

## **CMS 2008 Prescription Drug Profiles Public Use File (PUF)**

### **Data Dictionary and Codebook**

This is a profile-level file with the following variables. See the General Documentation for an overview of file contents, data source, information about suppression, and analytic utility.

#### **BENE\_SEX\_IDENT\_CD**

This field indicates the sex of the beneficiary. It is based on the BENE\_SEX\_IDENT\_CD variable in the Beneficiary Summary File (BSF).

<b>Variable Value</b>	<b>Formatted Value</b>	<b>Frequency</b>	<b>Frequency (%)</b>
1	Male	368,635,108	35.550
2	Female	668,300,568	64.450

Note: Percentages may not add up to 100% due to rounding.

#### **BENE\_AGE\_CAT\_CD**

This categorical variable is based on the beneficiary's age at end of the reference year (2008). In the event the beneficiary died during the reference year, the age at the date of death is used. It is based on the BENE\_AGE\_AT\_END\_REF\_YR variable in the Beneficiary Summary File (BSF).

<b>Variable Value</b>	<b>Formatted Value</b>	<b>Frequency</b>	<b>Frequency (%)</b>
1	Under 65	235,978,902	22.757
2	65 - 69	170,922,588	16.483
3	70 - 74	174,962,475	16.873
4	75 - 79	158,240,409	15.260
5	80 - 84	139,550,521	13.458
6	85 & Older	157,280,781	15.168

Note: Percentages may not add up to 100% due to rounding.

#### **RXNORM\_RXCUI**

This variable indicates the corresponding drug's RxNorm RxCUI (RxNorm Concept Unique Identifier). RxNorm is a normalized naming system for generic and branded drugs produced by

the National Library of Medicine. The PUF contains 1,254 distinct RXCUIs (including "missing" identified by a value of 0). The table below provides the frequency of the corresponding RxNorm RxCUI for the top ten drug names. It is determined by matching the PROD\_SRVC\_ID variable in the PDE File with the National Drug Code (NDC) in RxNorm.

Variable Value	Formatted Value	Frequency	Frequency (%)
6918	Metoprolol	31,144,021	3.003
10582	Thyroxine	31,011,803	2.991
29046	Lisinopril	28,515,150	2.750
36567	Simvastatin	27,511,334	2.653
214182	Acetaminophen/Hydrocodone	23,989,992	2.314
4603	Furosemide	23,948,935	2.310
0	Unknown/Missing	21,588,694	2.082
17767	Amlodipine	19,936,672	1.923
83367	Atorvastatin	18,445,966	1.779
6809	Metformin	18,139,813	1.749
-	All other values	792,703,296	76.447

Note: Percentages may not add up to 100% due to rounding.

### DRUG\_MAJOR\_CLASS

This categorical field indicates the major class of the drug. This variable is created using the VA\_CLASS variable in the Veterans Affairs National Drug File (VA-NDF) database. The DRUG\_MAJOR\_CLASS groups drugs into major (Level 0) categories based on mechanism and effect. The PUF contains 30 distinct drug classes (including "missing" identified by a value of 0). The ten most frequent values for major drug class and their frequencies are provided below.

Variable Value	Formatted Value	Frequency	Frequency (%)
CV000	Cardiovascular Medications	360,823,260	34.797
CN000	Central Nervous System Medications	201,792,480	19.460
HS000	Hormones/Synthetics/Modifiers	130,941,328	12.628
GA000	Gastrointestinal Medications	57,376,365	5.533
AM000	Antimicrobials	40,288,614	3.885
0	Unknown/Missing	37,572,643	3.623
MS000	Musculoskeletal Medications	37,458,339	3.612
BL000	Blood Products/Modifiers/Volume Expanders	33,199,098	3.202

Variable Value	Formatted Value	Frequency	Frequency (%)
RE000	Respiratory Tract Medications	31,524,278	3.040
OP000	Ophthalmic Agents	20,913,380	2.017
-	All other values	85,045,891	8.202

Note: Percentages may not add up to 100% due to rounding.

### DRUG\_CLASS

This categorical field indicates the class of the drug. This variable is created using the VA\_CLASS variable in the VA-NDF database. The DRUG\_CLASS groups drugs into minor (Level 1) categories based on mechanism and effect. The PUF contains 263 distinct drug classes (including "missing" identified by a value of 0). The ten most frequent values for drug class and their frequencies are provided below.

Variable Value	Formatted Value	Frequency	Frequency (%)
CV350	Antilipemic Agents	77,347,724	7.459
CV100	Beta Blockers/Related	61,072,911	5.890
CN609	Antidepressants	48,887,289	4.715
CN101	Analgesics	47,095,698	4.542
HS502	Blood Glucose Regulation Agents	46,937,962	4.527
CV800	Ace Inhibitors	44,210,199	4.264
0	Unknown/Missing	40,994,888	3.953
CV200	Calcium Channel Blockers	40,639,868	3.919
GA900	Gastric Medications, Other	39,490,486	3.808
HS851	Thyroid Modifiers	31,633,415	3.051
-	All other values	558,625,236	53.873

Note: Percentages may not add up to 100% due to rounding.

