PINC AI™ Healthcare Database:

Data That Informs and Performs

PINC AI™ Applied Sciences

March 2024

This white paper describes the attributes and capabilities of the PINC AI™ Healthcare Database (PHD), which pharmaceutical companies, device industries, academics, healthcare insurers, and healthcare policy makers use for clinical, financial, and outcomes analyses. Since 2000, more than 1150 publications using PINC AI™ healthcare data on multiple therapeutic and quality improvement/patient safety areas have appeared in more than 260 scholarly, peer-reviewed journals – more than 220 of these are Premier-authored papers (PINC AI™ Applied Sciences website). PINC AI Applied Sciences is responsible for leveraging the HIPAA-compliant PHD through its highly professional and experienced team.

SUGGESTED CITATION: PINC AI™ Applied Sciences, Premier Inc. PINC AI™ Healthcare Database: Data that informs and performs (White Paper). March 2024. https://offers.pinc-ai.com/PINC-AI-Healthcare-Database-White-Paper-LP.html

INTRODUCTION

Electronic healthcare databases house “big data” and offer several advantages as robust research tools.[1-3] A comprehensive hospital-based healthcare history of each patient in the healthcare system may be available. Patients and physicians are not involved in data collection, thus eliminating potential bias that may come with being observed. Since data accrue from a large diverse population, rare outcomes and long-term effects can be studied. The data reflect the state of clinical practice in the general population. These real-world databases lend themselves to creating profiles of drug use and physician prescribing practices and conducting post-marketing studies on treatment effectiveness, safety issues, and cost-effectiveness. Due to its electronic format, information is readily accessible. With appropriate approvals in place, healthcare databases from multiple sources including electronic medical records and claim databases may be able to be linked to provide additional analytic power. In addition, a new paradigm is emerging to use real-world evidence from a variety of real-world data sources to supplement randomized clinical trials and support regulatory decision-making for drugs and devices.[4, 5]

BACKGROUND

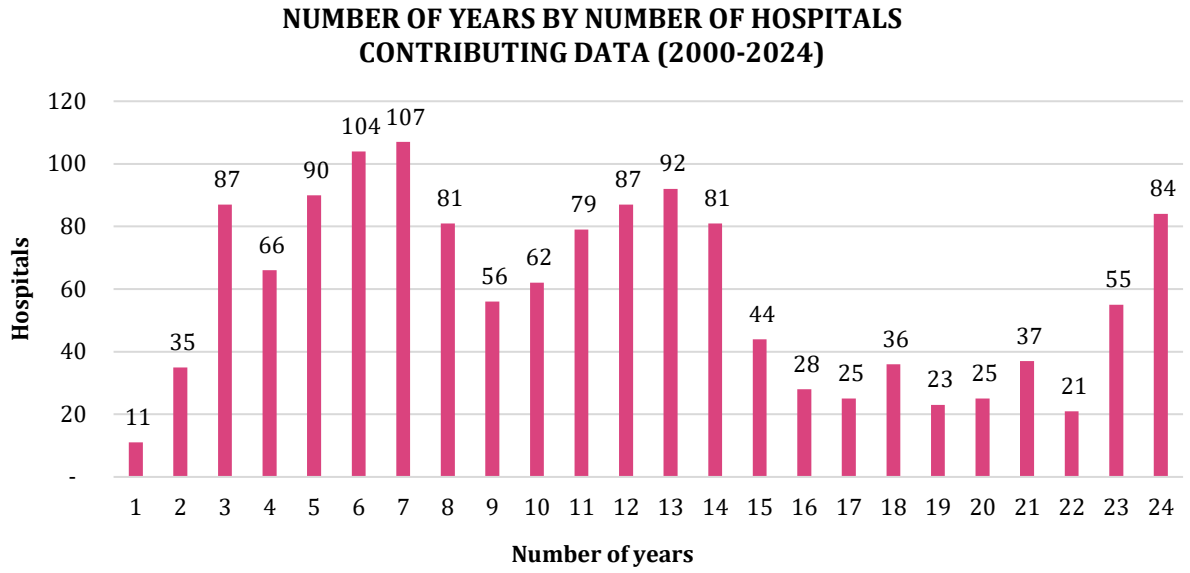
The PINC AI Healthcare Database (PHD), formerly known as the Premier Healthcare Database, is one of the most comprehensive electronic healthcare data repositories. It originated from the merger of Premier with American Healthcare Systems and SunHealth in 1997, when the largest healthcare network of its time was created.[6] The resulting Premier Alliance committed to investing significant financial resources into a database to improve quality of care.[7] Since that time, Premier Inc. has emerged as a leader in healthcare transformation. From Premier’s early beginnings as a group purchasing organization to its leadership in patient safety and quality improvement initiatives, the PHD has been a valuable resource for the pharmaceutical and device industries, academia, federal and national healthcare agencies (including the Centers for Disease Control and Prevention, Centers for Medicare Services,[8] Food and Drug Administration,[9] and National Institutes of Health), as well as for Premier’s member hospitals and health systems to improve healthcare delivery, quality, and cost reduction. More than 1,400 hospitals/healthcare systems contribute data to the PHD. It provides a unique source of real-world data to conduct evidence-based and population-based analyses of drugs, devices, other treatments, disease states, epidemiology, resource utilization, healthcare economics, and clinical outcomes.

