Texas All-Payor Claims Database

Research Accessible Dataset

Center for Health Care Data

Release Date: 2026-01-21 | Data Availability: 2019-01-01 – 2024-12-31

Last updated: January 21, 2026

UTHealth Houston
School of Public Health
Center for Health Care Data

Contents

1200 Pressler St.
Houston, TX 77030

Fax: 713.500.9493
CHCD@uth.tmc.edu
go.uth.edu/chcd

About the Texas All-Payor Claims Database

In September 2021, House Bill (HB) 2090—a health cost transparency law passed in the 87th Legislative session—established the Texas All-Payor Claims Database (TX-APCD) within The University of Texas Health Science Center at Houston (UTHealth Houston) and UTHealth Houston School of Public Health Center for Health Care Data (CHCD).

The TX-APCD includes medical, pharmacy, and dental insurance administrative claims data, as well as enrollment and provider files collected from commercial payors, government and municipal benefit plans (such as Employees' Retirement System), Medicaid Fee-For-Service, and Medicaid managed care organizations. HB 2090 requires insurers regulated by the Texas Department of Insurance (TDI) to submit claims data to the TX-APCD monthly beginning in 2023, including retroactively submitting historical claims from January 1, 2019. Centers for Medicare and Medicaid Services (CMS) Medicare Fee-For-Service and Employee Retirement Income Security Act (ERISA) employer self-funded health benefit plans are not subject to data submission mandates. The TX-APCD contains data from over 120 distinct payors and represents approximately 60% of Texans with medical coverage (80% of Texans if any medical, pharmacy, or dental coverage is considered). Data are submitted monthly with a one-month lag by insurers in data formats specified in the TX-APCD Common Data Layout (CDL), which is modeled after the APCD-CDL© established by the National Association of Health Data Organizations.

Administrative claims data capture information generated for the purposes of insurer reimbursement and include diagnoses, procedures, dates, costs, and other relevant information along with patient and provider demographics. Claims that were rejected in the adjudication process due to incompleteness, inaccuracy, or other administrative reasons must be included in the data submission. Duplicate claims or claims not related to an enrolled member may be excluded from submission.

Member enrollment data are also submitted by insurers for members covered for one or more days of the submitted month, whether the member utilized services during the reporting period or not. Enrollment data includes demographic information for each individual as well as information on the member's coverage. A provider file including National Provider Identifiers (NPI), specialty codes, and practice location is submitted for every provider in the insurer's provider network or every provider for whom claims were adjudicated for that submission period.

Research Accessible Dataset

The Texas All-Payor Claims Database Research Accessible Dataset (APCD-RAD) consists of administrative claims that have been cleaned and aggregated into a functional dataset. This dataset is maintained and distributed by the CHCD for use by qualified research entities, which include healthcare providers in Texas, 501(c)(3) organizations engaging in public-interest research of health care delivery, and institutions of higher education. The APCD-RAD is provided on a lag that reflects the amount of time required to receive, validate, and correct data where necessary, as well as time to process and clean received data. Currently, data is available for January 1, 2019 through December 31, 2024 for the following tables:

- Enrollment tables
  - Annual_enroll
  - Monthly_enroll
- Administrative Claims tables
  - Dental
  - Medical
  - Pharmacy
- Provider table

More specific information on the data contained within these tables can be found in the ‘Dataset Tables and Descriptions’ and ‘Dataset Table Usage Notes’ sections of this document.

Identifiability Considerations

Information in the APCD-RAD has been deidentified by removing Protected Health Information (PHI) such as names, dates of birth, dates of death, social security numbers (SSNs), addresses, original member IDs, and original claim IDs. A RAD-unique system-generated person ID tracks individuals across all plans and carriers and replaces personal identifiers (Additional information on how person IDs are assigned can be found in section pers_id under Dataset Table Usage Notes). Similarly, a RAD-unique system-generated claim ID replaces carrier-assigned claim IDs. Dates of birth are replaced by age at end of calendar year. Individual person addresses are not available for research use, but higher-level information such as ZIP code and county FIPS code are available pending additional justification and approval.

