UTHealth Houston
School of Public Health
Center for Health Care Data

Guide to Health Care Administrative Claims Data

UTHealth Houston School of Public Health
Center for Health Care Data

Table of Contents

UTHealth Houston
School of Public Health
Center for Health Care Data

Purpose

This document provides a high-level overview of administrative health care claims data and serves as a general reference for qualified researchers requesting access to claims datasets licensed and administered by the UTHealth Houston School of Public Health Center for Health Care Data.

While an introduction to claims data is included, researchers who are new to using administrative claims for research purposes are strongly encouraged to complete a dedicated tutorial or formal course to gain the foundational knowledge necessary for effective and responsible use of claims data.

Overview of Health Care Administrative Claims Data

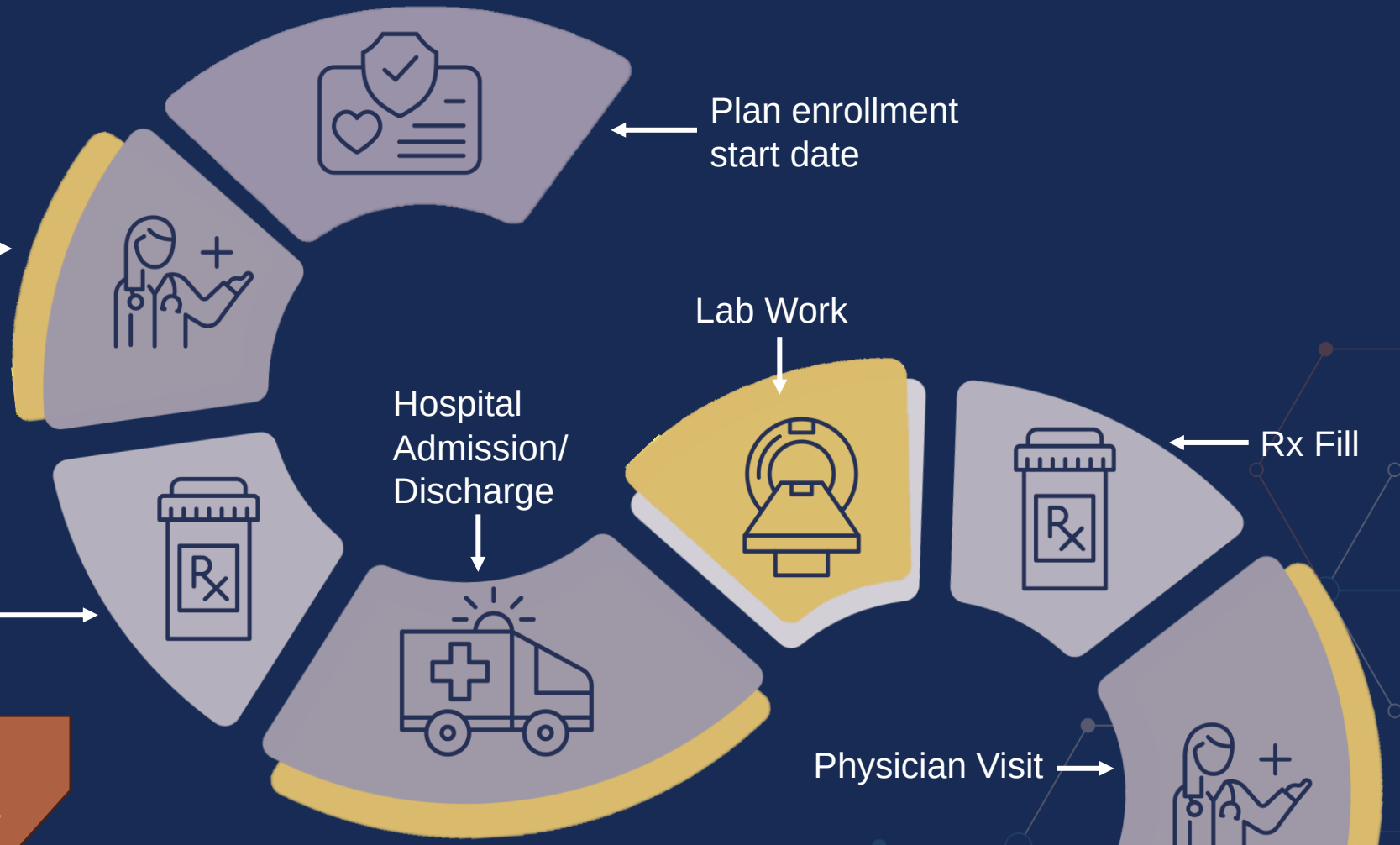
UTHealth Houston
School of Public Health
Center for Health Care Data