THE PINC AI HEALTHCARE DATABASE

Overview

The PHD comprises U.S. hospital-based, service-level, all-payer information on inpatient discharges, primarily from geographically diverse non-profit, non-governmental, and community and teaching hospitals and health systems from rural and urban areas. Hospitals and healthcare systems submit administrative, healthcare utilization, and financial data from patient encounters. Inpatient admissions include more than 169 million visits with more than 9 million per year since 2012, representing approximately 25% of annual United States inpatient admissions.[10] Outpatient encounters include more than 1 billion outpatient visits, with more than 90 million visits per year since 2012. Outpatient visits to emergency departments, ambulatory surgery centers, and alternate sites of care are included. The PHD contains information from more than 346 million unique patients. Using a unique masked identifier, patients can be tracked in the same hospital across the inpatient and hospital-based outpatient settings, and their hospital length of stay and readmissions to the same hospital can be assessed. Information in the PHD is de-identified and HIPAA-compliant in accordance with the HIPAA Privacy Rule.[11]

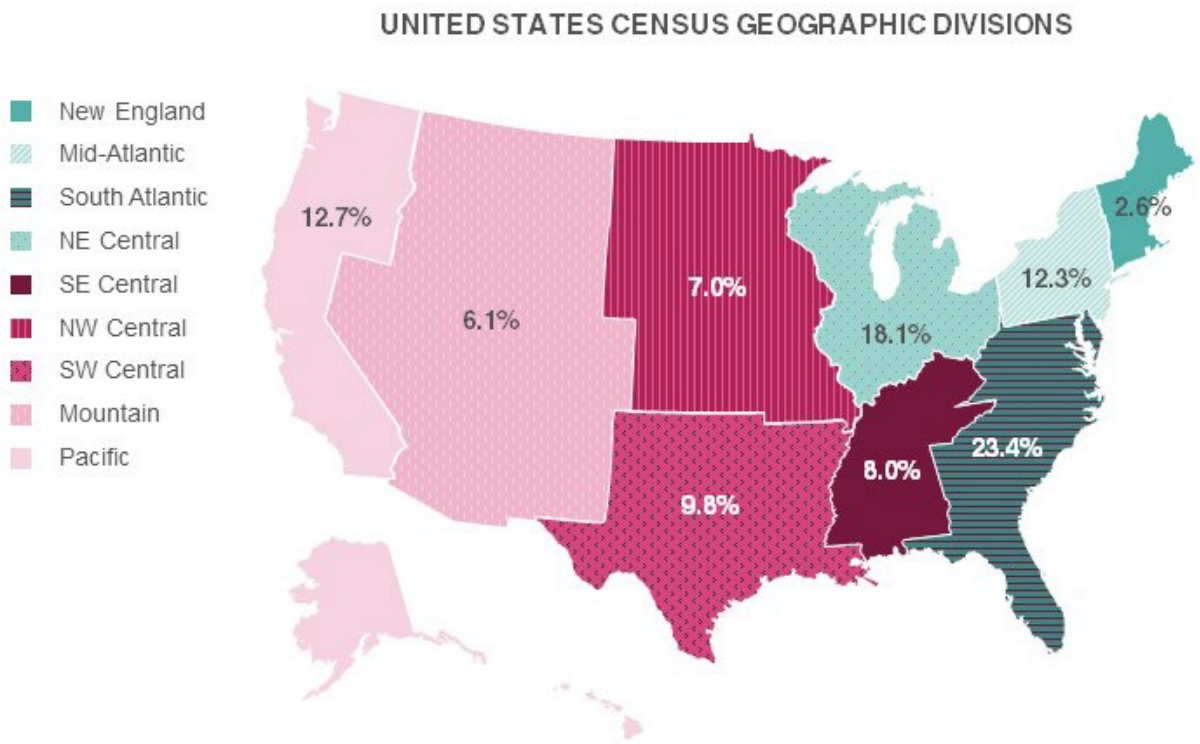
The PHD is updated weekly, with data accruing since January 2000. The number of years that hospitals/healthcare systems have contributed data to the PHD is represented in the graph below.



