

## **CMS 2008 Basic Stand Alone (BSA) Prescription Drug Events Public Use File (PUF)**

### **Data Dictionary and Codebook**

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

#### **PDE\_EVENT\_ID**

This field contains a cryptographic event ID. As this is an event-level file, each row on the file has a unique value for PDE\_EVENT\_ID. PDE\_EVENT\_ID cannot be used to link to any other data files released by CMS or other sources. Because of the manner in which the field was created, PDE\_EVENT\_ID can also serve as a random number, representing random draws from a uniform distribution.

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

This field indicates the sex of the beneficiary.

<b>Variable Value</b>	<b>Formatted Value</b>	<b>Frequency</b>	<b>Frequency (%)</b>
1	Male	17,754,634	35.279
2	Female	32,571,354	64.721

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.

<b>Variable Value</b>	<b>Formatted Value</b>	<b>Frequency</b>	<b>Frequency (%)</b>
1	Under 65	11,500,646	22.852
2	65 - 69	8,299,441	16.491
3	70 - 74	8,477,216	16.845
4	75 - 79	7,625,479	15.152
5	80 - 84	7,871,719	15.641
6	85 and older	6,551,487	13.018

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

## PDE\_DRUG\_CD

It is the categorical variable which represents the drug names which are derived from RxNorm database, reflecting the ingredient that gives the drug its distinctive properties. The PUF contains 1,022 distinct drug codes (including "missing" identified by a value of 0). If a drug consists of more than one ingredient, then each ingredient is separated by a "/". The table below provides the frequency of top 10 drug names. The ten most frequent drugs and their frequencies are provided below.

Variable Value	Formatted Value	Frequency	Frequency (%)
392	Metoprolol	1,537,213	3.055
553	Thyroxine	1,530,451	3.041
350	Lisinopril	1,416,475	2.815
518	Simvastatin	1,363,073	2.708
254	Furosemide	1,190,658	2.366
8	Glatiramer	1,173,283	2.331
0	Unknown/Missing	1,119,820	2.225
36	Amlodipine	988,083	1.963
610	Atorvastatin	920,900	1.830
373	Metformin	901,326	1.791
-	All other values	38,184,706	75.875

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

## PDE\_DRUG\_STR\_CD

This variable indicates the strength of the drug. It is derived from the drug name in the RxNorm database. Based on the active ingredient(s) of the drug, strengths are given to three significant digits, and can be expressed as ratios of two unit types where appropriate. The unit types are included in a separate variable. The PUF contains 665 distinct drug strength codes (including "missing" identified by a value of 0). If a drug consists of more than one ingredient, then the strength of each ingredient is separated by a "/" in PDE\_DRUG\_STR\_CD. The ten most frequent values for drug strength and their frequencies are provided below.

Variable Value	Formatted Value	Frequency	Frequency (%)
161	10	5,060,622	10.056
265	20	4,671,441	9.282
375	33200/37500/10000	3,300,936	6.559
391	5	2,628,102	5.222

Variable Value	Formatted Value	Frequency	Frequency (%)
175	100	2,496,149	4.960
284	25	2,405,345	4.780
407	50	2,374,372	4.718
421	500	1,829,970	3.636
0	Unknown/Missing	1,119,820	2.225
484	75	959,668	1.907
-	All other values	23,479,563	46.655

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

### PDE\_DRUG\_STR\_UNITS

This variable indicates the unit of the strength of the drug. It is derived from the drug name in the RxNorm database. The PUF contains 46 distinct drug strength units (including "missing" identified by a value of 0). If a drug consists of more than one ingredient, then the unit of strength for each ingredient is separated by a “,” in PDE\_DRUG\_STR\_UNITS. The ten most frequent values for unit of drug strength and their frequencies are provided below.

Variable Value	Formatted Value	Frequency	Frequency (%)
3	MG	38,303,423	76.111
4	MG, MG	4,507,358	8.956
18	MG/ML	1,991,051	3.956
0	Unknown/Missing	1,119,820	2.225
9	MG/ACTUAT	1,075,885	2.138
1	MEQ	867,568	1.724
39	UNT/ML	640,474	1.273
10	MG/ACTUAT, MG/ACTUAT	411,154	0.817
13	MG/MG	324,408	0.645
11	MG/HR	302,071	0.600
-	All other values	782,776	1.555

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

The following table provides the descriptions of the abbreviations:

Abbreviation	Description
MG	Milligram
ML	Milliliter
MEQ	Milliequivalent
UNT	Unit
HR	Hour
ACTUAT	Actuation <sup>(1)</sup>
SQCM	Square Centimeter

(1) Actuation refers to a measured dose per activation of a dispensing device; e.g., in an inhaler, the strength of the clinical drug is given by how much is dispensed with each actuation.

#### **PDE\_DRUG\_DOSE\_CD**

It is the categorical variable for the dose form of the drug such as topical solution or oral tablet. The PUF contains 71 distinct values for dose form (including "missing" identified by a value of 0). The ten most frequent values for dose form and their frequencies are provided below.