Healthcare Administrative Claims Data

A collection of data that is used for insurance billing purposes and captures information from all providers and payors.

Claims data can be used to study healthcare delivery, benefits, harms, and costs.

Tracking the Patient Journey with Claims Data



Yellow areas indicate where electronic health record (EHR) data can supplement claims data.

How Claims Data Adds Value

Study a large portion of the population

Learn more about patient procedures and diagnoses

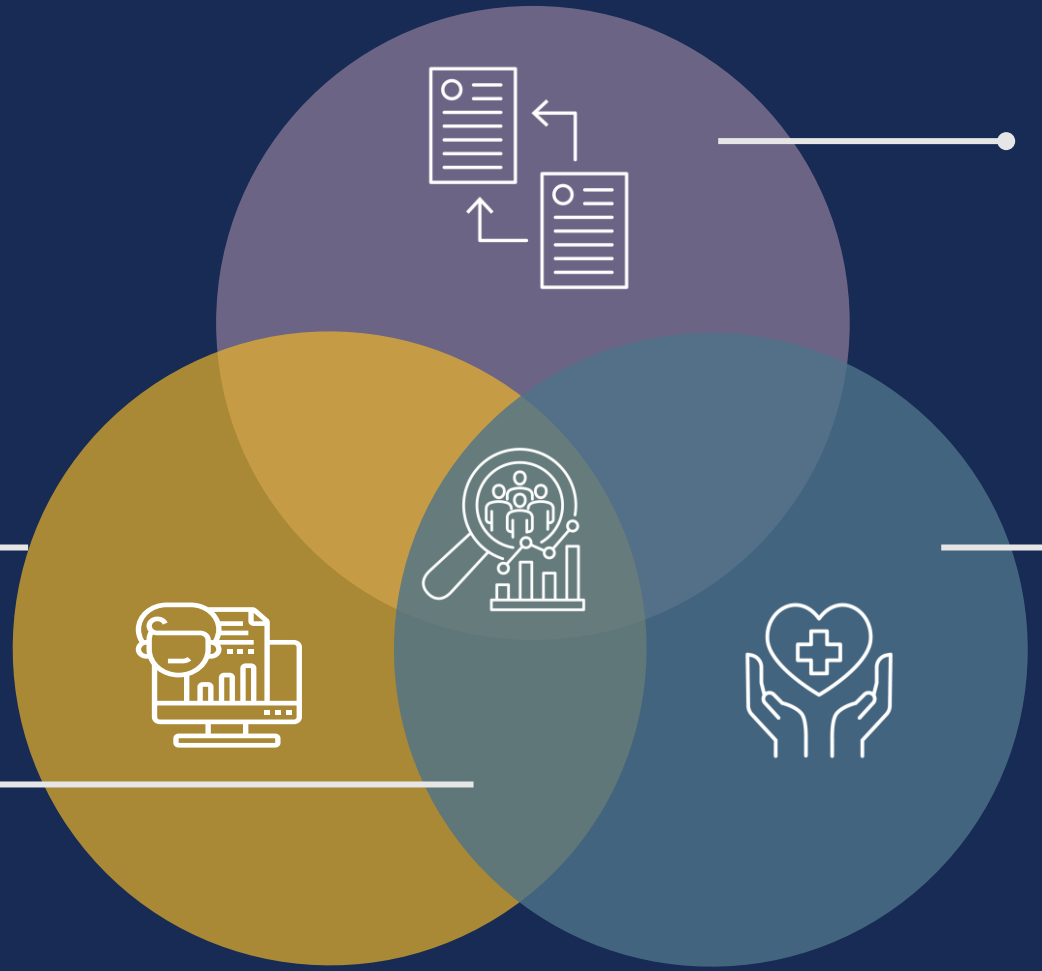
Follow insured patients over time

Use aggregated data on episodes of care (inpatient, ED visits, office visits, etc.)

