DEPARTMENT OF HEALTH

Minnesota All Payer Claims Database Provider Specialty Public Use File: A User Guide

SEPTEMBER 2026

Table of contents

Minnesota All Payer Claims Database Provider Specialty Public Use File: A User Guide 1

Background 2

Public Use File Overview 2

Design of the Public Use Files 2

Definition of a Provider Specialty 2

Data Elements 3

Exclusions from the Public Use File 4

Other Important Data Considerations 4

Appendix A: Interpreting PUF Data 6

Interpreting the Data Across Years 7

Appendix B: User Calculations 9

Aggregating Records 9

Counts and Dollar Amounts 9

Means 9

Medians 10

Standard Deviations 10

MN APCD
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, and prescription drugs.[1] To aid in the study of medical spending by provider specialty in Minnesota, MDH has prepared a PUF derived from medical professional claims. This document introduces the PUF, illustrates how to interpret PUF records, and includes technical instructions for users who wish to further aggregate PUF records.

Public use file overview

The provider specialty PUF was derived from medical professional claims filed by insurers. Each record in the PUF aggregates payments and procedures from professional claims (e.g., consultation, examination, surgery) associated with a provider medical specialty and an additional set of stratifying variables. PUF records may encompass a variety of procedures. The PUF does not include facility claims or pharmacy claims, but it does contain prescription drug costs to the extent that they exist within medical claims for professional services. The provider specialty PUF can be used to study variation in payments to providers by individual specialty and within or across combinations of additional stratifying variables, including payer type.

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

Design of the public use files

Definition of a provider specialty

Medical professional specialties were identified using Medicare Provider Enrollment, Chain, and Ownership System (PECOS) specialty codes, which are assigned to medical professionals who apply to enroll as providers under the Medicare program. PECOS divides medical professionals into practitioners and those who provide only ordering and referring services. The specialties included in the PUF are restricted to practitioners. For this group, PECOS identifies 86 specialties. A list of all 86 PECOS practitioner specialties is provided in the “PECOS Specialties” tab of the PUF. Specialties that do not appear in the PUF were redacted due to small cell size.

The majority of the claims included in the PUF had providers whose specialties were obtained from the PECOS database. The PECOs database is updated weekly. Practitioners who do not enroll as Medicare providers do not appear in the PECOS database. In order to include these practitioners in the PUF, their specialties were obtained from an alternative registry, the National Plan and Provider Enumeration System (NPPES), which identifies the specialties that practitioners reported in their applications for NPIs (may be less current than the PECOS data).[2]

The NPPES classification is more detailed than the PECOS classification, with each NPPES code being defined by the combination of a group, a classification, and in most but not all cases a specialization. Frequently, the combination of a NPPES group and classification corresponds to a single PECOS specialty, with the NPPES specialization representing the equivalent of a sub-specialty not distinguished in PECOS. To incorporate practitioners with this alternative system of specialties, Mathematica developed a mapping of NPPES specialties into the corresponding PECOS specialties. In all, 333 NPPES specialties mapped into PECOS specialties. A copy of this map is provided in the NPPES Map tab in the PUF.

Data elements

PUF records for each specialty are further stratified by:

- Payer type (commercial, Medicare, or Minnesota Health Care Programs)
- Site of service
- Patient's resource utilization band (RUB)
- Rural-urban commuting area (RUCA) classification of the provider's ZIP code
- RUCA classification of the patient's ZIP code

Site of service distinguishes among five locations based on place of service codes: (1) provider's office, clinic, or urgent care facility, (2) hospital or surgery center, (3) emergency room, (4) home or rehabilitation facility, and (5) any other practice setting. The patient's RUB is an indicator of the patient's utilization of medical resources during the year. Developed by Johns Hopkins University and produced with their Adjusted Clinical Groups (ACG®) software, RUB distinguishes among five levels of utilization, ranging from healthy users (1) to very high users (5). The RUCA classification is based on a coding of census tracts by their location and the commuting patterns of their residents. RUCA codes prepared by the Economic Research Service of the U.S. Department of Agriculture were mapped to Minnesota ZIP codes by Mathematica, and ZIP codes were then classified using a scheme recommended by the Washington State Department of Health as: (1) urban core, (2) suburban, (3) micropolitan, and (4) rural/small town. Providers with out-of-state ZIP codes were assigned a RUCA class of 5. Thus, each PUF record represents the claims associated with the same PECOS specialty, one of three payer types, and a specific combination of the four additional stratifiers.