Variable Value	Formatted Value	Frequency	Frequency (%)
41	Oral Tablet	33,018,757	65.610
36	Oral Capsule	3,541,581	7.037
13	Extended Release Tablet	3,526,898	7.008
9	Enteric Coated Capsule	1,669,483	3.317
11	Extended Release Capsule	1,207,174	2.399
0	Unknown/Missing	1,119,820	2.225
34	Ophthalmic Solution	775,384	1.541
25	Metered Dose Inhaler	604,599	1.201
10	Enteric Coated Tablet	602,342	1.197
17	Injectable Solution	509,456	1.012
-	All other values	3,750,494	7.452

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

## PDE\_DRUG\_CLASS\_CD

This categorical field indicates the class of the drug. It is a set of 28 different values. This variable is created using the VA\_CLASS variable in the VA-NDF database. The PDE\_DRUG\_CLASS\_CD groups drugs into major (Level 0) categories based on mechanism and effect. The PUF contains 28 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 (%)
9	Cardiovascular medications	16,042,489	31.877
8	Central nervous system medications	8,344,276	16.580
0	Unknown/Missing	7,263,793	14.433
14	Hormones/Synthetics/Modifiers	5,409,130	10.748
12	Gastrointestinal medications	2,383,120	4.735
17	Musculoskeletal medications	1,754,840	3.487
3	Antimicrobials	1,692,131	3.362
23	Respiratory tract medications	1,518,788	3.018
7	Blood products/Modifiers/Volume expanders	1,470,218	2.921
19	Ophthalmic agents	1,000,128	1.987
-	All other values	3,447,075	6.849

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

## PDE\_DRUG\_QTY\_DIS

The quantity of the drug dispensed, in terms of units, grams, milliliters, etc. This variable is created using the QTY\_DSPNSD\_NUM variable in the original PDE file.

Variable Value	Frequency	Frequency (%)
0	122,537	0.243
1	305,637	0.607
2	190,821	0.379
3	1,238,190	2.460
4	850,706	1.690
5	473,051	0.940
6	550,854	1.095
7	257,461	0.512
8	125,683	0.250

<b>Variable Value</b>	<b>Frequency</b>	<b>Frequency (%)</b>
9	260,570	0.518
10	1,335,539	2.654
15	1,899,884	3.775
20	1,054,114	2.095
25	221,214	0.440
30	23,851,594	47.394
40	429,163	0.853
50	775,451	1.541
60	7,236,999	14.380
70	37,231	0.074
80	151,979	0.302
90	5,044,267	10.023
100	1,987,479	3.949
200	1,251,257	2.486
300	216,210	0.430
400	119,730	0.238
500	226,883	0.451
600	8,832	0.018
700	2,367	0.005
800	3,026	0.006
900	21,895	0.044
1,000	20,966	0.042
2,000	11,092	0.022
3,000	2,501	0.005
4,000	36,452	0.072
5,000	445	0.001
6,000	517	0.001
7,000	175	0.000
8,000	242	0.000
9,000	17	0.000
10,000	47	0.000
12,000	14	0.000
13,000	20	0.000
14,000	4	0.000
15,000	1	0.000

Variable Value	Frequency	Frequency (%)
16,000	9	0.000
17,000	33	0.000
18,000	14	0.000
20,000	9	0.000
21,000	9	0.000
22,000	3	0.000
28,000	4	0.000
30,000	2,585	0.005
43,000	1	0.000
60,000	184	0.000
100,000	14	0.000
120,000	6	0.000

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

#### **PDE\_DRUG\_DAY\_SPLY\_CD**

The number of days' supply for the dispensed drug. This variable is created using the DAYS\_SUPLY\_NUM variable in the original PDE file.

Variable Value	Formatted Value	Frequency	Frequency (%)
1	1 - 7 days	3,193,136	6.345
2	8 - 14 days	2,188,673	4.349
3	15 - 21 days	1,814,563	3.606
4	22 - 29 days	2,336,092	4.642
5	30 - 59 days	35,507,905	70.556
6	60 - 89 days	640,703	1.273
7	90 or more days	4,644,916	9.230

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

#### **PDE\_DRUG\_COST**

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 original PDE file.

<b>Variable Value<sup>(1)</sup> (\$)</b>	<b>Frequency</b>	<b>Frequency (%)</b>
0	4,608,848	9.158
10	15,734,658	31.265
20	6,660,643	13.235
30	3,478,810	6.913
40	1,976,881	3.928
50	1,377,147	2.736
60	1,410,728	2.803
70	1,303,540	2.590
80	1,773,516	3.524
90	1,537,692	3.055
100	3,312,958	6.583
150	3,215,638	6.390
200	1,467,736	2.916
250	708,815	1.408
300	482,723	0.959
350	335,078	0.666
400	222,126	0.441
450	129,022	0.256
500	113,341	0.225
550	92,643	0.184
600	49,983	0.099
650	38,330	0.076
700	36,174	0.072
750	42,900	0.085
800	28,336	0.056
850	28,011	0.056
900	25,957	0.052
950	9,857	0.020
1,000	70,091	0.139
2,000	34,209	0.068
3,000	7,868	0.016
4,000	4,565	0.009
5,000	3,159	0.006
6,000	1,372	0.003
7,000	1,984	0.004



Variable Value <sup>(1)</sup> (\$)	Frequency	Frequency (%)
8,000	395	0.001
9,000	110	0.000
10,000	131	0.000
20,000	13	0.000

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

(1) Note that a total drug cost amount between \$0 and \$4.99 is rounded to \$0 according to the rounding rules. Hence, the corresponding value for \$10 in the PUF is a value between \$5 and \$14.99 in the initial 5% sample file.

### PDE\_DRUG\_PAT\_PAY\_CD

This categorical variable indicates how much of the total drug cost 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 original PDE file.

Variable Value	Formatted Value	Frequency	Frequency (%)
1	Patient did not pay any amount	13,252,666	26.334
2	Patient paid an amount greater than \$0 but less than the total drug cost	31,061,536	61.721
3	Patient paid the total drug cost	6,011,786	11.946

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.

Variable Value	Formatted Value	Frequency	Frequency (%)
0	Unknown/Missing	1,119,820	2.225
1	Brand name	16,112,188	32.016
2	Generic name	33,093,980	65.759

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