DEPARTMENT OF HEALTH

Minnesota All Payer Claims Database Prescription Drug Public Use Files: A User Guide

SEPTEMBER 2026

Table of contents

Minnesota All Payer Claims Database Prescription Drug Public Use Files: A User Guide 1
- Background 2
  - Public Use File Overview 2
  - Definition of a Prescription Drug 2
- Data Elements 3
- Metadata 4
  - Exclusions 4
- Other Important Data Considerations 4

Appendix B: Interpreting PUF Data 7
- Detail PUF 7
- Summary PUF 9

Appendix C: User Calculations 10
- Means for Alternative Units 10
- Aggregating Records 10
  - Counts 10
  - Means 11
  - Medians 11
  - Standard Deviations 11

MNAPCD
All Payer Claims Database

Background

The Minnesota Department of Health (MDH) maintains the Minnesota All Payer Claims Database (MN APCD), a repository of healthcare claims data that supports statewide analyses of healthcare costs, quality, and utilization. Under legislative mandate, MDH releases publicly available summary information from the MN APCD in the form of public use files (PUFs). PUF data are delivered in spreadsheets with aggregated records that prevent the identification of individual members, providers, and health plans. Currently available MN APCD PUFs, derived from medical and pharmacy claims, contain summary data on healthcare services, healthcare utilization, primary diagnoses, provider specialties, members, and prescription drugs.[1] This document introduces the Prescription Drug PUFs, illustrates how to interpret PUF records, and includes technical instructions for users who wish to further aggregate PUF records.

Public use file overview

Two versions of MN APCD prescription drug PUFs are available:

- The *Detail* PUF contains retail pharmacy claims data that have been aggregated by the first two segments of the National Drug Code (NDC)
- The S*ummary* PUF contains retail pharmacy claims data that have been aggregated by nonproprietary drug name

PUF levels of aggregation are further explained in the "Definition of a Prescription Drug" section.

Summary and *Detail* PUFs are stratified by payer type (commercial, Medicare, and Minnesota Health Care Programs).[2]

Prescription drugs administered in medical settings such as hospitals, infusion centers, nursing homes, or other medical offices—although often high in cost and significant drivers of growth—are not included in these PUFs. These drugs are generally billed in medical claims, as opposed to pharmacy claims, which were the basis for the PUFs.

Costs in the PUFs represent healthcare transactions *before* any applicable rebates. Currently, data on rebates do not exist in a transparent manner and are not required reporting elements under state law that authorizes maintaining the MN APCD. Although these PUFs represent the single largest collection of prescription drug use data for Minnesota, they do not represent prescription drug use by every Minnesotan. For example, the MN APCD does not include claims for certain payers, and the volume of available commercial data has been affected by a recent Supreme Court ruling (see "Other Important Data Considerations" section). **Users must carefully consider their use and interpretation of the data.**

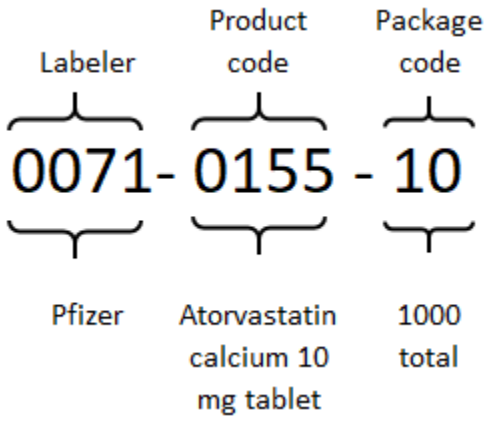
MDH developed the PUFs in partnership with Mathematica and welcomes questions from users at: health.APCD@state.mn.us. MDH appreciates user feedback about experience with the PUFs.

Definition of a prescription drug

Prescription drugs are broadly defined by active ingredient (nonproprietary drug name) or more narrowly by the specific product, containing the active ingredient, that a given pharmaceutical

company produces. Each product is assigned a unique NDC, which consists of three segments (Figure 1). The first segment identifies the labeler (i.e., the pharmaceutical company). The second segment, which is specific to the labeler, identifies a distinct product in terms of active ingredient(s), active strength, and dosage form. The third segment is a package code, which indicates how a drug is packaged for sale to pharmacies. Package codes may vary with the first two segments of an NDC, but this variation is not relevant to individual prescriptions. Multiple NDCs can share a single nonproprietary drug name.

Figure 1. Illustration of three segments of a National Drug Code.