The PHD contains information on hospital and visit characteristics, admitting and attending physician specialties, healthcare payers, and patient data from standard hospital discharge billing files. These data include demographics and disease states; admission and discharge diagnoses; information on billed services including costs at the departmental level such as medications and devices, laboratory tests performed, diagnostic and therapeutic services, microbiology test results (for a subset of hospitals), and patient disposition and discharge health status. For most data elements, less than 1% of patient records are missing information, and for key elements, such as demographics and diagnostic information, less than 0.01% are missing data.[10]

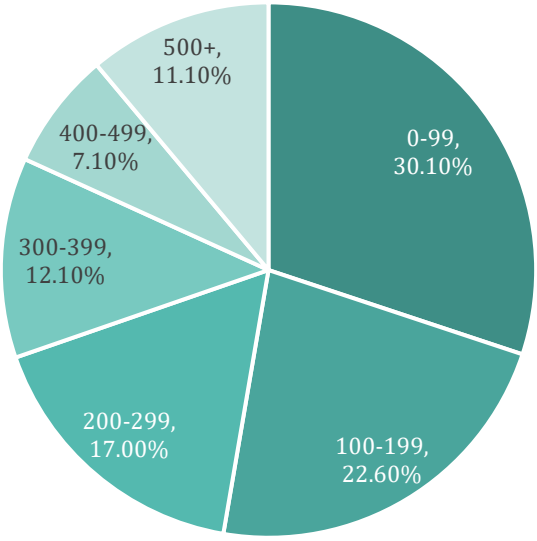
Hospital and Visit Characteristics

The PHD uses member hospital statistics provided from a combination of self-report and the American Hospital Association Annual Survey Database™.[12] Hospitals in the PHD represent the four geographic regions and their respective divisions the United States Census defines them (Northeast: *New England, Middle Atlantic*; Midwest: *East North Central, West North Central*; South: *South Atlantic, East South Central, West South Central*; West: *Mountain, Pacific*).



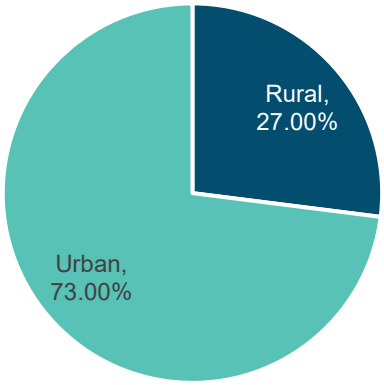
Hospital characteristics of bed capacity, urban and rural populations served, and teaching status are recorded for all hospitals contributing data.