Using Claims Data to Explore Diverse Health Topics

- Adolescent Health
- Chronic Disease
- Comorbidities
- Dental Health
- Healthcare Service Utilization
- Immunizations
- Mental Disorders
- Prescription Use
- Preventive Care
- Maternal Health
- Men’s Health
- Morbidity and Mortality
- Substance use disorder
- Treatment and Therapies
- Women’s Health
- And more!

Claims Data as a Complementary Data Source



Primary Data Collection

Electronic Health Records (EHR)

Patient Registry

Non-Medical Drivers of Health Datasets

Common Uses of Claims Data

Disease Surveillance

Data can be used to monitor disease prevalence, incidence, and health outcomes across regions in real time.

Predictive Analytics

Predictive analytics, along with statistical modeling, can be used to help identify potential risks, patterns of care, and more.

Program/Intervention Evaluation

Using data to examine the effectiveness of program delivery

Health Economics & Outcomes Research

An evaluation of services/interventions provided by healthcare companies to analyze costs, utilization, and health outcomes.

Elements in Claims Data for Analysis

UTHealth Houston
School of Public Health
Center for Health Care Data

Claims Data Tables

Claims data tables contain the following information:

- Member Enrollment
  - Patient and plan information
- Medical Claims
  - Includes inpatient, outpatient, home health, skilled nursing, professional, and others
- Pharmacy Claims
  - Filled prescriptions
- Dental Claims

Medical Claims

- Medical claims are submitted by providers for an encounter with a patient. This includes licensed healthcare professional (physicians, Physician Assistant, etc.) and facilities (hospitals, clinics, etc.)
- Information includes:
  - Patient Demographics, Diagnoses, and Procedures recorded at the time of encounter
  - Provider – Both physician and facility level. Billing and rendering providers.
  - Payment – Charged amount, paid amount, copay, etc.

UTHealth Houston
School of Public Health
Center for Health Care Data

What is in a medical claim?

- Patient ID
- Claim ID
- Provider (billing, rendering, etc.)
- Diagnosis codes
  - ICD10-CM
- Procedure codes
  - CPT/HCPCS or ICD10-PCS
- Modifiers and Units
- Place of Service
- Type of Bill
  - Facility Type Code (link) and Service Classification Type Code (link)
- Diagnosis Related Group (DRG):
  - All Patient Refined Diagnosis Related Group (APR DRG) (link)
  - Medicare Severity DRG (MS-DRG) by Center for Medicare and Medicaid Services (CMS) (link)
- Revenue codes
- Payment amounts
  - Allowed, paid, charged, etc.

*It is up to the data requestor to identify which codes are relevant for their research. We highly recommend performing literature reviews to develop comprehensive lists of diagnosis, procedure, and other relevant codes.

Pharmacy Claims

- Pharmacy claims are submitted by a pharmacy when a member fills the prescription.
- Drugs administered by a physician are coded in medical claims
  - Drugs given by Provider as IV Injections are coded in medical claims
  - HCPCS J Codes: Charges only cover the cost of the drug/product. Charges associated with the administration are not included
- Not all medical members have pharmacy (Rx) benefits
  - Medicaid – all enrollees get Rx benefits
  - Medicare – Requires enrollment in a Part D pharmacy plan

UTHealth Houston
School of Public Health
Center for Health Care Data

What is in a pharmacy claim?