Data elements

The PUFs include a number of data elements, including drug characteristics, utilization measures, calculated metrics and rankings, and summary data on the demographics of prescription drug users. NDCs allow linking of descriptive drug characteristics from reference data (Medi-Span[3]) to measures of prescription drug use and spending from claims data. Drug characteristics in the PUFs, which differ between *Summary* and *Detail* files (Figure 2), include:

- Nonproprietary name
- Proprietary name
- Brand/generic classification
- Therapeutic class
- Labeler name
- Dosage form (e.g., tablet, cream, injection)
- Active strength (e.g., a numerical value)
- Active ingredient unit (e.g., milligrams, milligrams per milliliter)

Figure 2. Example of select drug characteristic variables in *Detail* PUF.

Nonproprietary name	Proprietary name	Brand / generic	Therapeutic Class	Labeler	Form	Strength	Unit
ACETAMINOPHEN	UP & UP ACETAMINOPHEN	Generic	Analgesics & Anesthetics	TARGET	TABLET	500	MG/ 1

The assignment of therapeutic classes to drugs relied on the Medi-Span reference dataset. Therapeutic classes provide a helpful categorization with which to easily view and understand prescription drug data that span thousands of drugs.

At both the *Summary* and *Detail* level, the PUFs provide measures of the number of prescriptions and the number of unique members with a prescription, as well as calculated metrics of mean, median, and standard deviation for:

- Days' supply
- Quantity dispensed
- Health plan paid amount
- Member paid amount
- Total paid amount[4]

Additionally, the *Summary* PUF contains drug rankings by a number of dimensions (e.g., cost per prescription) and distributions of prescription drug user demographic characteristics, including:

- Member age group
- Member sex
- Rural/urban classification of member home ZIP code

Metadata

Exclusions

MN APCD retail pharmacy claims meeting any of the following criteria were excluded from the PUFs to optimize the analytic usefulness of the files:

- Duplicate or denied claim
- Member with out-of-state residence
- Member with unknown sex
- Negative quantity dispensed
- Negative paid amounts

Additionally, claims with an NDC that could not be matched to Medi-Span (reference database) were excluded. After claims were aggregated to the unit of analysis for each PUF, rows with fewer than 11 unique members or 5 unique prescribers were removed to prevent identification and comply with applicable statutes and data use agreements. Approximately 4% of claims and 6% of total paid amount from the MN APCD are redacted from the Prescription Drug PUFs.

Other important data considerations

Minnesota policymakers structured the requirements for data submission under the MN APCD to focus on payers under its jurisdiction and payers who represent the primary volume of healthcare services in the state. As such, the MN APCD was not designed to capture pharmacy

(or medical) claims for individuals who are covered by Tricare, Veterans Affairs, the Indian Health Service, or Workers' Compensation. Additionally, the MN APCD does not include:

- Prescription drug use or spending by Minnesotans who are uninsured
- Claims for services provided by plans that do not cover general medical care, such as accident-only, vision, or dental plans
- Low-volume submitters of pharmacy claims, defined as having less than $300,000 in claims volume (exempt from submission to the MN APCD)
- Written prescriptions that were never filled

As noted earlier, prescription drugs administered in medical settings such as hospitals, infusion centers, nursing homes, or other medical offices *are* submitted to the MN APCD. However, the Prescription Drug PUFs include only prescription drugs obtained through a retail pharmacy.

There are a number of additional data characteristics that PUF users should consider. We have referred to most of these characteristics throughout the document but provide additional details here:

- **What price data are available?** Pricing for prescription drugs is opaque and complex. It evolves from negotiations between multiple parties across the supply chain and is influenced by a range of incentives. Absent robust transparency laws, the final price—paid by Medicaid, Medicare, or as a whole or by individual commercial payers—is a closely guarded trade secret. Data systems like the MN APCD generally capture the paid amount *before* rebate transactions occur. This means the actual transacted price is overstated for many drugs in the PUFs.

  Similarly, the MN APCD and the PUFs do not capture the influence of coupons or other discounts from list prices that pharmaceutical manufacturers selectively grant members. Interpretation of cost per prescription should consider supply measures such as quantity dispensed and days' supply.

- **Are self-insured claims part of the data?** In a decision released on March 1, 2016, the U.S. Supreme Court upheld a lower court's ruling that self-insured health plans, or large employers who retain the insurance risk for their employees, could not be required to submit claims data to a state's APCD (Gobeille v. Liberty Mutual Insurance Co.). The court found that requiring private self-insured plans to participate in state APCDs was preempted by the Employee Retirement Income Security Act (ERISA). The ruling does not prohibit the voluntary submission of self-insured plan data to the MN APCD, a decision that rests with employers themselves and not their brokers.