Variables containing PHI or other sensitive data require additional justification and approvals to be included in a data extract.

Accessing the APCD-RAD

Access to the APCD-RAD is only granted to "qualified research entities," which include healthcare providers in Texas, 501(c)(3) organizations engaging in public-interest research of health care delivery, and institutions of higher education. For further information on how “qualified research entities” are defined as well as other helpful information, please review the FAQs for TX-APCD Data Requestors on the CHCD TX-APCD website (https://sph.uth.edu/research/centers/center-for-health-care-data/texas-all-payor-claims-database/frequently-asked-questions/).

For those who qualify, data access is granted on a secure UTHealth server where data management and statistical software programs are provided. Analytics must be completed within the secure servers—**NO DATA CAN BE DOWNLOADED, EXPORTED, OR COPIED**. Security monitoring software is installed on all UTHealth servers to enforce this mandate. If a researcher requires a specific program outside of the pre-installed list, let the CHCD team know during the data request process so that the required software can be installed. Please see the Data Request page on the CHCD website for more details (https://sph.uth.edu/research/centers/center-for-health-care-data/data-request).

Inclusions/Exclusions

APCD-RAD includes all medical, pharmacy, and dental administrative claims data as well as corresponding enrollment and provider information from all payors operating within the state of Texas with the following exceptions:

- ERISA plans—employer self-funded benefit plans governed by the Employee Retirement Income Security Act of 1974—are not required to report to TX-APCD, although some information is voluntarily contributed.
- Medicare Fee-For-Service is not regulated by state law and does not submit data to TX-APCD
- The APCD-RAD includes only the most recent version of a claim - interim/partially adjudicated claim versions are excluded from APCD-RAD.
- Voided, reversed, and duplicated claim lines are also excluded. Denied claims are retained with an indication of denied status.
- Services that bypass insurance—i.e. paid completely out-of-pocket by patients—are not captured by the TX-APCD.

Dataset Tables and Descriptions

Enrollment Tables

Annual_enroll

The *Annual_enroll* table contains a single record per person per calendar year. The table includes information on member demographics (see note about race/ethnicity reporting in 'Dataset Table Usage Notes'), primary and secondary medical coverage types and designs, and number of months with medical, pharmacy, dental, and supplementary medical insurance benefits or other insurance types, and an indicator for high-deductible plans.

Monthly_enroll

The *Monthly_enroll* table contains a single record per person per month. The table includes information on member demographics (see note about race/ethnicity reporting in 'Dataset Table Usage Notes'), primary and secondary medical coverage types and designs, pharmacy, dental, vision, supplementary medical insurance benefits or other insurance types, and an indicator for high-deductible plans.

Claims Tables

Dental

The *Dental* table includes line-level data on dental claims that were billed to insurance. The table includes Common Dental Terminology (CDT) codes, location within the oral cavity or specific tooth, costs associated with that claim line, and provider information. NOTE: Although a diagnosis code column exists in this table, dental claims are not required to include a diagnosis and so this column is largely null.

Medical

The *Medical* table includes line-level data on medical claims that were billed to insurance. This covers professional and institutional claims and encounters that can occur in hospital or office settings. Claims can encompass physician visits, hospitals, laboratory and diagnostic services, skilled nursing facilities, home health, hospice care, durable medical equipment, and more. Each claim line includes variables such as ICD-CM diagnosis codes, procedure codes (e.g. Healthcare Common Procedural Coding System (HCPCS) codes, Current Procedural Terminology (CPT) and their modifiers, ICD-PCS), service units, bill type, place of service, revenue code, dates of service, costs associated with that claim line, claim status, network indicator, and provider information. NOTE: professional claims (billed by healthcare professionals) and institutional claims (billed by facilities) are submitted to insurance using different forms (typically CMS-1500/837P for professional and UB-04/837I for facility claims). As these forms have different fields, data fields populated in APCD-RAD will also differ depending on claim type.