The contents of the PUF are described in a data dictionary that appears as a tab in the PUF. To reduce the granularity of the data, users can aggregate PUF records in particular ways. For

example, sums of dollars paid and counts of procedures performed can be added across any combination of records while counts of unique providers have very limited additivity, as a given provider is likely to contribute to multiple records. For aggregation guidance, see Appendix B.

Exclusions from the public use file

The provider specialty PUF was generated from claims for professional services rendered by in-state or out-of-state practitioners to Minnesota residents. Facility, or institutional, claims were not included. Claims were dropped from PUF due to any of the following reasons: duplicate or denied status, identification as a facility claim, missing payer information, negative reported amount paid by the insurer or member, out-of-state patient residence, provider NPI that could not be matched to a PECOS or NPPES record, excluded specialty, site of service location coded as not applicable to a professional claim, patient ID that could not be matched to the member file for the purpose of assigning the patient's RUB, and provider or patient ZIP code that could not be matched to a census tract for the purpose of assigning a RUCA code.

Each claim for professional services corresponds to a procedure performed by the provider—for example, a consultation, examination, or surgery. After claims were aggregated to produce a preliminary, pre-redacted version of the PUF, records representing fewer than 11 unique providers or fewer than 11 unique patients were redacted to prevent identification of individual providers or patients. Approximately 59% of claims and 74% of total paid amount from the MN APCD are redacted from the Provider Specialty PUF.

Other important data considerations

The MN APCD includes medical and pharmacy claims for Medicare, Minnesota Health Care Programs, and most commercial plans. The MN APCD was not designed to include claims for healthcare covered by Tricare, Veterans Affairs, the Indian Health Service, Workers' Compensation, or for care provided to Minnesotans without health insurance. It also does not include claims for services provided by plans that do not cover general medical care, such as accident-only, vision, or dental plans. In addition, data from certain low-volume carriers (less than $3 million in medical claims or less than $300,000 in pharmacy claims) are exempt from submission to the MN APCD. Lastly, it should be noted that claims data are only as accurate as the coding on submitted claims.

In a decision released on March 1, 2016, the U.S. Supreme Court upheld a lower court's ruling that self-insured health plans could not be required to submit claims data to a state's APCD (Gobeille v. Liberty Mutual Insurance Co.). The court found that requiring self-insured plans to submit medical and pharmacy claims was preempted by the Employee Retirement Income Security Act (ERISA). The decision does not prohibit the voluntary submission of self-insured plan data to the MN APCD. The effect of this decision was to substantially reduce the volume of commercial claims and enrollment that ERISA-subject self-insured plans reported to the MN APCD. Summing commercial counts and costs in the PUF would therefore result in a considerable underestimate of use and spending across the whole commercial market. The

calculation of averages and medians are not expected to be materially impacted by the reduction in the data volume.

Appendix A: Interpreting PUF data

The following tables show subsets of data from the PUF to illustrate how to interpret key data elements. Table 1 reports mean provider charges and payments for procedures performed (or services billed) by cardiologists in an urban office setting, by payer, for patients with a RUB of 3 living in an urban core. Each row in the table corresponds to a subset of the columns in a single record in the PUF. The entries in the first row indicate the number of unique providers (for example, 341 providers in 2022) and the total number of procedures (26,675 procedures in 2022) for patients covered by commercial insurers. Continuing this example, providers serving commercial members charged an average of $301.59 per procedure in 2022 and were reimbursed an average of $224.87—including $143.29 paid by commercial insurers and $81.58 paid by commercial members.

Similar numbers of providers performed procedures for patients covered by commercial insurers and Medicare, with fewer providers serving patients in Minnesota Health Care Programs. Nevertheless, it is likely that many of the providers were the same across the three payers. Compared with procedures covered by commercial insurers, providers performed many more procedures covered by Medicare in all years, but only about a third as many procedures covered by Minnesota Health Care Programs. When interpreting the data, note the previously described data considerations regarding the impact of Gobeille v. Liberty Mutual Insurance Co. on commercial claims.

The charges per procedure did not differ greatly across the three payers, but providers received much lower average total (insurer plus member) payments from the public insurers than the commercial insurers—for example, in 2022, $143.29 from commercial insurers, $55.35 from Medicare, and $49.74 from Minnesota Health Care Programs. Differences in the mean amounts paid by members are more striking than the differences in the amounts paid by their insurers—for example, ranging from $2.66 among Minnesota Health Care Program enrollees in 2022 to $81.58 among commercial insurance enrollees.

