CMS 2010 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).

Variable Value	Formatted Value	Frequency	Frequency (%)
1	Male	407,083,843	36.671
2	Female	703,005,311	63.329

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 (2010). 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).

Variable Value	Formatted Value	Frequency	Frequency (%)
1	Under 65	265,185,291	23.889
2	65 - 69	184,585,324	16.628
3	70 - 74	188,676,103	16.996
4	75 - 79	162,999,181	14.683
5	80 - 84	142,604,898	12.846
6	85 & Older	166,038,357	14.957

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 1,229 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 (%)
36567	Simvastatin	39,128,791	3.525
10582	Thyroxine	33,273,050	2.997
6918	Metoprolol	32,037,672	2.886
29046	Lisinopril	31,832,596	2.868
214182	Acetaminophen/Hydrocodone	28,189,242	2.539
17767	Amlodipine	25,878,893	2.331
7646	Omeprazole	25,374,407	2.286
4603	Furosemide	23,892,708	2.152
6809	Metformin	20,564,292	1.852
0	Unknown/Missing	18,057,071	1.627
-	All other values	831,860,432	74.936

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	378,626,099	34.108
CN000	Central Nervous System Medications	223,877,246	20.168
HS000	Hormones/Synthetics/Modifiers	139,153,649	12.535
GA000	Gastrointestinal Medications	63,692,300	5.738
AM000	Antimicrobials	43,564,170	3.924
MS000	Musculoskeletal Medications	41,040,797	3.697
0	Unknown/Missing	38,210,581	3.442
BL000	Blood Products/Modifiers/Volume Expanders	34,811,852	3.136

Variable Value	Formatted Value	Frequency	Frequency (%)
RE000	Respiratory Tract Medications	33,873,494	3.051
OP000	Ophthalmic Agents	23,251,483	2.095
-	All other values	89,987,483	8.106

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	87,293,392	7.864
CV100	Beta Blockers/Related	63,740,993	5.742
CN609	Antidepressants	54,113,342	4.875
CN101	Analgesics	52,048,935	4.689
HS502	Blood Glucose Regulation Agents	48,957,230	4.410
CV800	Ace Inhibitors	46,045,847	4.148
0	Unknown/Missing	44,412,723	4.001
CV200	Calcium Channel Blockers	43,524,946	3.921
GA900	Gastric Medications, Other	43,126,014	3.885
HS851	Thyroid Modifiers	33,652,929	3.032
-	All other values	593,172,803	53.435

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	57,403	0.005

Variable Value	Formatted Value	Frequency	Frequency (%)
0	Unknown/Missing	18,057,071	1.627
1	Brand name	291,380,901	26.248
2	Generic name	800,593,779	72.120

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	406,619	0.037
1	MAPD	342,061,439	30.814
2	PDP	762,130,548	68.655
3	Other	5,490,548	0.495

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	929,731	0.084
0	Unknown/Missing	19,130	0.002
1	Dual eligible	496,930,662	44.765

Variable Value	Formatted Value	Frequency	Frequency (%)
2	Low income subsidy	64,628,904	5.822
3	No Subsidy	547,580,727	49.328

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	3,260,205	0.294
0	No information	75,320,581	6.785
1	Catastrophic	87,375,217	7.871
2	Deductible	86,358,657	7.779
3	Initial coverage limit	181,794,463	16.377
4	Non-covered drug	7,297,816	0.657
5	Pre-initial coverage limit	668,682,215	60.237

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	6,985,424	0.629
0	No plan benefit information	74,532,350	6.714
1	Defined standard	125,459,412	11.302
2	Actuarially equivalent	331,695,876	29.880
3	Basic alternative	196,061,975	17.662
4	Enhanced alternative	375,354,117	33.813

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	21,564,131	1.943
1	Family medicine	327,145,574	29.470
2	Internal medicine	434,639,707	39.154
3	Psychiatry & neurology	41,886,388	3.773
4	Specialist	34,710,116	3.127
5	Other	250,143,238	22.534

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	24,657,251	2.221
0	Unknown/Missing	718,209,097	64.698
1	No gap coverage	178,465,198	16.077
2	Plan offers gap coverage	188,757,608	17.004

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	30,628,824	2.759
0	Unknown/Missing	30,809,905	2.780
1	-	766,296,678	69.030
2	-	229,911,687	20.711
3	-	45,950,409	4.139
4	-	6,170,754	0.556
5	-	320,804	0.029
6	-	93	0.000

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	1,037,014	0.093
0.494 – 1.000	27,979,299	2.520
1.001 – 1.250	354,471,287	31.932
1.251 – 1.500	476,365,930	42.912
1.501 – 4.224	250,235,624	22.542

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_SUPLY_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	291,142,194	26.227
30.1 – 33.0	168,752,366	15.202
33.1 – 39.0	293,094,680	26.403
39.1 – 916.0	357,099,914	32.169

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	336,987,095	30.357
$13.1 - $25.0	285,529,664	25.721
$ 25.1- $95.0	227,726,639	20.514
$95.1 - $47,512.0	259,845,756	23.408

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	89,929,766	8.101
$0 - $1.2	235,366,983	21.203
$1.3 - $4.1	257,305,948	23.179

Variable Value Range	Frequency	Frequency (%)
$4.2 - $7.5	240,950,805	21.706
$7.6 - $5,501.0	286,535,652	25.812

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 (1,110,089,154) is the total number of Part D events for the 100% population of Fee-for-Service beneficiaries in 2010.

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	25,215,800	2.272
2	16 – 20	18,342,202	1.652
3	21 – 50	49,116,962	4.425
4	51 – 100	49,752,131	4.482
5	101 – 500	196,598,475	17.710
6	501 +	771,063,584	69.460

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