Pharmacy

The *Pharmacy* table contains claims data where each row represents one fill instance of one drug or medical supply billed to insurance. A claim is filed when a prescription is filled, not at the time of a physician's prescription. The table contains variables such as fill date, refill number, National Drug Code (NDC), generic indicator, compounding indicator, formulary indicator (indicates if drug is on insurer's list of approved drugs), quantity dispensed, number of days supply, dispensing method, costs associated with that claim, network indicator, claim status, prescriber, network, and pharmacy information.

Provider Table

Provider

The *Provider* table includes provider-level information for all in-network providers, regardless of whether they submitted any claims to the insurance, and out-of-network providers who submitted claims for that month. The information contains National Provider Identifier (NPI), provider name, address, up to 5 specialties, and entity type (person or group).

Dataset Table Usage Notes

Enrollment

APCD requires reporting of all covered members, including out-of-state members (such as college-age dependents living out of state) and members who do not incur any claims during a given timeframe (a.k.a. "silent" members). However, from January 1, 2019 to February 28, 2023, the collected data is "historical data" compiled from submitters who may have had to retroactively reconstruct data to meet APCD reporting requirements. Comparison of enrollment data counts suggest that there may be missing enrollment records during the historical time period.

The race/ethnicity fields in enrollment tables are NOT required fields in TX-APCD and are therefore frequently null. Estimates from 2024 show that race and ethnicity are missing in over 50% of member records from commercial payors. Population of these fields is also non-random – please use race/ethnicity data cautiously.

Pers_id

Pers_id (individual person identifier) is a derived identifier created by CHCD in order to determine who is “an individual.” This identifier remains constant for a person across medical, pharmacy, and dental tables, across time, across addresses, and across different insurance plans and data submitters. It is constructed by comparing identifiers such as full name, date of birth, social security number, address, and other useful fields across all submitted data, even for cases where a claim line does not have a corresponding enrollment record. The algorithm employs a search methodology such that limited misspellings and/or name changes (such as last name changes due to marriages/divorces), address changes, and typos in other fields will still result in correct matches. Efforts have been made to match newborns (e.g. Baby Girl Smith) to their named records, although we should note that sometimes newborns are billed under the maternal health plan using the mother’s member ID—in these cases, we are unable to clearly differentiate maternal claim rows from pediatric claim rows.

Where there is ambiguity, the CHCD chooses to err or the side of fewer, more distinct matches. If we are unable to determine if two rows belong to two different individuals or a single individual, we choose to believe it is two individuals. Additionally, there are a few edge cases where two individuals may be confused for a single individual. An example would be similarly named twins (e.g. Tim Jones and Tom Jones, sharing the same date of birth) residing in the same household who either do not have social security numbers on record (payors are not required to report SSN) or have had a social security number error on some administrative form. These cases are rare, and should not have an outsized impact on most research processes.

Cohort Selection

Cohort selection is a very important part of any research project. For example, when determining who is going to be in a cohort, you may want to exclude people who only have dental enrollment records (17% of 2024 enrollment data), or people who only have medical but no pharmacy records (10% of 2024 enrollment data) if prescription drugs are a determining factor in analysis, or alternatively people with only pharmacy records and no medical enrollment. Additionally, there are many healthcare plans that do not offer full coverage – e.g. Medicaid CHIP perinatal plans that only cover a select range of obstetric services and nothing else, or catastrophic plans that do not offer any coverage until a member meets a high deductible (typically over $10,000).

Note that someone who appears to have dental and/or pharmacy coverage only without medical coverage in the APCD-RAD does not imply that the person does not have medical insurance. More than half of employer-sponsored health plans are self-funded and are not mandated to report to the Texas APCD. Over 40% of Medicare plans are Medicare Fee-For-Service and also do not report to the Texas APCD. It is not only possible but likely that an individual who appears in the APCD-RAD as having "only dental" also has concurrent medical and/or pharmacy coverage through a non-reporting avenue.