- Patient ID
- Allowed amount
- Drug information:
  - Therapeutic Class, Generic vs. Brand Name Drug, Dose, Strength
  - National Drug Classification (NDC) Codes
- Dispensing information:
  - Date filled, New or Refill, Days Supply, Quantity Dispensed
- Pharmacy Information

UTHealth Houston
School of Public Health
Center for Health Care Data

National Drug Classification (NDC) Codes

- There are three segments to each NDC code
  - Labeler code – 4 or 5 digits – not necessarily the manufacturer – assigned by Food & Drug Administration (FDA)
  - Product code – 3 or 4 digits – assigned by labeler
  - Packaging code -2 or 1 digit – assigned by labeler

NDC	Labeler	Product	Package
00054003621	00054 Roxane Lab	0036 Clarithromycin 250mg tab	21 60 tablet bottle
0002515152031	00025 Pfizer US	1520 Celebrex 100 mg cap	31 100 capsule bottle

UTHealth Houston
School of Public Health
Center for Health Care Data

NDC Codes – Continued

- NDC are displayed on drug packaging in a 10-digit format: 4-4-2 or 5-3-2 or 5-4-1
- CMS: Uses NDC 11 (5-4-2) by adding a leading zero to comply with the format.

Converting NDCs from 10-digits to 11-digits					
10-Digit Format on Package	10-Digit Format on Example	11-Digit Format	11-Digit Format Example	Actual 10-digit NDC Example	11-Digit Conversion of Example
4-4-2	9999-9999-99	5-4-2	**0**9999-9999-99	0002-7597-01	**0**0002-7597-01
5-3-2	99999-999-99	5-4-2	99999-**0**999-99	50242-040-62	50242-**0**040-62
5-4-1	99999-9999-9	5-4-2	99999-9999-**0**9	60575-4112-1	60575-4112-**0**1

UTHealth Houston
School of Public Health
Center for Health Care Data

Dental Claims

Professional Claims

- Some dental procedures are medical claims – dental surgery, orthodontal, periodontal
- Dental claims use Current Dental Terminology (CDT) codes
- Claim fields are different from medical—for example, claim fields document which tooth received treatment

Common Payment Fields in Claims Data

- Charges – amount charged by the provider of services
- Excluded/Discount – not allowed or discounted via contractual arrangements
- Allowed: maximum amount plan will pay for covered health care service
  - Also called “eligible expense,” “payment allowance,” or “negotiated rate.”
  - If provider charges more than allowed amount, patient may have to pay difference
- Paid: Total payments paid to a provider for a service. Payment equals amount eligible for payment under medical plan terms after applying rules such as discounts, but **before** applying Coordination of Benefits (COB), Copayments, and Deductibles
- Out-of-pocket costs: deductibles, coinsurance, copayments for covered services plus all costs for services that aren't covered
- Coordination of Benefits (COB): a patient is covered under more than one insurance plan

What Claims Data Does Not Include

UTHealth Houston
School of Public Health
Center for Health Care Data

What is NOT in a medical claim

- Detailed clinical information
  - Lab values: Lab values are typically found in electronic health records - not claims data.
  - Patient vitals: Patient vitals such as height, weight, blood pressure, etc.
- Race/ethnicity (for commercial insurance providers)
  - Commercial insurance providers typically do not record information related to race and ethnicity.
- Nonmedical Drivers of Health
  - Claims data generally do not include information related to patients' socio-economic status, education, transportation, food insecurity, etc. However, there may be opportunities to leverage NMDOH databases alongside claims data to understand nonmedical drivers of health for patients.

NOT in a medical claim – continued

- Information on family members
  - Spouses and/or children of the insured may not be found in claims data, as they may be covered under separate insurance plans
  - Some plans are individual enrollment only
- Subjective information
  - Medical claims do not include healthcare providers' opinions, feelings, or beliefs about the patient's condition
- Premiums
  - Insurance premiums—the regular payments members make to keep coverage—are not included in medical claims.