Table 1. Mean 2022 payments for procedures performed by cardiovascular disease specialists in an urban office setting for patients with RUB code 3 and urban residence, by payer type.

Unique providers	Payer	Number of procedures	Mean charge	Mean insurer paid	Mean member paid	Mean total paid
341	Commercial	26,675	$301.59	$143.29	$81.58	$224.87
354	Medicare	42,582	$264.99	$55.35	$16.34	$71.70
309	Minnesota Health Care Programs	9,768	$272.37	$49.74	$2.66	$52.41

Table 2 reports mean 2022 provider charges and payments by payer for procedures performed by orthopedic surgeons in a hospital or surgery center, by providers with an urban location, and for patients with a RUB of 3 living in an urban area. Data in the rows indicate the number of unique providers, the number of procedures performed, the mean amounts charged, the mean amounts paid by insurers and members, and the mean total amount paid. For example, in 2022, 360 providers performed a total of 11,037 procedures for patients with commercial insurance coverage; they charged an average of $4,125.83, for which they were paid an average of $2,868.65 per procedure. Insurers paid on average $2,571.21 of this amount, with members paying an average of $297.43.

Table 2. Mean 2022 payments for procedures performed by orthopedic surgery specialists in a hospital or surgery center: urban providers for patients with RUB code 3 and urban residence, by payer type.

Unique providers	Payer	Number of procedures	Mean charge	Mean insurer paid	Mean member paid	Mean total paid
360	Commercial	11,037	$4,125.83	$2,571.21	$297.43	$2,868.65
317	Medicare	8,133	$3,061.54	$589.89	$52.40	$642.30
330	Minnesota Health Care Programs	6,825	$1,709.36	$269.43	$0.83	$270.26

Charges, member payments, insurer payments, and total payments were substantially less for Minnesota residents in Minnesota Health Care Programs or Medicare, compared with those amounts for commercial patients. For example, in 2022, providers were paid an average of $270.26 per procedure provided to patients in Minnesota Health Care Programs, $642.30 per procedure provided to patients in Medicare, and $2,868.65 per procedure provided to patients with commercial insurance. Patients in Minnesota Health Care Programs paid an average of $0.83 per procedure, while Medicare patients paid an average of $52.40, and commercially insured patients paid $297.43. The mix of procedures likely differed across the payers (reflecting, in particular, differences in the average age of enrollees in Medicare versus other payer types), contributing to the differences in both charges and payments.

Interpreting the Data Across Years

Table 3 reports the number of procedures and mean amounts paid per procedure by commercial insurers for the same records as shown in Tables 6, including records from both 2021 and 2022. Comparing the calculated mean amounts across years is straightforward. For example, commercial insurers and members paid a mean amount of $2,812.46 per procedure in 2021, for 11,677 procedures. In 2022, commercial insurers and members paid 2.0% more per procedure ($2,868.65) for 5.5% fewer procedures (11,037), compared with 2021.

Table 3. 2021-2022 percent change in the number of procedures and mean payment per procedure performed by orthopedic surgeons in a hospital or surgery center: urban physicians for patients with RUB code 3 and urban residence.

Payer	2021 number of procedures	2021 mean total paid	2022 number of procedures	2022 mean total paid	2021 to 2022 change in number of procedures (%)	2021 to 2022 change in mean total paid (%)
Commercial	11,677	$2,812.46	11,037	$2,868.65	-5.5%	2.0%
Medicare	7,632	$642.47	8,133	$642.30	6.6%	0.0%
Minnesota Health Care Programs	7,398	$247.86	6,825	$270.26	-7.1%	9.0%

Appendix B: User calculations

Aggregating records

Users may wish to construct totals, means, or other statistics across payer type or across one or more of the other stratifiers. Aggregation methods vary by type of statistic.

Counts and Dollar Amounts

Counts of procedures performed and amounts of dollars paid (that is, any of the four variables CHARGE_AMT_SUM, INSURER_AMT_SUM, MEMBER_PAID_AMT_SUM, and TOTAL_PAID_AMT_SUM) are additive. Each of these variables can be summed across any number or combination of records. There is no duplication in these quantities across records. There is also no duplication across specialties in the counts of unique number of providers with the same combination of all the other stratifiers (that is, payer type, patient RUB, provider RUCA class, and patient RUCA class). Within a given combination of stratifiers, for example, one can sum the number of unique providers across all specialties and obtain an unduplicated count of providers *within that combination of stratifiers*.

