel-specific files containing estimation variables to facilitate longitudinal analysis are available for downloading in the data section of the MEPS Web site.

References

Cohen, S.B. (1998). Sample Design of the 1996 Medical Expenditure Panel Survey Medical Provider Component. *Journal of Economic and Social Measurement.* Vol 24, 25-53.

Cohen, S.B. (1997). *Sample Design of the 1996 Medical Expenditure Panel Survey Household Component.* Rockville (MD): Agency for Health Care Policy and Research; 1997. MEPS Methodology Report, No. 2. AHCPR Pub. No. 97-0027.

Cohen, J.W. (1997). *Design and Methods of the Medical Expenditure Panel Survey Household Component.* Rockville (MD): Agency for Health Care Policy and Research; 1997. MEPS Methodology Report, No. 1. AHCPR Pub. No. 97-0026.

Cohen, S.B. (1996). The Redesign of the Medical Expenditure Panel Survey: A Component of the DHHS Survey Integration Plan. *Proceedings of the COPAFS Seminar on Statistical Methodology in the Public Service.*

Cox, B.G. and Cohen, S.B. (1985). Chapter 8: Imputation Procedures to Compensate for Missing Responses to Data Items. In *Methodological Issues for Health Care Surveys.* Marcel Dekker, New York.

Moeller J.F., Stagnitti, M., Horan, E., et al. *Outpatient Prescription Drugs: Data Collection and Editing in the 1996 Medical Expenditure Panel Survey (HC-010A).* Rockville (MD): Agency for Healthcare Research and Quality; 2001. MEPS Methodology Report No. 12. AHRQ Pub. No. 01-0002.

Monheit, A.C., Wilson, R., and Arnett, III, R.H. (Editors). *Informing American Health Care Policy.* (1999). Jossey-Bass Inc, San Francisco.

Shah, B.V., Barnwell, B.G., Bieler, G.S., Boyle, K.E., Folsom, R.E., Lavange, L., Wheeless, S.C., and Williams, R. (1996). *Technical Manual: Statistical Methods and Algorithms Used in SUDAAN Release 7.0,* Research Triangle Park, NC: Research Triangle Institute.

Zodet, M., Hill, S., and Miller, G.E. "Comparison of Retail Drug Prices in the MEPS and MarketScan: Implications for MEPS Editing Rules." Agency for Healthcare Research and Quality Working Paper, 2009.

D. Variable-Source Crosswalk

MEPS HC-110A: 2007 Prescribed Medicines Events

Survey Administration Variables

Variable	Description	Source
DUID	Dwelling unit ID	Assigned in sampling
PID	Person number	Assigned in sampling
DUPERSID	Sample person ID (DUID + PID)	Assigned in sampling
RXRECIDX	Record ID – Unique Prescribed Medicine Identifier	Constructed
LINKIDX	Link to condition and other event files	CAPI derived
PANEL	Panel indicator	Assigned in sampling
PURCHRD	Round in which the Rx/prescribed medicine was obtained/purchased	CAPI derived

Prescribed Medicines Events Variables

Variable	Description	Source
RXBEGDD	Day person first used medicine	PM11OV1
RXBEGMM	Month person first used medicine	PM11OV2
RXBEGYRX	Year person first used medicine	PM11
RXNAME	Medication name (Imputed)	Imputed
RXNDC	National drug code (Imputed)	Imputed
RXQUANTY	Quantity of Rx/prescribed medicine (Imputed)	Imputed
RXFORM	Form of Rx/prescribed medicine (Imputed)	Imputed
RXFRMUNT	Unit of measurement for form of Rx/prescribed medicine (Imputed)	Imputed
RXSTRENG	Strength of Rx/prescribed medicine dose (Imputed)	Imputed
RXSTRUNT	Unit of measurement for strength of Rx/prescribed medicine dose (Imputed)	Imputed
PHARTP1-PHARTP12	Type of pharmacy provider – (1st-12th)	PM16