### PDE\_DRUG\_TYPE\_CD

This categorical field indicates the type of the drug: unknown, generic drug, or brand name drug. It is determined by matching the PROD\_SRVC\_ID variable in the PDE File with the National Drug Code (NDC) in RxNorm.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	57,640	0.006

Variable Value	Formatted Value	Frequency	Frequency (%)
0	Unknown/Missing	21,588,694	2.082
1	Brand name	332,370,502	32.053
2	Generic name	682,918,840	65.859

Note: Percentages may not add up to 100% due to rounding.

## PLAN\_TYPE

This categorical variable indicates the type of plan associated with events in the profile. It is derived from the Contract\_ID code in the Part D Plan Characteristics File. This field summarizes plan type in four categories: (1) Prescription Drug Plans (PDP) (Contract\_ID starts with letter "S"), (2) Medicare Advantage Part D (Contract\_ID starts with letter "H" or "R"), (3) Other Plan (Contract\_ID starts with letter "E"), and (\*) Suppressed.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	388,427	0.037
1	MAPD	281,429,669	27.141
2	PDP	751,030,679	72.428
3	Other	4,086,901	0.394

Note: Percentages may not add up to 100% due to rounding.

## COVERAGE\_TYPE

This categorical variable indicates beneficiary liability of cost-sharing. The beneficiary cost share group code associated with each event is taken from the month in which the event is recorded. This variable is created using the CST\_SHR\_GRP\_CD\_01 to CST\_SHR\_GRP\_CD\_12 variables in the Beneficiary Summary File. Cost-sharing codes are classified into five categories: (0) Missing/Unknown, (1) Dual Eligible (variable values of 01, 02 or 03 in the month of the event), (2) Low Income Subsidy (LIS) (variable values of 04, 05, 06, 07, or 08 in the month of the event), (3) No Subsidy (variable values of 00, XX, 09, 10, 11, 12 or 13 in the month of event), and (\*) Suppressed.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	903,776	0.087
0	Missing/Unknown	14,327	0.001
1	Dual eligible	459,364,576	44.300

Variable Value	Formatted Value	Frequency	Frequency (%)
2	Low income subsidy	59,610,372	5.749
3	No Subsidy	517,042,625	49.863

Note: Percentages may not add up to 100% due to rounding.

### **BENEFIT\_PHASE**

Indicates the benefit phase in which the claim was expected to occur based on a data of service ordering of the beneficiary's claims, the beneficiary's accumulated gross drug and out-of-pocket costs, and the plan's deductible, initial coverage limit (ICL) and out-of-pocket threshold (OOPT) amount. This variable is created using the BENEFIT\_PHASE variable in the PDE File.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	3,114,941	0.300
0	No information	63,292,090	6.104
1	Catastrophic	87,707,528	8.458
2	Deductible	60,673,465	5.851
3	Initial coverage limit	188,564,388	18.185
4	Non-covered drug	8,556,625	0.825
5	Pre-initial coverage limit	625,026,639	60.276

Note: Percentages may not add up to 100% due to rounding.

### **DRUG\_BENEFIT\_TYPE**

Indicates the type of Part D benefit structure used by the plan benefit package. This variable is created using the DRUG\_BENEFIT\_TYPE variable in the Part D Plan Characteristics File.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	7,093,921	0.684
0	No plan benefit information	62,768,244	6.053
1	Defined standard	150,260,008	14.491
2	Actuarially equivalent	202,870,889	19.564
3	Basic alternative	275,737,542	26.592
4	Enhanced alternative	338,205,072	32.616

Note: Percentages may not add up to 100% due to rounding.

## PRESCRIBER\_TYPE

This categorical field indicates the type of the prescriber. This variable is created using the PRIMARY\_TAXONOMY\_CD variable in the Prescriber Characteristics File. It is based on the National Uniform Claims Committee (NUCC) taxonomy code reported as primary by the prescriber in the National Plan & Provider Enumeration System (NPPES) data base. Prescriber codes are classified into six categories including the “suppressed” identified by “\*”.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	21,613,408	2.084
1	Family medicine	295,253,119	28.474
2	Internal medicine	393,486,675	37.947
3	Psychiatry & neurology	35,515,898	3.425
4	Specialist	35,507,851	3.424
5	Other	255,558,725	24.646

Note: Percentages may not add up to 100% due to rounding.