BED CAPACITY

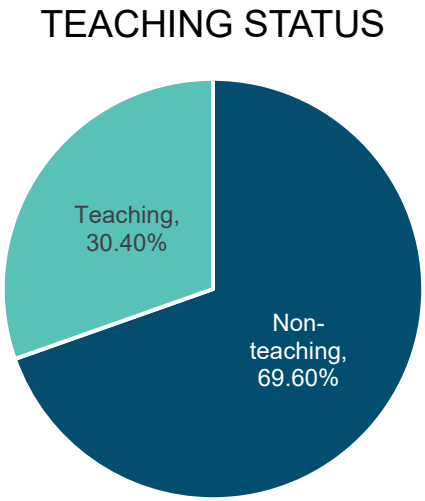


The United States Census defines an urban area as a territory in which core census groups or blocks have a population density of at least 1000 people per square mile, and surrounding census blocks have an overall density of at least 500 people per square mile. Rural areas are considered territory outside the definition of urban.[13]

POPULATION SERVED



A teaching facility has either a medical school affiliation reported to the American Medical Association or a documented affiliation agreement with a medical school accredited by the Association of American Medical Colleges Liaison Committee on Medical Education. These organizations must sponsor or participate significantly in at least four approved active residency programs. At least two of the approved residency programs should be in medicine, surgery, obstetrics/gynecology, pediatrics, family practice, or psychiatry.



Comparisons of the member hospital characteristics from the PHD with those from the American Hospital Association (AHA) demonstrate a similar distribution, although the AHA has a greater number of smaller member hospitals.

	PHD (2023)	AHA DATABASE (2020)*
	N (%)	N (%)
All Facilities	1054 (100)	4306 (100)
Midwest	255 (24.20)	1306 (30.30)
Northeast	139 (13.20)	517 (12.00)
South	448 (42.50)	1610 (37.40)
West	212 (20.10)	873 (20.30)
Rural	314 (29.80)	1044 (24.20)
Urban	740 (70.20)	3262 (75.80)
000-099	351 (33.30)	2235 (51.90)
100-199	231 (21.90)	847 (19.70)
200-299	175 (16.60)	474 (11.00)

	PHD (2023)	AHA DATABASE (2020)*
300-399	120 (11.40)	287 (6.70)
400-499	58 (6.50)	157 (3.60)
500+	109 (10.30)	306 (7.10)
Non-Teaching	733 (69.50)	2351 (54.60)
Teaching	321 (30.50)	1955 (45.40)

**Latest available data from the AHA*

Hospital-Level Encounter Data

Visit-level information in the PHD includes admitting and attending physician specialties, point of origin, type of admission, and discharge status (including mortality). Definitions are based on the elements in hospital claims derived from the uniform billing form (UB-04) and categorized into PINC AI standard definitions, as well as the PINC AI proprietary data dictionary.

HOSPITAL-LEVEL ENCOUNTER DATA

POINT OF ORIGIN

NON-HEALTHCARE FACILITY

- Patients coming from home or workplace, or patients receiving care at home (such as home health service)

CLINIC OR PHYSICIAN'S OFFICE TRANSFER

- Ambulatory Surgery Center; another Home Health Agency
- Distinct unit to another in same hospital with separate claim
- Healthcare facility or born outside hospital
- Hospice, under plan or enrolled
- Skilled nursing facility
- Hospital (different facility)
- Other/unknown such as born inside hospital, Court/Law Enforcement, Information not available

TYPE OF ADMISSION

EMERGENCY

- Patient requires immediate medical intervention as a result of severe, life threatening or potentially disabling conditions. Generally, the patient was admitted through the emergency room

URGENT

- Patient required immediate attention for the care and treatment of a physical or mental disorder. Generally, the patient was admitted to the first available and suitable accommodation

TRAUMA CENTER

- Visits to a trauma center/hospital as licensed or designated by the State or local government authority authorized to do so, or as verified by the American College of Surgeons and involving a trauma activation

ELECTIVE

- Patient's condition permitted adequate time to schedule the availability of suitable accommodations.