  Although Minnesota is working with self-insured employers and brokers to encourage reporting and create conditions that are conducive to doing so without additional burden, the court's ruling resulted in a substantial reduction in the volume of commercial claims beginning in the spring of 2016. Summing commercial prescription counts and costs of 2016 (and later) data in the PUFs would therefore result in a considerable underestimate of use and spending across the whole commercial market. To the extent that the demographics of fully and self-insured employees differ, user

characteristics could also be affected. The calculation of averages or per-unit measures for commercial enrollees are not expected to be materially impacted by the reduction in the data volume.

Appendix B: Interpreting PUF data

The following tables show subsets of data from the *Summary* and *Detail* PUFs to illustrate how to interpret key data elements. The sample *Detail* PUF data are derived from three records representing a single two segment NDC and all payer types. The two segment NDC, 60505-2580, is for a generic version of atorvastatin calcium (nonproprietary name), which is used to treat high cholesterol and sold under the brand name Lipitor. This particular NDC is a tablet containing 40 milligrams of the active ingredient and produced by Apotex.

Detail PUF

Table 1A shows a variety of mean paid amounts across payer types for the selected NDC in 2021 and 2022. Commercial insurers in 2021 accounted for 26,521 scripts with a mean days' supply of 73.8. The insurers paid an average of $5.06 per script while members paid an average of $9.80. The mean total amount paid was $14.87, which includes a very small amount paid by another source (not shown separately). Medicare accounted for 89,966 scripts with a mean days' supply of 62.0, lower than the commercial average. Medicare paid an average of $14.18 per script while members paid an average of $3.66. The mean total amount paid was $17.89. Minnesota Health Care Programs (MHCP) accounted for many fewer scripts than the other two payers at 19,779 with a mean days' supply of 40.8, about half that of the commercial insurers and Medicare. Minnesota Health Care Programs paid an average of $9.17 per script while members paid an average of $0.84. The mean total amount paid was $10.04.

Table 1A. Select mean cost statistics for two segment NDC 60505-2580 by payer type, 2021 and 2022.

Year	Product NDC	Payer	Number scripts	Days supply mean	Insurer paid mean	Member paid mean	Total paid mean
2021	60505-2580	Commercial	26,521	73.8	$5.06	$9.80	$14.87
2021	60505-2580	Medicare	89,966	62.0	$14.18	$3.66	$17.89
2021	60505-2580	MHCP	19,779	40.8	$9.17	$0.84	$10.04
2022	60505-2580	Commercial	32,447	74.0	$3.91	$8.32	$12.24
2022	60505-2580	Medicare	81,150	62.3	$13.76	$3.05	$16.86
2022	60505-2580	MHCP	20,412	46.1	$14.56	$0.87	$15.48

Comparing the calculated mean amounts across years—for example, calculating the unit or the percent change—is, in general, straightforward. All mean (or average) values can be compared directly. For example, as shown in Table 1B, the mean days' supply in 2022 was higher for all payers. The mean total price paid declined between 2021 and 2022 for commercial payers and Medicare but not for Minnesota Health Care Programs.

Comparing sum values—specifically, the number of scripts as shown in Table 1A, or the amounts paid—is more complicated, but only for commercial payers. Recall that commercial

scripts are understated in 2016 and subsequent years due to unreported self-insured commercial claims. However, sum values for Medicare and Minnesota Health Care Programs can be compared across all PUF years—with the caveat that underlying enrollment in those programs has changed due the numbers of Minnesotans becoming eligible (in particular, for Medicare) and Minnesota Health Care Program changes that affect eligibility.

Table 1B. Select percentage change cost statistics for two segment NDC 60505-2580 by payer type, 2021–2022.

Year	Product NDC	Payer	Number scripts	Days supply mean	Insurer paid mean	Member paid mean	Total paid mean
2021–2022	60505-2580	Commercial	22.3%	0.3%	-22.7%	-15.1%	-17.7%
2021–2022	60505-2580	Medicare	-9.8%	0.5%	-3.0%	-16.7%	-5.8%
2021–2022	60505-2580	MHCP	3.2%	13.0%	58.8%	3.6%	54.2%

Additional statistics for the same NDC are shown in Tables 2 and 3, focusing on average costs per user, script, and days' supply. For most patients, the cost per user is probably an annual cost. The exceptions are patients who started or stopped taking the drug during the year or were switched to a different NDC or changed payer during the year.[1] In 2020, the cost per user was over twice the cost per script—for example, among commercial payers, $42.71 per user compared with $14.87 per script. This implies that on average the users of this NDC filled over two scripts during the year. The cost per script in Table 2 is identical to the mean total paid in Table 1A (which is calculated per script). The cost per days' supply is more uniform across payers than the previous two measures: in 2021, $0.20 for commercial payers, $0.28 for Medicare, and $0.24 for Minnesota Health Care Programs. The cost per unit dispensed—in this case a single 40 milligram tablet—is slightly more varied at $0.19 for commercial payers, $0.28 for Medicare, and $0.25 for Minnesota Health Care Programs.