Counts of unique providers within the *same* specialty are *not* additive across different categories of stratifiers, however. This is because a given provider can see patients with different types of payers, different RUB classes, and different RUCA classes. Depending on the specialty, a provider may also see patients in more than one type of site. It is even possible that a provider might bill from more than one location having different provider RUCA classes. Therefore, we do not recommend summing the counts of unique providers within the same specialty across records of any type, as the sums are almost certain to overstate the true number of unique providers and may do so by several times over, depending on what records are summed.

Means

When records in the PUF are aggregated, the mean of the aggregate record (or the grand mean for this set of records) can be calculated as the weighted average of the means of the individual records, where the weights are the numbers of procedures performed. A more direct calculation is to sum the corresponding totals from which the means were calculated and divide this sum by the sum of the procedures performed.[3] This calculation is illustrated in Table 4 using data from the three records in Table 1. This represents an aggregation of records over payer type within a single combination of specialty, site of service, patient RUB, provider RUCA, and patient RUCA.

Table 4. Calculation of the mean of an aggregate of records.

Record	Number of procedures	Mean total paid	Sum total paid	Mean of aggregate (grand mean)*
1	26,675	$2,868.65	$5,998,665.74	N/A
2	42,582	$642.30	$3,053,303.48	N/A
3	9,768	$270.26	$511,951.13	N/A
Sum	79,025	N/A	$9,563,920.35	$121.02

*Sum of *sum total paid* divided by sum of *number of procedures*

Medians

One cannot determine the exact median of a statistic without access to the underlying microdata (in this case the individual claims). Unlike means, the median or weighted median of a set of subgroup medians (for example, the medians of amounts paid by payer type) is not the median of the overall group (that is, the median of the amounts paid across all three payer types). However, with a very large number of subgroups and none of them substantially larger than the others, the weighted median of the subgroup medians provides a good approximation of the median of the overall group. One can apply a calculation similar to the one illustrated in Table 4 to obtain the approximate median for an aggregate of PUF records (Table 5).

Table 5. Calculation of the approximate median of an aggregate of records.

Record	Number of procedures	Median total paid	Product*	Approximate median of aggregate (grand median)**
1	26,675	$80.40	$2,144,670.00	N/A
2	42,582	$26.25	$1,117,777.50	N/A
3	9,768	$12.65	$123,565.20	N/A
Sum	79,025	N/A	$3,386,012.70	$42.85

**Number of procedures* multiplied by *median total paid*

**Sum of *product* divided by sum of *number of procedures*

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 4 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 the standard deviation for an aggregate of records.

Record	Number of procedures	Standard deviation of total amount	(1)	Mean total paid	(2)	(3)	(4)
1	26,675	$397.23	4,209,092,874.61	$224.87	$10,784.82	287,685,140	N/A
2	42,582	$100.91	433,605,186.15	$71.70	$2,432.46	103,579,113	N/A
3	9,768	$81.66	65,136,497.50	$52.41	$4,707,33	45,981,219	N/A
Sum	79,025	N/A	4,707,834,558.26[a]	N/A	N/A	437,245,474[b]	255.16

[a] Within group sum of squares

[b] Between group sum of squares

Example column calculations in Table 6:

(1) 4,209,092,874.61 = 26,675 * $(397.23)^2$

(2) 10,784.82 = $(224.87 – 121.02)^2$

(3) 287,685,140 = 10,784.82 * 26,675

(4) 165.63 = $\sqrt{(4{,}707{,}834{,}558.26 + 437{,}245{,}474)/79{,}025}$

[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] PECOS was chosen as a primary source over NPPES because the specialties reported in PECOS can be more current than those captured in NPPES.

[3] Note that the mean payment for a given record was computed by dividing the total payments by the number of procedures. Because the mean payment was rounded to two decimal places, however, multiplying the mean payment amount by the number of procedures does not reproduce the total paid amount exactly.

Minnesota Department of Health
Health Economics Program
PO Box 64882
St. Paul MN 55164-0882
(651) 201-3550
www.health.state.mn.us/healtheconomics

MN APCD
All Payer Claims Database

DEPARTMENT OF HEALTH