OTHER/UNKNOWN

- Newborn
- Information not available

DISCHARGE STATUS

TRANSFER TO ANOTHER HEALTHCARE FACILITY

- Discharged/transferred to another type of healthcare institution not defined elsewhere in this list with or without a planned acute care hospital inpatient readmission
- Nursing facility certified under Medicaid but not certified under Medicare with or without a planned acute care hospital inpatient readmission

- Facility that provides custodial or supportive care with or without a planned acute care hospital inpatient readmission
- Short term general hospital for inpatient care with or without a planned acute care hospital inpatient readmission
- Skilled nursing facility (SNF) with Medicare certification with or without a planned acute care hospital inpatient readmission
- Designated cancer center or children's hospital with or without a planned acute care hospital inpatient readmission
- Federal healthcare facility with or without a planned acute care hospital inpatient readmission
- Inpatient rehabilitation facility (IRF) including rehabilitation distinct part units of a hospital with or without a planned acute care hospital inpatient readmission
- Medicare certified long term care hospital (LTCH) with a planned acute care hospital inpatient readmission
- Critical Access Hospital (CAH)
- Other short term general hospital for inpatient care

TRANSFER TO ANOTHER CARE FACILITY (CONTINUED)

- Intermediate care facility (ICF)
- Another type of institution for inpatient care (including distinct parts); Psychiatric hospital or psychiatric distinct unit of a hospital
- Skilled nursing facility (SNF) with Medicare certification in anticipation of covered skilled care
- Within this institution to a hospital-based Medicare approved swing bed, nursing facility, designated cancer center or children's hospital; another type of healthcare institution not defined elsewhere in code list; long term care hospitals; another rehabilitation facility

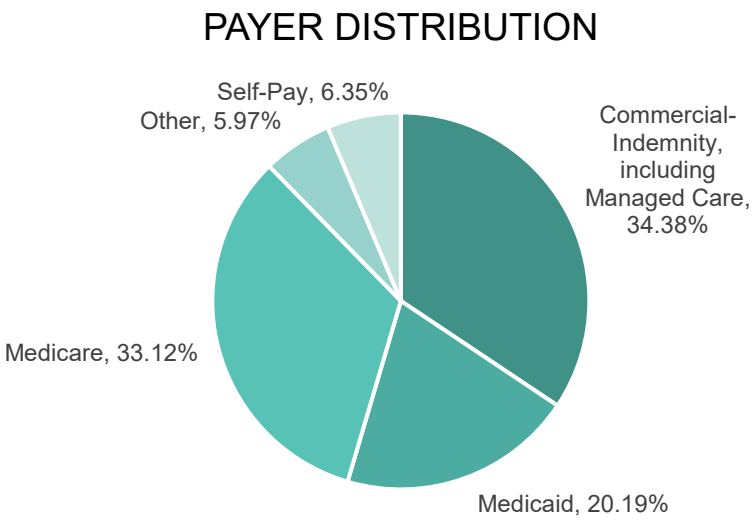
HOME

- Home/self-care
- Home or self-care with a planned acute care hospital readmission
- Home care of organized home health service organization
- Home under care of organized home health service organization with a planned acute care hospital inpatient readmission

HOSPICE-EXPIRED-OTHER/UNKNOWN

- Left against medical advice
- Court/law enforcement with or without a planned acute care hospital inpatient readmission
- Still patient or expected to return for outpatient services
- Information not available

Type of primary payer coverage information adds an important dimension to the PHD and is a useful surrogate for socioeconomic status of patients.