2024 Estimates		
Insurance type	**Count of Persons**	**Percent**
Full medical and pharmacy coverage	12,620,000	59.5%
Dental coverage only	3,610,000	17.0%
Full medical coverage, no pharmacy coverage	1,960,000	9.3%
Pharmacy coverage only	1,140,000	5.4%
Dual Medicare/Medicaid eligible	840,000	4.0%
Pharmacy and/or dental coverage, no Medical	470,000	2.2%
Partial medical coverage	440,000	2.1%
Other	100,000	0.5%

Of people with full medical coverage, the types of medical coverage are as follows:

2024 Estimates, people with full medical insurance only		
Insurance type	**Count of Persons**	**Percent**
Commercial	8,140,000	54.1%
Medicaid	4,830,000	32.1%
Medicare Advantage	2,080,000	13.8%

See Appendix 1 for list of medical plans that count as "full coverage"

Dual Enrollment

INFORMATION ON DUAL-ELIGIBLE INDIVIDUALS AND CLAIMS ARE CURRENTLY NOT AVAILIABLE IN THIS DATA RELEASE. Dual enrollment refers to the case when an individual has both Medicare and Medicaid coverage. In cases when an individual has both Medicare Advantage (reports to APCD) and Medicaid, we would expect to see at least two distinct claims for each service with Medicare acting as the primary payor and Medicaid acting as the secondary payor. If the individual has Medicare FFS (does NOT report to APCD) and Medicaid, the dataset will contain only the claim submitted by Medicaid, and we expect to see paid amounts that reflect the typically smaller proportion paid for by a secondary payor. While services should be consistent regardless of what source the data is coming from, the sum of costs for services would differ significantly. These situations should be considered if a researcher is interested in calculating costs for dual-eligible individuals.

Primary/Secondary Medical Insurance

Similar to duals, some individuals will have both a primary medical insurance policy as well as a secondary (or more) medical insurance policy to cover costs and services not covered by the primary medical insurance policy. For individuals with more than one concurrent medical policy, it is possible to see more than one claim reference the same service. In such cases, it would be expected that the primary coverage would pay the major portion of the bill and report the secondary coverage responsibility as a Coordination of Benefit (COB) amount. The secondary carrier should report similar paid amounts but in different columns, with the amount paid by the primary carrier reported as Other Insurance Paid Amount.

Unlike the case of dual-enrollees, these individuals and both claims are retained in the current data release of APCD-RAD. Please take these cases into consideration when calculating costs or counting services so as not to overcount a single event that was billed on two claims.

Enrollment table aggregation logic

The monthly enrollment table (*Monthly_enroll*) is an aggregation of data from over 120 data submitters. Since enrollees can be covered by multiple payors for different types of insurance or have supplemental plans, it is often the case that the TX-APCD receives more than one row of enrollment data for an individual in a given month, generally from more than one data submitter. For example, someone may have a primary medical insurance policy, a prescription drug insurance policy, and a dental insurance policy. That would mean a minimum of 3 rows of data for that individual per month were submitted. In order to consolidate records, individual-level information is aggregated using a “most frequent, then most recent” algorithm. This means that for a given data period (month or year), the most frequent value is used. In the case of ties, the more recent value is selected. There are three scenarios where this logic is implemented differently.

1. In the case of individual-level information unlikely to change, the most common value across all received data is used. The variables that are processed this way are age, sex, race, ethnicity, and hispanic indicator.
2. Individual-level information that is expected to change more frequently is aggregated on a monthly basis. The variables that are processed this way are ZIP code, FIPS code, and state. Note that location information is processed as a block, and not as individual columns.
3. In the case of the yearly table (*Annual_enroll*), if there is a tie (say someone has a commercial primary medical insurance for 6 months and primary Medicare Advantage medical primary for 6 months), then the most recent value is used.