What is NOT in a pharmacy claim

- Non-prescription items such as:
  - Over the counter medications
  - Food or dietary supplements
- Medication that was prescribed, but not filled by the patient
- Medication that was taken in a hospital (this will be in the medical claim)
- Lab results and diagnostic test results
- Subjective information (opinions, feelings, or beliefs about the patient's condition)
- Details on patients within the same household
- Prescription adherence

Insurance Types

UTHealth Houston
School of Public Health
Center for Health Care Data

Insurance Types

- **Commercial**
  - These include both HMO and PPO plans. Provide plans to individuals and groups, often through employers. Common carriers that offer commercial plans are Blue Cross Blue Shield, Aetna, United Health Care, etc. These include both HMO, POS and PPO plans.
  - Medicare Advantage (Part C): a Medicare-approved plan offered by private insurers and offers a bundled Medicare plan (Part A, B and often a prescription drug benefit). It may also include extra benefits such as dental, vision, etc.

Insurance Types - Continued

- **Government-sponsored plans**

  Medicare Fee For Service: This federal health insurance program is administered by the Centers for Medicare & Medicaid Services and includes Parts A (hospital) and B (medical), and Part D (when enrolled). This program is for:

  - those ages 65 and older
  - younger people with disabilities
  - People with End Stage Renal Disease or Amyotrophic Lateral Sclerosis (ALS)

  Medicaid: This federal health insurance program is administered by the state of Texas and provides free or low-cost health coverage. In Texas, coverage is for:

  - Low-income children and families (e.g. CHIP)
  - Pregnant women with qualifying incomes
  - People with disabilities
  - Seniors ages 65 and older with limited income

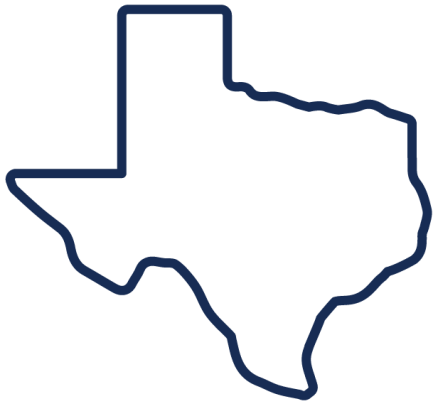
- Texas programs include STAR, STAR kids, STAR health, and STAR + PLUS.

Texas All-Payor Claims Database (TX-APCD)

UTHealth Houston
School of Public Health
Center for Health Care Data

Texas All-Payor Claims Database (TX-APCD)

The Texas All-Payor Claims Database is a large, statewide data system that identifies cost and quality drivers from billions of medical, pharmacy and dental claims from a wide range of payors — private insurance companies, Medicaid, Medicare, and some self-funded employer health plans.



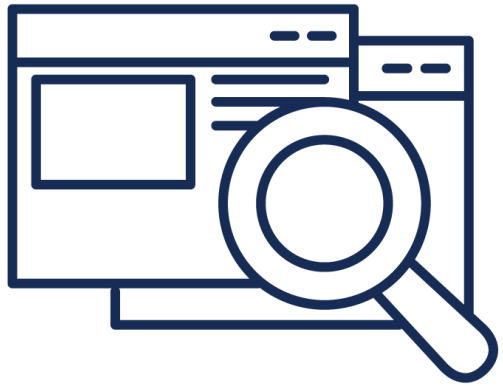
Its main purposes are to:

- Increase transparency in health care cost and quality
- Support evidence-based policymaking by giving lawmakers and agencies better insight into health spending, utilization, and trends
- Provide data access to qualified researchers to identify ways to improve population health, reduce costs, and address variations in care
- Inform consumers through public tools that compare costs and quality across providers

UTHealth Houston
School of Public Health
Center for Health Care Data

TX-APCD: What is NOT included

- Claims from persons residing outside of Texas
- Claims from persons insured under plans outside of Texas
- Claims from persons who are Texas residents and:
  - Are covered under Federal insurance plans; e.g.,
    - Veterans Administration
    - Tricare
    - Medicare Fee for Service (however, the CHCD provides this data from another resource to be used in conjunction with the TX-APCD)
  - Insured under an employer ERISA plan that opted to no participate
  - Uninsured individuals or individuals who choose to pay cash at the time of service



Who are TX-APCD Submitters?

The Texas Department of Insurance (TDI) requires **all payors** that are regulated by TDI to submit claims to TX-APCD.