Patient- and Service-Level Data

Patient demographics include age, sex, race (white, black, other) and ethnicity (Hispanic, non-Hispanic). International Classification of Diseases (ICD) Diagnosis Codes for each hospital encounter (ICD-9 for discharge dates prior to10/1/2015 or ICD-10 for discharge dates on or after 10/1/2015) identify disease states and comorbid conditions. ICD procedure codes (version 9 and 10 as described above) as well as hospital-submitted Current Procedural Terminology (CPT) and Healthcare Common Procedure Coding System (HCPCS) codes identify diagnostic and therapeutic procedures ordered during hospital encounters. There is no limit on the number of ICD diagnosis codes that are provided; therefore, all codes a hospital provides are contained in the PHD.

Detailed pharmacy data including brand/generic drug names, strength, dosing, route of administration, day of service charge, and quantity charged are also available in the PHD. Medical devices and supply utilization can also be identified with day of service charge. Additionally, the PHD contains microbiology laboratory results information from more than 516 hospitals, including specimen ID, test name, test day of service and time, specimen source, result, sensitivity data, and observation status (i.e., final, corrected) cumulatively from 2009 forward. Since 2017, more than 404 hospitals provide in-hospital laboratory results. Patient vitals including height, weight, blood pressure, heart rate, and temperature are also available, and a select group of hospitals also provide respiratory function.

The PHD has the required information necessary for generating several clinical algorithms. 3M™ All Patient Refined™ Diagnosis Related Group (APR™-DRG), Severity of Illness (APR-SOI), and Risk of Mortality (APR-ROM) account for age, procedures, and clinical severity of primary diagnosis and all secondary diagnoses assigned during hospitalization and computed for each patient at the time of hospital discharge.[14] The Elixhauser Comorbidity Index measures patient comorbidity based on ICD-9 and ICD-10 diagnosis codes and has been found to predict hospital resource use and in-hospital mortality.[15-17] The Charlson Comorbidity Index measures overall health status through assessment of comorbidities at time of discharge.[16,18,19] Each comorbidity category has an associated weight based on the adjusted risk of one-year mortality, and the sum of all the weights results in a single comorbidity score for a patient. A score of zero indicates that no comorbidities were found. The higher the score, the more likely the predicted outcome will result in higher resource use or mortality. Cunningham Bleeding Score identifies oral anticoagulant bleeding-related hospitalizations calculated from a primary discharge.[20]

Billing and Financial Data

Costs (fixed and variable) to the hospital and charges to the payer can be determined from the PHD. The PHD charge master (CDM) is a comprehensive table of items billable to a hospital patient or to a patient's health insurance provider. It includes hospital services, medical procedures, equipment fees, supplies, drugs, and diagnostic evaluations such as imaging and laboratory tests. Days of services/supplies/drugs ordered, delivered/administered, and billed are documented. Each item in the CDM is assigned a unique identifier code and a set price used to generate patient bills. Each hospital system maintains its own CDM. Reimbursement data are unavailable within these data, as they remain proprietary information hospitals and payers hold.

Two versions of cost are available. Each hospital uses their own cost accounting systems to determine procedural costs. For hospitals that are identified as ratios of costs to charges (RCC), the hospitals provide PINC AI Applied Sciences with charge data, and the teams work with the hospitals to assign Medicare Cost to Charge Ratios (MCCR) to the data provided. Regardless of the source of the cost and charge data, they are reviewed and validated against the data from the hospital at the visit and total numbers within certain variances, before use in the database.