Variable	Description	Source
RXFLG	Flag variable indicating imputation source for NDC on pharmacy donor record	Constructed
PCIMPFLG	Flag indicating type of household to pharmacy prescription match	Constructed
CLMOMFLG	Charge/payment, Rx claim filing, and OMTYPE =2 or =3 (insulin and diabetic supply equipment events) status	CP01/Constructed
INPCFLG	Flag indicating if the person has at least one record in the pharmacy component	Constructed
SAMPLE	Flag indicating if a respondent received a free sample of this drug in the round	CAPI derived
RXICD1X	3 digit ICD-9 condition code	PM09
RXICD2X	3 digit ICD-9 condition code	PM09
RXICD3X	3 digit ICD-9 condition code	PM09
RXCCC1X	Modified Clinical Classification Code	Constructed/Edited
RXCCC2X	Modified Clinical Classification Code	Constructed/Edited
RXCCC3X	Modified Clinical Classification Code	Constructed/Edited
PREGCAT	Multum Pregnancy Category	Cerner Multum, Inc.
TC1	Multum Therapeutic Class #1	Cerner Multum, Inc.
TC1S1	Multum Therapeutic Sub-Class #1 for TC1	Cerner Multum, Inc.
TC1S1_1	Multum Therapeutic Sub-Sub-Class for TC1S1	Cerner Multum, Inc.
TC1S1_2	Multum Therapeutic Sub-Sub-Class for TC1S1	Cerner Multum, Inc.
TC1S2	Multum Therapeutic Sub-Class #2 for TC1	Cerner Multum, Inc.
TC1S2_1	Multum Therapeutic Sub-Sub-Class for TC1S2	Cerner Multum, Inc.
TC1S3	Multum Therapeutic Sub-Class #3 for TC1	Cerner Multum, Inc.
TC1S3_1	Multum Therapeutic Sub-Sub-Class for TC1S3	Cerner Multum, Inc.
TC2	Multum Therapeutic Class #2	Cerner Multum, Inc.
TC2S1	Multum Therapeutic Sub-Class #1 for TC2	Cerner Multum, Inc.
TC2S1_1	Multum Therapeutic Sub-Sub-Class for TC2S1	Cerner Multum, Inc.
TC2S1_2	Multum Therapeutic Sub-Sub-Class for TC2S1	Cerner Multum, Inc.
TC2S2	Multum Therapeutic Sub-Class #2 for TC2	Cerner Multum, Inc.
TC3	Multum Therapeutic Class #3	Cerner Multum, Inc.
TC3S1	Multum Therapeutic Sub-Class #1 for TC3	Cerner Multum, Inc.

Variable	Description	Source
RXSF07X	Amount paid, self or family (Imputed)	CP11/Edited/ Imputed
RXMR07X	Amount paid, Medicare (Imputed)	CP12/CP13/Edited/ Imputed
RXMD07X	Amount paid, Medicaid (Imputed)	CP12/CP13/Edited/ Imputed
RXPV07X	Amount paid, private insurance (Imputed)	CP12/CP13/Edited/ Imputed
RXVA07X	Amount paid, Veteran's Administration (Imputed)	CP12/CP13/Edited/ Imputed
RXTR07X	Amount paid, TRICARE (Imputed)	CP12/CP13/Edited/ Imputed
RXOF07X	Amount paid, other Federal (Imputed)	CP12/CP13/Edited/ Imputed
RXSL07X	Amount paid, state and local government (Imputed)	CP12/CP13/Edited/ Imputed
RXWC07X	Amount paid, Worker's Compensation (Imputed)	CP12/CP13/Edited/ Imputed
RXOT07X	Amount paid, other insurance (Imputed)	CP12/CP13/Edited/ Imputed
RXOR07X	Amount paid, other private (Imputed)	Constructed/Imputed
RXOU07X	Amount paid, other public (Imputed)	Constructed/Imputed
RXXP07X	Sum of payments RXSF07X – RXOU07X (Imputed)	CP12/CP13/Edited/ Imputed

Weights

Variable	Description	Source
PERWT07F	Poverty/mortality/nursing home adjusted person-level weight	Constructed
VARSTR	Variance estimation stratum, 2007	Constructed
VARPSU	Variance estimation PSU, 2007	Constructed

Attachment 1

Definitions of Abbreviations for RXFORM

Dosage Form	Definition
-7	refused
-8	don't know
-9	not ascertained
ACC	accessory
ADR	acetic acid drop
AE	aerosol
AER	aerosol
AERO	aerosol
AEROSOL	
AMP	ampule
ARO	aerosol solid
AUTO INJ	auto-injection
BACK SUPPORT BELT	
BAG	
BAL	balm
BALM	
BAN	bandage
BANDAGE	
BAR	
BATTERY	
BENCH	
BOT	bottle
BOTTLE	
BOX	
BOXES	
BRACE	
BRIEF	
BUT	butterfly
C	capsules, or cream (varies)
C12	12 hour extended-release capsule
C24	24 hour extended-release capsule
CA	capsule
CANE	
CAP	capsule
CAP DR	delayed-release capsule
CAP ER	extended-release capsule
CAP SA	slow-acting capsule
CAPLET	
CAPLT	caplet