Table 2. Select unit cost statistics for two segment NDC 60505-2580 by payer type, 2021 and 2022.

Year	Product NDC	Payer	Number scripts	Cost per user	Cost per script	Cost per days supply	Cost per unit dispensed
2021	60505-2580	Commercial	26,521	$42.71	$14.87	$0.20	$0.19
2021	60505-2580	Medicare	89,966	$61.82	$17.89	$0.28	$0.28
2021	60505-2580	MHCP	19,779	$41.61	$10.04	$0.24	$0.25
2022	60505-2580	Commercial	32,447	$34.07	$12.24	$0.16	$0.16
2022	60505-2580	Medicare	81,150	$56.10	$16.86	$0.27	$0.26
2022	60505-2580	MHCP	20,412	$57.57	$15.48	$0.33	$0.34

[1] Patients continuing with the same NDC but changing payer type will appear in one row of the PUF for part of the year and another row for the rest of the year.

Summary PUF

Table 3 shows selected cost statistics from the *Summary* PUF for all atorvastatin calcium scripts in the MN APCD in 2021 and 2022 (including the NDC used to populate Tables 1 and 2.) Grouping by nonproprietary name combines brand and generic formulations of the drug, as well as all dosage forms and active strengths for the drug sold in Minnesota in each year.

Table 3. Select mean cost statistics for atorvastatin calcium by payer type, 2021 and 2022.

Year	Payer	Number scripts	Days supply mean	Insurer paid mean	Member paid mean	Total paid mean	Cost per days supply
2021	Commercial	397,163	72.1	$7.77	$7.64	$15.43	$0.21
2021	Medicare	974,987	71.9	$13.15	$4.25	$17.46	$0.24
2021	MHCP	46,518	46.8	$9.75	$0.68	$10.49	$0.22
2022	Commercial	418,679	72.9	$7.84	$7.72	$15.57	$0.21
2022	Medicare	928,725	72.9	$14.85	$3.30	$18.21	$ 0.24
2022	MHCP	207,833	55.4	$13.10	$0.68	$13.86	$0.24

In 2022, Medicare accounted for the largest number of atorvastatin calcium scripts (928,725), followed by commercial payers (418,679) and Minnesota Health Care Programs (207,833). Scripts paid by commercial payers or Medicare averaged more than a 70-day supply, while scripts paid by Minnesota Health Care Programs averaged just 55.4 days. The insurer paid, member paid, and total paid amounts were lowest for Minnesota Health Care Programs.

Appendix C: User calculations

Means for alternative units

The mean costs reported on each PUF are calculated over the number of prescriptions filled. Users may also wish to calculate average expenditures over different units—for example, per unique member, per days' supply, or per unit dispensed. Means for alternative units have already been computed in each PUF for the total paid amount, but it can also be calculated for the member-paid or insurer-paid amounts. For example, the average member cost per script can be calculated by dividing the total member cost by the number of scripts, and the average member cost per member can be calculated by dividing the total member cost by the number of unique members. Similarly, the average member cost per days' supply can be calculated by dividing the mean member cost by the mean days' supply, and the average member cost per unit dispensed can be calculated by dividing the mean member cost by the mean quantity dispensed.

Examples of these cost calculations are shown in Table 4 for the single NDC 60505-2580 used in the earlier illustrations, by payer type in 2022. The average cost per member is calculated by dividing the sum of the member paid amounts by the number of unique members. The average member cost per day is calculated by dividing the mean member paid amount by the mean days' supply.

Table 4. Select unit cost statistics for two segment NDC 60505-2580 by payer type, 2022.

Year	Payer	Unique members	Member paid sum	Cost per member	Days supply mean	Member paid mean	Member cost per day
2022	Commercial	11,658	$270,056.76	$23.16	74.0	$8.32	$0.11
2022	Medicare	24,391	$247,858.95	$10.16	62.3	$3.05	$0.05
2022	MHCP	5,489	$17,856.90	$3.25	46.1	$0.87	$0.02

Aggregating records

Users may wish to aggregate PUF records across payer type or combine selected drugs. Aggregation methods vary by type of statistic.

