Clinical data

Clinical data is a staple resource for most health and medical research. Clinical data is either collected during the course of ongoing patient care or as part of a formal clinical trial program. Clinical data falls into six major types:

Electronic health records

Administrative database

Claims data

Disease registries

Health surveys

Clinical trials data

For additional administrative, registry and survey sources such as healthdata.gov, see: Statistics Sources: Health Sciences

Electronic Health Record

The purest type of electronic clinical data which is obtained at the point of care at a medical facility, hospital, clinic or practice. Often referred to as the electronic medical record (EMR), the EMR is generally not available to outside researchers. The data collected includes administrative and demographic information, diagnosis, treatment, prescription drugs, laboratory tests, hospitalization, patient insurance, etc. Administrative Data

Often associated with electronic health records, these are primarily hospital discharge data reported to a government agency like AHRQ.

Healthcare Cost & Utilization Project (H-CUP)

HCUPnet is a free, on-line query system based on data from the Healthcare Cost and Utilization Project (HCUP). It provides access to health statistics and information on hospital inpatient and emergency department utilization. The project includes a number of datasets and sample studies listed under the information icon.

Claims Data

Claims data describe the billable interactions (insurance claims) between insured patients and the healthcare delivery system. Claims data falls into four general categories: inpatient, outpatient, pharmacy, and enrollment. The sources of claims data can be obtained from the government (e.g., Medicare) and/or commercial health firms (e.g., United HealthCare).

Basic Stand Alone (BSA) Medicare Claims Public Use Files (PUFs)

This is the Basic Stand Alone (BSA) Public Use Files (PUF) for Medicare claims. This is a claim-level file in which each record is a claim incurred by a 5% sample of Medicare beneficiaries. Claims include inpatient/outpatient care, prescription drugs, DME, SNF, hospice, etc. There are some demographic and claim-related variables provided in every PUF.

Medicaid Statistical Information System

MSIS is the basic source of state-submitted eligibility and claims data on the Medicaid population, their characteristics, utilization, and payments and is available by clicking on the link on the left-side column.

PharMetrics Integrated Database (IMS Rx Benchmark)

The PharMetrics Integrated Database is the largest non-Payer owned integrated claims database of commercial insurers in the U.S. This de-identified, Integrated Database includes medical and pharmacy claims for more than 70 million members from more than 100 health plans across the U.S. The Integrated Database includes inpatient and outpatient claims, diagnoses and procedures based on ICD-9 and CPT-4 codes, as well as retail and mail order pharmacy claims.

Disease Registries

Disease registries are clinical information systems that track a narrow range of key data for certain chronic conditions. The most common conditions captured include cancer, diabetes, heart disease, and asthma. Registries often provide critical information for managing patient conditions.

```
  Global Alzheimer's Association Interactive Network (GAAIN)
  The Global Alzheimer’s Association Interactive Network (GAAIN) is a collaborative project that will provide researchers around the globe with access to a vast repository of Alzheimer’s disease research data and the sophisticated analytical tools and computational power needed to work with that data.
  National Cardiovascular Data Registry (NCDR)
  The NCDR® is the American College of Cardiology’s worldwide suite of data registries helping hospitals and private practices measure and improve the quality of cardiovascular care they provide. The NCDR encompasses six hospital-based registries and one outpatient registry. There are currently more than 2,400 hospitals and nearly 1,000 outpatient providers participating in NCDR registries.
  National Program of Cancer Registries
  CDC provides support for states and territories to maintain registries that provide high-quality data. Data collected by local cancer registries enable public health professionals to understand and address the cancer burden more effectively.
  National Trauma Data Bank
  The National Trauma Data Bank® (NTDB) is the largest aggregation of trauma registry data ever assembled. The goal of the NTDB is to inform the medical community, the public, and decision makers about a wide variety of issues that characterize the current state of care for injured persons.
  Surveillance, Prevention, and Management of Diabetes Mellitus DataLink (SUPREME DM)
```

Health Surveys

In order to provide an accurate evaluation of the population health, national surveys of the most common chronic conditions are generally conducted to provide prevalence estimates. National surveys are one of the few types of data collected specifically for research purposes, thus making it more widely accessible.

```
  Medicare Current Beneficiary Survey
  The Medicare Current Beneficiary Survey (MCBS) is a continuous,
multipurpose survey of a nationally representative sample of the Medicare
population. The central goals of MCBS are to determine expenditures and
sources of payment for all services used by Medicare beneficiaries.
  National Health & Nutrition Examination Survey (NHANES)
  The National Health and Nutrition Examination Survey (NHANES) is a program
of studies designed to assess the health and nutritional status of adults
and children in the United States. The survey is unique in that it combines
interviews and physical examinations.
  National Long Term Care Survey
  The National Long Term Care Survey is funded through a Cooperative
Agreement between the National Institute on Aging (NIA) and Duke University.
It is a longitudinal survey designed to study changes in the health and
functional status of older Americans (aged 65+). It also tracks health
expenditures, Medicare service use, and the availability of personal,
family, and community resources for caregiving.
  National Medical Expenditure Survey
  The Medical Expenditure Panel Survey (MEPS) is a set of large-scale
surveys of families and individuals, their medical providers, and employers
across the United States. MEPS is the most complete source of data on the
cost and use of Healthcare and health insurance coverage.
  National Center for Health Statistics
  A rich source of health data and statistics on a variety of topics.
  CMS Data Navigator
  Center for Medicare & Medicaid Services - Research, Statistics, Data &
Systems
  National Health and Aging Trends Study (NHATS)
  NHATS is a study of Medicare beneficiaries age 65 years and older. The
study is being conducted by the Johns Hopkins University Bloomberg School of
Public Health, with data collection by Westat, and support from the National
Institute on Aging. NHATS is intended to foster research that will guide
efforts to reduce disability, maximize health and independent functioning,
and enhance quality of life at older ages.
```

Clinical Trial Registries & Results Databases

```
  ClinicalTrials.gov
  o Registry and results database.
  o Information on publicly and privately supported clinical studies from
around the world.
  Current Controlled Trials
  o Registry of randomized controlled trials.
  o Can search by ISRCTN.
  European Union Clinical Trials Database
  o Protocol and results information on interventional clinical trials
conducted in the EU from May 1, 2004 onwards.
  o Good source of pediatric drug development trials.
  o Includes pediatric clinical trials which form part of a pediatric
investigation plan (PIP), including those where the investigator sites are
outside the EU.
```

```
  IFPMA Clinical Trials Portal
  o Portal created by the International Federation of Pharmaceutical
Manufacturers & Associations
  o Search engine provides access to ongoing trials and trial results from
ClinicalTrials.gov, Current Controlled Trials, Japan Pharmaceutical
Information Center, and member company corporate websites.
  o Best single place to search for pharmaceutical company sponsored
clinical trials.
  WHO International Clinical Trials Registry Platform
  o Access to data provided by clinical trial registries around the world
that meet WHO criteria for content and quality.
  Immune Tolerance Network TrialShare
  o Clinical and mechanistic assay data from 35 clinical trials.
  o Thousands of biological specimens from clinically well-phenotyped study
participants.
  Drug Delivery Clinical Trials Database
  o Tool for identifying active clinical trials studying drug delivery
platforms, formulations, and PK/PD interactions.
```