Aggregation logic deprioritizes nulls and non-informative values, such as ‘U’ for unknown. A person would only have missing or unknown information in an all-time column if no informative information has ever been submitted by any data submitter. Similarly, an individual would only

have missing or unknown information in a monthly column if no informative information is submitted from any payor for that month.

The type of insurance is determined using several columns from the TX-APCD CDL, including insurance/product category type, medical/pharmacy/dental/behavioral health coverage under this plan indicator columns, and primary insurance indicator. Once the plan type and plan design have been determined, a plan type hierarchy is used to determine the individual's primary health plan, secondary health plan, pharmacy plan, and so forth. This information is then consolidated onto a single row. Only the predominant medical plan type for a given year is shown on the annual enrollment table. If a person is enrolled in commercial insurance 4 months out of the year but has Medicare Advantage for the remaining 8 months, they will appear on the annual enrollment table as having Medicare Advantage as their predominant medical plan, with 12 months of medical coverage. In the case of ties, the more recent value is used.

Medical

The *medical* table includes information from roughly 70 different data submitters. TX-APCD data submission guidelines require submission of data for all interim stages of adjudication. Roughly 95% of medical claims are directly paid out upon the initial bill and do not require an extended adjudication process. However, the remaining 5% do, and the adjudication process can take place over months or years and can have many interim versions. The APCD-RAD has logic in place to determine the most "final" adjudicated version of a claim as of the time of processing. Denied claim lines, if they are part of the "final" version of a claim, are retained. Voided claims—typically claims that have some sort of error—are excluded from the APCD-RAD. Interim versions of claims are also excluded in favor of the final version

- Note that it is possible and expected for a single medical service to be represented on more than one claim. Some possible reasons are:
    - A medical service may have both a primary and a secondary payor (in which case both claims should be considered in analysis). These claims can be distinguished from "duplicate" claims as they may have a non-zero coordination of benefits (cob) column or non-zero other insurance paid amount (oth_ins_pd_amt) column and different values for plan_paid_amount.
    - The claim processor uses a third-party administrator (TPA) and both (or more) parties have submitted data for the same service.
    - In the case of "duplicate" claims due to multiple data processors not resulting from a primary/secondary coordination of benefits scenario, CHCD has made efforts to reconcile known instances. However, it is likely that some residual percentage of "duplicate" claims persist within the database.

Services that are paid completely out-of-pocket by the patient and are not billed to insurance are not captured in the TX-APCD.

Pharmacy

The *pharmacy* table includes information from roughly 60 different data submitters. Similar to the medical table, logic has been implemented to determine the "final" adjudicated version of a claim. Denied claim lines are retained, voided claim lines are excluded, and interim claim versions are also excluded. Additionally, claim lines with invalid NDC codes have also been excluded.

- Similar to the *medical* table, a single prescription fill may appear more than once in the pharmacy table.
    - A prescription fill may have both a primary and secondary payor (in which case both claims should be considered). These claims can be distinguished from “duplicate” claims as they may have a non-zero coordination of benefits (cob) column or non-zero other insurance paid amount (oth_ins_pd_amt) column and different values for plan_paid_amount.
    - More than one claim may be submitted for the same prescription from multiple data processors. In the pharmacy space, in additional to third-party administrators (TPAs), we may also receive “duplicate” claims from a pharmacy benefit manager (PBM).
    - In the case of “duplicate” claims due to multiple data processors not resulting from a primary/secondary coordination of benefits scenario, CHCD has made efforts to reconcile known instances. However, it is likely that some residual percentage of “duplicate” claims persist within the database.