Counts

Counts of prescriptions filled and paid amounts can be summed across PUF records directly; however, summing counts of unique members across PUF records is more complex. Depending on the rows being summed, a given member may appear in more than one PUF record – most commonly by obtaining prescriptions for more than one drug. In such cases, the sum will overstate the number of unique members by counting some individuals more than once. Summing unique members across payer type within the same NDC or nonproprietary drug is

less likely to produce an overestimate of unique members than is summing across drugs, but changes in payer type within a given year do occur.

Means

When records in the PUF are aggregated, the mean of the aggregate record (i.e., the grand mean for the set of records) can be calculated as the weighted average of the means of the individual records, where the weights are the numbers of prescriptions. This calculation is illustrated in Table 5 using data from the three 2022 records in Table 5. This represents an aggregation of records for a single NDC over payer types.

Table 5. Calculation of mean for an aggregate of PUF records (grand mean)

PUF record	Number of scripts	Mean total paid	Number of scripts x mean total paid	Grand mean
1	32,447	$12.24	397,151.28	N/A
2	81,150	$16.86	1,368,189 .00	N/A
3	20,412	$15.48	315,977.76	N/A
Sum	134,009	N/A	2,081,318.04	$15.53*

*2,081,318.04/ 134,009

Medians

Calculating the exact median of a measure requires access to the underlying microdata (i.e., individual claims). Unlike means, the weighted median of a set of individual PUF records is not the median of the aggregate PUF record. However, with a large number of PUF records, none of which having a substantially greater number of claims, the weighted median of the individual record medians provides a good approximation of the median of the aggregate record. The calculation illustrated in Table 5 can be used to obtain the approximate median for an aggregate of PUF records by substituting the variable median total paid for mean total paid.

Standard deviations

Calculating the standard deviation for an aggregate of PUF records is more complex than calculating the mean, as it requires performing several computational operations on the data from the individual records. The operations described below are illustrated in the corresponding numeric columns in Table 6. Columns with non-numeric names represent PUF data.

(1) Square the standard deviation from each record and multiply it by the number of scripts. Summing these products across records yields the *within group sum of squares*.[a]

(2) Calculate the difference between each record mean and the grand mean (see Table 5 for grand mean calculation) and square this difference.

(3) Multiply the squared difference from (2) by the number of scripts. Summing these values across records yields the *between group sum of squares*.[b]

(4) Sum the *within group sum of squares* and the *between group sums of squares*, and divide the result by the total number of scripts in the aggregate record to calculate a mean squared deviation or variance. Take the square root of the variance to obtain the standard deviation of the aggregate record.

Table 6. Calculation of standard deviation for an aggregate of records.

PUF record	Number of scripts	Total paid SD	(1)	Total paid mean	(2)	(3)	(4)
1	32,447	$9.46	2,903,733.95	$12.24	10.82	351,209.57	N/A
2	81,150	$13.44	14,658,416.64	$16.86	1.77	143,546.24	N/A
3	20,412	$64.65	85,314,454.47	$15.48	0.00	51.03	N/A
Aggregate	134,009	N/A	102,876,605.06[a]	$15.53	N/A	494,806.84[b]	27.8

[a]*within group sum of square*
[b]*between group sum of squares*

Example column calculations in Table 4:

(1) $9.46^2 * 32,447 = 2,903,733.95$

(2) $(12.24 – 15.53)^2 = 10.82$

(3) $10.82 * 32,447 = 351,209.57$

(4) $\sqrt{\frac{102,876,605.06a + 494,806.84b}{134,009}}$ = 27.8 (standard deviation of aggregate record)

[1] At this time, all PUFs are available free of charge to the user community. PUFs may be downloaded online by completing a survey form: https://survey.vovici.com/se/56206EE333F13F0F.

[2] Prescription drug data for 2015 are not currently available in the PUFs.

[3] Copyright ©2022 Wolters Kluwer Clinical Drug Information Inc. Publication of research findings or reference to Medi-Span in these PUFs does not constitute an endorsement by Wolters Kluwer Clinical Drug Information Inc.

[4] As noted earlier, all paid amount variables in the PUFs represent pre-rebate transactional payments.

Minnesota Department of Health
Health Economics Program
PO Box 64882
St. Paul MN 55164-0882
(651) 201-4520
www.health.state.mn.us/healtheconomics