- Medical plans
- Dental plans
- Behavioral Health plans
- Medicare Advantage plans
- Medicare Supplemental plans (voluntary)
- County and Municipal Sponsored Plans
- State Plans
- Medicaid Managed Care Organizations (MCO's)
- Medicaid fee for service
- Non-ERISA self-funded plans & Non-Preempted ERISA)100% of Medicare Fee for Service claims are available through CHCD

Monthly Submission Claims

- Enrollment Claims
- Provider Claims
- Medical Claims
- Dental Claims
- Pharmacy Claims

Planning Your Project Utilizing Claims Data

UTHealth Houston
School of Public Health
Center for Health Care Data

Using claims to answer your question

- Who is in my population and how can I identify them?
  - Inclusion and exclusion criteria
- What is considered an event or case?
- How much time should be covered?
  - You only need data over many years if you are: (1) studying a very rare condition/event and are concerned about sample size, or (2) specifically measuring trends over time.
  - Restricting to less time can make overall data management easier. Claims data can get very large, making programming and file storage efforts cumbersome and costly
- Am I interested in costs, and if so, which costs? Cost charged to insurance? Cost paid by patients? Amounts received by providers?

Prepare ahead of time

- If you have not frequently used claims for research, we recommend a literature review of your topic searching specifically for ways to identify your target population, event, case, or others within administrative claims.
- It can also be beneficial to collect information from clinicians or billing departments, if you work in a medical setting, for advice on which codes are commonly used.
- While the CHCD is well-versed in claims data, we are not medical experts and cannot tell you which diagnosis codes or procedure codes should be used for your study. **You are the Subject Matter Expert.**

What to Prepare

- Diagnosis codes
  - ICD-10-CM, APR DRG, MS DRG
- Procedure codes
  - CPT/HCPCS, ICD-10-PCS
  - Modifiers
- Type of bill, place of service, revenue codes
- Drug codes
  - NDC, therapeutic classes, generic names
- Provider codes
  - NPI, Specialty, NPI Taxonomy

Example 1 - Diabetes Prevalence

- Research question: What is the prevalence of diabetes in Texas?
  - **Who is in my population?**
    - Everyone? Adults 18+? Pregnant women? Children and adolescents? 65+?
    - Should anyone be excluded? Those with diseases that cause secondary diabetes? Cystic fibrosis, pancreatitis, corticosteroid use?
  - **Who counts as diabetic?**
    - Type1, 2, or gestational?
    - Is one instance of a diabetes diagnosis sufficient? Two diagnoses at least 30 days apart for increased positive predicted value? Should they also have a prescription for insulin?
  - **When?**
    - If this is a 1-year cross-sectional, should you only consider those with a full year of insurance coverage? Is less time of surveillance sufficient?
    - If someone is diagnosed in the year before your research year and is still alive the following year, should we still consider them as diabetic?

UTHealth Houston
School of Public Health
Center for Health Care Data

Example 2: Mammography with questions answered

- Research question: What was the screening mammography rate for Harris County in 2024?
  - **Who is in my population?**
    - Women, age 40-75, living in Harris County at the end of the measurement period, continuously enrolled in insurance for the entire measurement period, exclude those who were in long-term care, were frail or had an advanced illness, or had a previous bilateral mastectomy or two unilateral mastectomies (right and left)
  - **What counts as a mammogram?**
    - At least 1 mammogram identified using HCPCS/CPT procedure codes during the measurement period.
  - **When?**
    - Current recommendations are for every 2 years, so for the rate in 2024, we must include mammograms from 2023 & 2024.

Example 2 Mammography (cont.): Prepared Codes

- 538 ICD-10-CM Diagnosis codes identifying
  - Absence of left or right breast, advanced illnesses, frailty, history of mastectomy, hospice encounter
- 121 HCPCS/CPT/ICD-10-PCS Procedure codes
  - Mastectomy, frailty device, frailty encounter, hospice encounter, palliative care
- 14 Revenue codes for hospice
- 6 Mammography CPT codes

- **This is just to identify the numerator and denominator!!!**

For any questions regarding health care data, or to submit a data request, please contact our team for more information.

Thank you!

UTHealth Houston
School of Public Health
Center for Health Care Data