Dosage Form	Definition
CAPS	capsules
CAPSULE	
CAPSULE SA	slow-acting capsule
CATHETER	
CC	cubic centimeter
CER	extended-release capsule
CHAMBER	
CHEW	chewable tablet
CHEW TAB	chewable tablet
CHEW TABS	chewable tablets
CHEWABLE	
CHW	chewable tablets
CLEANSER	
COLLAR	
COMBO	
COMPOUND	
CON	condom
CONDOM	
CONTAINER	
COTTON	
CPSR	slow-release capsule
CR	cream
CRE	cream
CREA	cream
CREAM	
CRM	cream
CRYSTAL	
CTB	chewable tablets
CTG	cartridge
CUTTER	
DEV	device
DEVICE	
DIA	diaper
DIAPER	
DIAPHRAM	
DIS	disk, or dermal infusion system
DISK	
DOS PAK	dose pack
DR	drop
DRE	dressing
DRESSING	
DRO	drop
DROP	
DROPS	

Dosage Form	Definition
DROPS OPTH OTI	ophthalmic/otic drops
DROPS SUSP	drops suspension
DRP	drop
DRPS	drops
DSK	disk
DSPK	tablets in a dose pack
EAR DROP	
EAR DROPS	
EAR SUSP	ear suspension
EC TABS	enteric coated tablets
ECC	enteric coated capsules
ECT	enteric coated tablets
ELI	elixir
ELIX	elixir
ELIXIR	
ELX	elixir
EMERGENCY KIT	
EMO	emollient
EMU	emulsion
EMULSION	
ENEMA	
ERTA	extended-release tablets
EXTN CAP	extended-release capsule
EXTRACT	
EYE DRO	eye drop
EYE DROP	
EYE DROPS	
EYE SO	eye solution
FIL	film
FILM ER	film, extended-release
FILMTAB	
FILMTABS	
FLOWMETER	
FOA	foam
FOAM	
GAU	gauze
GAUZE	
GEF	effervescent granules
GEL	
GEL CAP	gel capsule
GER	granules, extended release
GFS	gel-forming solution
GLOVE	
GRA	granules

Dosage Form	Definition
GRR	grams
GTT	drops
GUM	
HOSE	medical hosiery
HU	capsule
ICR	control-release insert
IMPLANT	
IN	injectible
INH	inhalant
INH AER	inhalant aerosol
INHAL	inhalant
INHAL SOL	inhalant solution
INHALER	
INHL	inhalant
INJ	injectible
INJECTION (S)	
INSERT	
INSULIN	
IUD	intrauterine devise
IV	intravenous
JEL	jelly
JELLY	
KIT	
L	lotion
LANCET	
LANCET (S)	
LI	liquid
LINIMENT	
LIQ	liquid
LIQUID	
LOLLIPOP	
LOT	lotion
LOTION	
LOZ	lozenge
LOZENGE	
MASK	
MCG	microgram
METER	
MG	milligram
MIS	miscellaneous
MIST	
MONITOR	
MOUTHWASH	
NAS	nasal spray

Dosage Form	Definition
NASAL	
NASAL INHALER	
NASAL POCKET HL	nasal inhaler, pocket
NASAL SOLN	nasal solution
NASAL SPR	nasal spray
NASAL SPRAY	
NDL	needle
NE	nebulizer
NEB	nebulizer
NEBULIZER	
NMA	enema
NMO	nanomole, millimicromole
ODR	ophthalmic drop (ointment)
ODT	oral disintegrating tablet
OIL	
OIN	ointment
OINT	ointment
OINT TOP	topical ointment
OINTMENT	
ONT	ointment
OP	ophthalmic solution
OP DROPS	ophthalmic drops
OP SOL	ophthalmic solution
OPH	ophthalmic
OPH S	ophthalmic solution or suspension
OPH SOL	ophthalmic solution
OPH SOLN	ophthalmic solution
OPHTH DROP (S)	ophthalmic drops
OPHTH OINT	ophthalmic ointment
OPHTH SOLN	ophthalmic solution
OPT SLN	ophthalmic solution
OPT SOL	ophthalmic solution
OPTH	ophthalmic solution or suspension or ointment
OPTH S	ophthalmic solution or suspension
OPTH SLN	ophthalmic solution
OPTH SOL	ophthalmic solution
OPTH SUSP	ophthalmic suspension
OPTIC	
ORAL	
ORAL INHL	oral inhalant
ORAL INHALER	
ORAL PWD	oral powder

