ASCENT
Advancing the Science of Palliative Care Across the Lifespan

Using Claims and Electronic Health Record Data in Palliative Care Research

What are claims and EHR data?

Claims and electronic health record (EHR) data are types of real-world data that are useful for palliative care research.

Claims data are based on bills submitted by health care providers to payers for payment. Claims data are informed by EHR data.

EHR data are the detailed information recorded by health care providers during a health care encounter.

The information included in each type of dataset is listed to the right.

Claims and EHR datasets include different levels of detail

Information	Claims	EHR
Patient Demographics	✓	✓
Patient Social History		✓
Care Provider	✓	✓
Timing and Type of Encounter	✓	✓
Symptoms		✓
Laboratory Results		✓
Diagnoses	✓	✓
Procedures	✓	✓
Prescription Medications	✓	✓
Provider Notes (Unstructured Text)		✓
Costs of Services	✓	✓

Why should you consider claims or EHR data for palliative care research?

Claims and EHR data allow you to make inferences about the effectiveness and costs of care provided in real-world settings.

Claims and EHR data facilitate investigation of processes and outcomes for seriously ill individuals for whom prospective data collection may be infeasible or overly burdensome.

Claims datasets and some EHR datasets are large, which facilitate the study of relatively rare diseases and improve statistical power to detect clinically meaningful effects of interventions.

What should you know before using claims or EHR data in research?

Important contextual data may not be available.

- Standardized data on patient preferences, goals of care, and availability of in-home support from friends or family are not typically available in claims or EHR data.
- Structured data provide information that a palliative care consultation was recorded, but they do not provide detailed information on what comprised the consultation.
- Not every aspect of care during an encounter will be recorded. Informal symptom assessments may guide care decisions but not be recorded.

Not all ICD codes are reliably recorded.

- Codes that reflect social needs and that are not required for billing are less likely to be recorded (e.g., the Z51.5 code for specialty palliative care).

Patients enter and leave datasets at different times.

- Patients will seek care at different points of an illness, so the date of a first diagnosis in claims or EHR data may reflect early- or late-stage disease.
- Patients may change providers or insurers, affecting availability of history and data.

Claims and EHR data reflect a treated sample.

- Patients who receive palliative care are generally more seriously ill than patients who do not. Individuals who do not access care will not be included in the data.
- Analyses of claims and EHR data must account for systematic differences between patients receiving and not receiving care to avoid erroneously inferring that care is causing adverse effects.

Where can you find claims and EHR data?

Source	Details and Links
Centers for Medicare and Medicaid Services	CMS datasets include Medicare Fee-For-Service claims, Medicare Advantage encounter data, and Transformed Medicaid Statistical Information System (available via ResDAC)
Healthcare Cost and Utilization Project	Healthcare Cost and Utilization Project includes the largest collection of longitudinal hospital care data in the United States.
National Institutes of Health	All of Us Research Hub is one of the largest biomedical data resources of its kind, storing health data from participants from across the United States.
National Patient-Centered Clinical Research Network	PCORnet Front Door is an access point for potential investigators seeking to leverage the PCORnet infrastructure and collaborate on patient-centered research.
State-level Departments of Health	State all-payer claims databases (e.g., New York all-payer claims data public use files) include datasets from public and private payers providing insurance to state residents.
Veterans Health Administration	The VA's Corporate Data Warehouse maintains information on patients treated within the VA, which can be used by VA researchers in IRB approved studies.
Individual hospitals/health systems	Individual hospitals and health systems house proprietary data warehouses that aggregate patient data and require authorized, often IRB-approved requests for access.
Commercial aggregate data sources	Commercial aggregate data sources (e.g., Optum, Health Care Cost Institute, Epic Cosmos) provide large-scale, real-world data.

Last updated: May 6, 2026
Created with support from NIH U54AG093230