Closed Claims Data

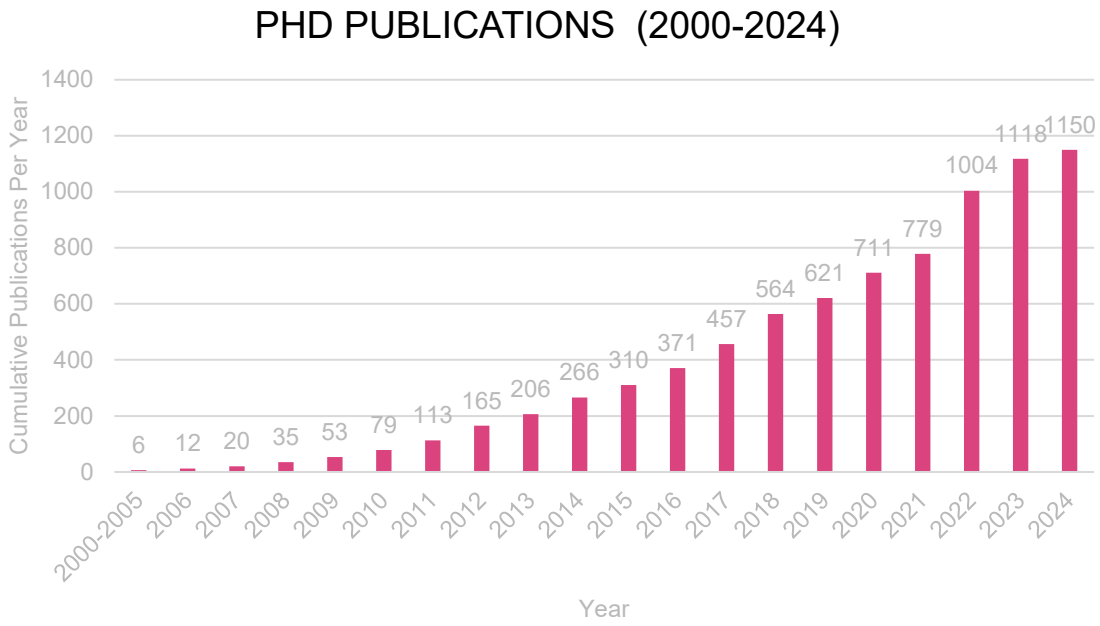
The closed claims data include the combined healthcare claims of patients from dozens of commercial health plans. As of July 2023, it reflects approximately 25 million covered lives from July 1, 2016 forward, and includes the commercially insured population as well as Medicaid and Medicare Advantage. It also includes enrollment history and healthcare claims for both pharmacy and medical benefits from all settings of care, medical claims such as diagnoses and procedures for exact dates of service, and pharmacy claims such as prescribed medications, prescription date, days' supply, and quantity dispensed. The commercial closed claims data are available as a linked source to the PHD. The database is updated monthly and is subject to an approximate 90-day lag.

Uses of Data

Data from the PHD provides insights into patient characteristics, patient care patterns, outcomes and burden of illness over time, and for a wide range of therapeutic areas and products/devices. Market assessments, comparative effectiveness analyses, cost analyses, and cost-effectiveness studies are frequently conducted using the PHD. Data licensing permits

external companies to perform analytics using the data. Additional approvals may be obtained for specific projects that would allow the healthcare database to be linked to other databases such as electronic medical records and claim databases.

To date, 1,150 articles using the PHD have appeared in 264 scholarly journals - between 2000 and March 2024, 228 Premier-authored papers and 922 papers using PHD data that did not have a PINC AI™ Applied Sciences researcher as an author have been published. The graph below highlights the number of publications using the PHD† over time.