Dental

The *dental* table includes information from roughly 60 different data submitters. Similar to the medical table, logic has been implemented to determine the “final” adjudicated version of a claim. Denied claim lines are retained, and interim claim versions are also excluded in favor of the final version

Provider

The *provider* table is an aggregation of all provider data submitted by all payors. The raw provider data is highly repetitive, with limited variation on a month-to-month basis. The APCD-RAD provider table selects only distinct values from the TX-APCD raw data to reduce redundancy. Please note that provider information is provided “as-is” and is not necessarily confirmed, cross-checked, or verified in any way by data submitters as most fields are not relevant to claims processing. Notably, NPIs may sometimes appear to be shared between healthcare providers or differ in other ways from the database maintained by the National Plan and Provider Enumeration System (NPPES) due to any number of reasons including using attending or billing provider NPI in the rendering provider field or using the provider group NPI in place of a single provider’s. Caution is recommended when using data from this table.

User Compliance

The TX-APCD obtains data that contains Current Procedural Terminology (CPT) Codes owned and copyrighted by the American Medical Association (AMA), and created the RAD with such codes in place. The TX-APCD acknowledges that all rights in the CPT coding are owned and retained by AMA.

1. User is responsible for obtaining licensure from the American Medical Association (AMA) for use of the CPT and HCPC codes in research, reporting, and publication.
2. User is prohibited from making AMA CPT codes publicly available, creating derivative works (including translating), transferring, selling, leasing, licensing, or otherwise making available to any unauthorized party the Deliverables, or a copy or portion of AMA CPT.
3. User expressly acknowledges that use of the AMA CPT codes is at User’s sole risk and the AMA CPT is provided “as is” without warranty of any kind. The TX-APCD does not provide text descriptions of the codes, nor cover license fees for external users.
4. Users are required to keep records and submit reports, including information reasonably necessary for the verification of royalties paid to the AMA for each relevant research project, and shall provide such information to the TX-APCD upon request.

The user is responsible for ensuring compliance with all applicable laws and regulations applicable to data security and privacy. The TX-APCD does not provide legal advice.

User must acknowledge the TX-APCD as the source of data in any reports, publications, demonstrations, or other display of data findings. The following is the approved citation:

UTHealth Houston School of Public Health, Center for Health Care Data. (2025). Texas All-Payor Claims Database (TX-APCD). ___. Retrieved [year(s)].

Appendix 1: List of Medical plan designs that count as “full coverage”

Medical Plan Type(s)	Medical Plan Design	Full Coverage	Notes
Medicaid	CHIP	yes	
Medicaid	CHIP PERI	no	CHIP Perinatal only covers a select range of obstetric services for expectant/postpartum individuals
Commercial	Cobra	yes	
Commercial, Medicare Advantage	EPO	yes	
Commercial, Medicare Advantage, Medicaid	FFS	yes	
Commercial, Medicare Advantage	HMO	yes	
Medicaid	HTW	no	Healthy Texas Women (HTW) provides a variety of women’s health and family planning services
Medicare	Imputed	no	This designation is for individuals who are dual-eligible and have Medicaid enrollment records but no corresponding Medicare enrollment records. They may or may not have full coverage, but there is insufficient information to tell definitively
Commercial	POS	yes	

1200 Pressler St.
Houston, TX 77030

Fax: 713.500.9493
CHCD@uth.tmc.edu
go.uth.edu/chcd

Medical Plan Type(s)	Medical Plan Design	Full Coverage	Notes
Commercial, Medicare Advantage	PPO	yes	
Medicare Advantage	SNP	yes	Medicare Special Needs Plans (SNPs) provide benefits to people with specific severe and chronic diseases such as cancer, diabetes, and end-stage renal disease (ESRD). They offer the same coverage as Medicare A & B, but may also offer additional coverage, such as additional days in the hospital. Enrollment in a SNP is contingent on meeting (and continuing to meet) special conditions of the plan.
Medicaid	STAR	yes	
Medicaid	STAR Health	yes	
Medicaid	STAR Kids	yes	
Medicaid	STAR+PLUS	yes	
Commercial, Medicare Advantage, Medicaid	Unknown	no	These plans may or may not offer full coverage, but as more specific information is not provided, we cannot assume that they do.

1200 Pressler St.
Houston, TX 77030

Fax: 713.500.9493
CHCD@uth.tmc.edu
go.uth.edu/chcd