Dosage Form	Definition
ORAL RINSE	
ORAL SOL	oral solution
ORAL SUS	oral suspension
ORAL SUSP	oral suspension
OTI	otic solution
OTIC	
OTIC SOL	otic solution
OTIC SOLN	otic solution
OTIC SUSP	otic suspension
PA	tablet pack, pad or patch (varies)
PAC	pack
PAD	
PADS	
PAK	pack
PAS	paste
PASTE	
PAT	patch
PATCH	
PCH	patch
PDR	powder
PDS	powder for reconstitution
PEDIATRIC DROPS	
PI1	powder for injection, 1 month
PI3	powder for injection, 3 months
PIH	powder for inhalation
PKG	package
PKT	packet
PLASTER	
PLEDGETS	
PO-SYRUP	syrup by mouth (oral syrup)
POPSICLE	
POUCH	
POW	powder
POWD	powder
POWDER	
POWDER/SUSPENS	powder/suspension
PRO	prophylactic
PULVULE	
PWD	powder
PWD F/SOL	powder for solution
RCTL SUPP	rectal suppository
RECTAL CREAM	
REDITABS	
RIN	Rinse

Dosage Form	Definition
RING	
RINSE	
ROLL	
S	syrup, suspension, solution (varies)
SA CAPS	Slow-acting capsules
SA TAB	Slow-acting tablet
SA TABLETS	Slow-acting tablets
SA TABS	Slow-acting tablets
SAL	Salve
SCRUB	
SER	extended-release suspension
SET	
SGL	soft b23gel cap
SHA	shampoo
SHAM	shampoo
SHMP	shampoo
SHOE	
SLT	sublingual tablet
SL TAB	sublingual tablet
SO	solution
SOA	Soap
SOL	solution
SOLN	solution
SOLUTION	
SP	spray
SPG	sponge
SPN	
SPONGE	
SPR	spray
SPRAY	
STOCKING	
SRN	syringe
STP	Strip
STR	Strip
STRIP	
STRIPS	
SU	suspension, solution, suppository, powder, or granules for reconstitution (varies)
SUB	sublingual
SUP	suppository
SUPP	suppository
SUPPOSITORIES	
SUPPOSITORY	

Dosage Form	Definition
SUS	suspension
SUS/LIQ	suspension/liquid
SUSP	suspension
SUSPEN	suspension
SUSPENDED RELEASE CAPLET	
SUSPENSION	
SWA	Swab
SWAB	
SWABS	
SYP	syrup
SYR	syrup
SYRINGE	
SYRP	syrup
SYRUP	
T	tablet
T12	12 hour extended-release tablet
T24	24 hour extended-release tablet
TA	tablet
TAB	tablet
TAB CHEW	chewable tablet
TAB DR	delayed-release tablet
TAB EC	enteric coated tablet
TAB SL	Slow-acting tablet
TAB SUBL	sublingual tablet
TABL	tablet
TABLET	
TABLET CUTTER	
TABLET SPLITTER	
TABLETS	
TABS	tablets
TAP	Tape
TAPE	
TB	tablet
TBCH	chewable tablet
TBS	tablets
TBSL	sublingual tablet
TBSR	Slow-release tablet
TCP	tablet, coated particles
TDM	extended-release film
TDR	orally disintegrating tablets
TDS	transdermal system
TEF	effervescent tablet
TER	extended-release tablet

Dosage Form	Definition
TES	Test
TEST	
TEST STRIP	
TEST STRIPS	
TIN	tincture
TOP CREAM	topical cream
TOP OINT	topical ointment
TOP SOL	topical solution
TOP SOLN	topical solution
TOPICAL	
TOPICAL CREAM	
TOPICAL SOLUTION	
TRO	troche
TTB	Time release tablet
TUB	Tube
TUBE	
UNDERWEAR	
UNIT DOSE	
UNT	Unit
VAGINAL CREAM	
VAPORIZER	
VIA	Vial
VIAL	
VIAL(S)	
VIL	Vial
WAF	wafer
WALKER	
WASH	
WIPES	
Z-PAK	

Attachment 2

Definitions of Codes and Abbreviations for RXFRMUNT

Code	Description
-7	refused
-8	don't know
-9	not ascertained
CC	Cubic centimeter
GM	Gram
L	Liter
ML	milliliter
OZ	ounce

Attachment 3

Definitions of Abbreviations, Codes and Symbols for RXSTRUNT

Abbreviations, Codes and Symbols	Definition
-7	refused
-8	don't know
-9	not ascertained
%	percent
09	compound
ACTIVATION	activation
ACTUATION	actuation
CC	cubic centimeters
CM2	square centimeter
DOSE	dose
DRP	drop
EL	ELISA (enzyme linked immunosorbent assay)
G	gram
GM	gram
GR	grain
HR or HRS	hour, hours
INH	inhalation
IU	international unit
MCG	microgram
MEQ	microequivalent
MG	milligram
ML	milliliter
MMU	millimass units
OZ	ounce
PACKET	packet
PFU	plaque forming units
SQ CM	square centimeter
U	units