## GAP\_COVERAGE

Indicator for whether the plan benefit package offers gap coverage. This variable is created using the GAP\_COVERAGE variable in the Part D Plan Characteristics File. It summarizes gap coverage as: (0) Unknown/Missing, (1) No gap coverage (variable values of "N"), (2) Offers gap coverage (all other variable values), and (\*) Suppressed.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	24,292,041	2.343
0	Unknown/Missing	191,048	0.018
1	No gap coverage	822,619,215	79.332
2	Plan offers gap coverage	189,833,372	18.307

Note: Percentages may not add up to 100% due to rounding.

## TIER\_ID

This variable represents the minimum cost sharing tier in which the product was placed in the sponsor's formulary. This variable is created using the TIER\_ID variable in the PDE File. As each plan may have different tier definitions, descriptions for variable values 1-6 in the table are left blank.

Variable Value	Formatted Value	Frequency	Frequency (%)
*	Suppressed	29,273,966	2.823
01	-	675,470,419	65.141
02	-	243,306,020	23.464
03	-	48,450,769	4.672
04	-	2,694,802	0.260
05	-	694,740	0.067
06	-	55,806	0.005
NA	Drug on the PDE does not link to the plan's formulary	36,798,106	3.549
XX	Unable to link to plan	191,048	0.018

Note: Percentages may not add up to 100% due to rounding.

### MEAN\_RXHCC\_SCORE

This is the average of the RxHCC score associated with each beneficiary with a prescription drug event in the profile. The RxHCC score is created by the CMS-RxHCC risk-adjustment model. The model assigns a risk score to reflect the health status of each beneficiary according to demographic variables and the beneficiary's diagnosis history. A higher risk score correlates to higher estimated costs for a beneficiary. Payments to Medicare Part D Plans are adjusted by CMS according to the risk scores of the beneficiaries.

Variable Value Range	Frequency	Frequency (%)
Unknown/Missing	919,341	0.089
0.556 – 1.000	41,343,665	3.987
1.001 – 1.250	394,672,192	38.061
1.251 – 1.500	420,881,216	40.589
1.501 – 4.110	179,119,262	17.274

Note: Percentages may not add up to 100% due to rounding.

### AVE\_DAYS\_SUPPLY

This is the arithmetic average number of days' supply of medication dispensed by the pharmacy for all events in profile. This variable is created using the DAYS\_SUPPLY\_NUM variable in the PDE

file. The following summary shows the average number of days' supply rounded to the nearest day.

Variable Value Range	Frequency	Frequency (%)
0 – 30.0	294,184,549	28.371
30.1 – 33.0	207,012,194	19.964
33.1 – 39.0	281,070,968	27.106
39.1 – 245.0	254,667,965	24.560

Note: Percentages may not add up to 100% due to rounding.

### **AVE\_TOT\_DRUG\_COST**

This is the arithmetic average dollar amount for all events in the profile on the entire cost of the prescription based on the sum of ingredient cost paid, dispensing fee paid, and total amount attributed to sales tax. This variable is created using the TOT\_RX\_CST\_AMT variable in the PDE file. The following summary shows the average total drug cost rounded to the nearest dollar.

Variable Value Range	Frequency	Frequency (%)
\$0- \$13.0	275,145,174	26.533
\$13.1 - \$25.0	256,284,218	24.716
\$ 25.1- \$95.0	252,138,387	24.316
\$95.1 - \$41,255.0	253,367,897	24.434

Note: Percentages may not add up to 100% due to rounding.

### **AVE\_PTNT\_PAY\_AMT**

This is the arithmetic average dollar amount for all events in the profile on what the beneficiary paid that is not reimbursed by a third party (e.g., copayments, coinsurance, deductible or other patient pay amounts). This variable is created using the PTNT\_PAY\_AMT variable in the PDE file. The following summary shows the average patient payment amount rounded to the nearest dollar.

Variable Value Range	Frequency	Frequency (%)
Unknown/Missing	90,771,365	8.754



Variable Value Range	Frequency	Frequency (%)
\$0 - \$1.2	252,865,199	24.386
\$1.3 - \$4.1	223,930,127	21.595
\$4.2 - \$7.5	232,924,698	22.463
\$7.6 - \$4,208.0	236,444,287	22.802

Note: Percentages may not add up to 100% due to rounding.

### **PDE\_CNT**

This variable contains the number of Part D events associated with each profile. The sum of this variable (1,036,935,676) is the total number of Part D events for the 100% population of Fee-for-Service beneficiaries in 2008.

### **BENE\_CNT\_CAT**

This variable is based on the number of unique beneficiaries associated with the profile. It is categorized into six values.

Variable Value	Formatted Value	Frequency	Frequency (%)
1	11 – 15	23,939,132	2.309
2	16 – 20	17,767,290	1.713
3	21 – 50	48,287,101	4.657
4	51 – 100	49,665,241	4.790
5	101 – 500	200,300,473	19.317
6	501 +	696,976,439	67.215

Note: Percentages may not add up to 100% due to rounding.