†*A small number of publications use* PINC AI *data sources other than the PHD (i.e., Operation Advisor, QUEST)*

BIBLIOGRAPHY

1. Corrao G, Mancia G. Generating evidence from computerized healthcare utilization databases. *Hypertension*. 2015;65(3):490-8. doi:10.1161/HYPERTENSIONAHA.114.04858
2. Schneeweiss S, Avorn J. A review of uses of health care utilization databases for epidemiologic research on therapeutics. *J Clin Epidemiol*. 2005;58(4):323-37. doi:10.1016/j.jclinepi.2004.10.012
3. Suissa S, Garbe E. Primer: administrative health databases in observational studies of drug effects--advantages and disadvantages. *Nat Clin Pract Rheumatol*. 2007;3(12):725-32. doi:10.1038/ncprheum0652
4. Corrigan-Curay J, Sacks L, Woodcock J. Real-World Evidence and Real-World Data for Evaluating Drug Safety and Effectiveness. *JAMA*. 2018;320(9):867-868. doi:10.1001/jama.2018.10136
5. Sherman RE, Anderson SA, Dal Pan GJ, et al. Real-World Evidence - What Is It and What Can It Tell Us? *N Engl J Med*. 2016;375(23):2293-2297. doi:10.1056/NEJMsb1609216
6. Scott L. SunHealth, AmHS/Premier to merge. *Mod Healthc*. Nov 27 1995;25(48):2-3.
7. Hensley S. Catalyst for change. Premier alliance to invest millions in database to improve quality of care. *Mod Healthc*. 1999;29(25):128.
8. Becker C. Time to pay for quality. CMS will partner with premier in trial project to give financial bonuses to hospitals that deliver the best care. *Mod Healthc*. 2003;33(26):6-7, 16, 1.
9. Premier databases to support FDA surveillance related to drug safety. *Health Care Strateg Manage*. 2001;19(11):10.
10. American Hospital Association. Trends in inpatient utilization in community hospitals, 1995-2016. U.S. Census Bureau; 2016.
11. Health & Human Services. Guidance Regarding Methods for De-identification of Protected Health Information in Accordance with the Health Insurance Portability and Accountability Act (HIPAA) Privacy Rule. Accessed March 1, 2022. https://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html
12. American Hospital Association. Fast Facts on U.S. Hospitals. Accessed March 1, 2022. https://www.aha.org/statistics/fast-facts-us-hospitals
13. U.S. Census Bureau. Urban and rural. Accessed March 1, 2022. https://www.census.gov/programs-surveys/geography/guidance/geo-areas/urban-rural.html#:~:text=The%20Census%20Bureau's%20urban%2Drural,non%2Dresidential%20urban%20land%20uses.
14. Averill RF, Goldfield N, Hughes JS, et al. *3M APR DRG Classification System*. 2008. https://www.hcup-us.ahrq.gov/db/nation/nis/v261_aprdrg_meth_ovrview.pdf
15. Elixhauser A, Steiner C, Harris DR, Coffey RM. Comorbidity measures for use with administrative data. *Med Care*. 1998;36(1):8-27. doi:10.1097/00005650-199801000-00004
16. Li B, Evans D, Faris P, Dean S, Quan H. Risk adjustment performance of Charlson and Elixhauser comorbidities in ICD-9 and ICD-10 administrative databases. *BMC Health Serv Res*. 2008;8:12. doi:10.1186/1472-6963-8-12
17. Quan H, Sundararajan V, Halfon P, et al. Coding algorithms for defining comorbidities in ICD-9-CM and ICD-10 administrative data. *Med Care*. 2005;43(11):1130-9. doi:10.1097/01.mlr.0000182534.19832.83
18. Charlson ME, Pompei P, Ales KL, MacKenzie CR. A new method of classifying prognostic comorbidity in longitudinal studies: development and validation. *J Chronic Dis*. 1987;40(5):373-83. doi:10.1016/0021-9681(87)90171-8
19. Deyo RA, Cherkin DC, Ciol MA. Adapting a clinical comorbidity index for use with ICD-9-CM administrative databases. *J Clin Epidemiol*. 1992;45(6):613-9. doi:10.1016/0895-4356(92)90133-8
20. Cunningham A, Stein CM, Chung CP, Daugherty JR, Smalley WE, Ray WA. An automated database case definition for serious bleeding related to oral anticoagulant use. *Pharmacoepidemiol Drug Saf*. 2011;20(6):560-6. doi:10.1002/pds.2109