Data and Statistics for Public Health

These resources can be found in the Matas Library Statistical Portal
http://libguides.tulane.edu/DataSets

Key Concepts: Overview

Broad Categories: Individual and Population
1. Individuals: Counts are of individuals who are provided particular health services health-related events or services
2. Population: Using sampling techniques to identify an appropriate survey population

Important Features of Health Statistics and Data Sets

The Context of Health Statistics: Health statistics are collected to meet the needs of the collector.
Note: Read the notes describing the reasons for and methods of data collection.
Some information is collected in more than one survey and estimates of the same statistic may vary among surveys.

Key Features of Health Statistics: population based and collected/analyzed over time

Health Statistics Come from Diverse Sources: Studies use administrative data: enrollment or eligibility information, claims information, and managed care encounters. Surveys are designed to collect specific data and are often conducted by trained personnel who administer them by telephone or in-person.

Health Data on the Web

Statistical Information on the Internet: Web sites are produced by organizations for a variety of purposes.
Statistical Information Not on the Internet: In some cases only summaries or partial data sets are available, and the full sets must be purchased.
Finding Statistics and Data Sets on the Internet: A portal or Research Guide is a gateway to other web sites. A search engine is a computer program that retrieves documents, files or data from a database or from a computer network (especially from the Internet)

Web Sites

Health Statistics Portals/Gateways
Partners in Information Access for the Public Health Workforce (phPartners.org)
http://phpartners.org/

MedlinePlus
http://medlineplus.gov/
See Health Topic – Health Statistics

Directory of HHS Data Resources
http://aspe.hhs.gov/datasets

Statistical Abstract of the United States – U.S. Census Bureau
http://www.census.gov/compendia/statab/

FedStats – Federal Government
http://fedstats.sites.usa.gov/
Provides access to official statistics collected and published by more than 100 federal agencies, much of it at the state level.
Federal Government Health Statistics Agencies gather, analyze, and report statistical data useful for public health purposes.

**National Center for Health Statistics (NCHS)**
http://www.cdc.gov/nchs/
This is the principal health statistics agency in the U.S. It is part of the Centers for Disease and Control and Prevention (CDC).

**Agency for Healthcare Research and Quality (AHRQ)**
http://www.ahrq.gov/
AHRQ is the lead scientific research federal agency charged with supporting research designed to improve the quality of healthcare, reduce its cost, improve patient safety, decrease medical errors, and broaden access to essential services.

**Substance Abuse and Mental Health Services Administration (SAMHSA)**
http://www.samhsa.gov/
SAMHSA’s Office of Applied Studies (OAS) provides the latest national data on 1) alcohol, tobacco, marijuana, and other drug abuse, 2) drug related emergency department episodes and medical examiner cases, and 3) the nation’s substance abuse treatment system.

**Centers for Medicare and Medicaid Services (CMS) (formerly HCFA)**
http://www.cms.gov/
CMS offers researchers and other health care professionals a broad range of quantitative information from estimates of future Medicare and Medicaid spending to enrollment, spending, and claims data.

**Data Sets: Federal**

**National Library of Medicine**

Health Data Tools and Statistics - Partners in Information Access for the Public Health Workforce
http://phpartners.org/health_stats.html
This public health portal has categorized data tools and statistics into these subcategories: Health Statistics, National Public Health Data Sets, State and Local Public Health Data Sets, Public Health Infrastructure Data, Search for Other Tools, and Tools for Data Collection and Planning.
Health Services and Sciences Research Resources (HSRR) – National Information Center on Health Services Research and Health Care Technology (NICHSR)
Directory of research datasets and instruments used in health services research and public health. Entries have URLs for more information, and most have pre-formulated PubMed/MEDLINE searches.

Toxicology and Environmental Health – Specialized Information Services
http://sis.nlm.nih.gov/Tox/ToxMain.html
Hazardous Substances Data Bank has chemical data valuable for environmental health concerns.

TOXNET – Specialized Information Services, National Library of Medicine
A variety of databanks and databases on toxicology, hazardous chemicals, and related areas including TRI, Toxics Release Inventory, and HSDB, Hazardous Substances Data Bank.

Agency for Healthcare Research and Quality
HCUP – Healthcare Cost and Utilization Project
http://hcupnet.ahrq.gov/
Provides access to health statistics and information on hospital inpatient and emergency department utilization. HCUP databases bring together the data collection efforts of state data organizations, hospital associations, private data organizations, and the federal government to create a national information resource of discharge-level health care data. HCUP includes the largest collection of longitudinal hospital care data in the United States, with all-payer, discharge-level information beginning in 1988.

MEPS – Medical Expenditure Panel Survey
http://meps.ahrq.gov/mepsweb/
MEPS is a survey of the civilian population living in U.S. communities. MEPS produces nationally representative statistics on health care expenses, including the type of medical services used, how frequently they are used, the cost of services, and how they are paid for, as well as health conditions and health insurance availability and coverage. MEPS also collects extensive information on employer-based health insurance plans.

Centers for Disease Control and Prevention (CDC)
Behavioral Risk Factor Surveillance System (BRFSS)
http://www.cdc.gov/brfss/
BRFSS is the world’s largest telephone survey; it tracks health risks of adults in the United States. A federal and state partnership allows data collection at the national, state, and local levels.

National Hospital Care Survey (NHCS)
http://www.cdc.gov/nchs/nhcs.htm
The NHCS is a collection of health care provider surveys, obtaining information about the facilities that supply health care, the services rendered, and the characteristics of the patients served.

National Health Interview Survey (NHIS)
http://www.cdc.gov/nchs/nhis.htm
NHIS is a continuing nationwide survey of the U.S. civilian noninstitutionalized population conducted in households. Each week a probability sample of households is interviewed by trained personnel of the U.S. Bureau of the Census to obtain information about the health and other characteristics of each living member of the sample household.
National Health and Nutrition Examination Survey (NHANES)
http://www.cdc.gov/nchs/nhanes.htm
NHANES is a survey that collects information about the health and diet of people in the United States. It is unique in that it combines a home interview with health tests that are done in a mobile examination center.

National Immunization Survey
http://www.cdc.gov/vaccines/imz-managers/coverage/imz-coverage.html
A large on-going survey of immunization coverage among U.S. preschool children (19 - 35 months old).

National Notifiable Disease Surveillance System (NNDSS)
http://wwwn.cdc.gov/nndss/
State health departments report notifiable infectious diseases to CDC.

National Vital Statistics System
http://www.cdc.gov/nchs/nvss.htm
Compiles information from states on vital events - births, deaths, marriages, divorces, and fetal deaths.

Youth Risk Behavior Surveillance System (YRBSS)
http://www.cdc.gov/nccdphp/dash/yrbs/index.htm
The YRBSS includes national, state, and local school-based surveys of representative samples of 9th through 12th grade students. These surveys are conducted every two years, usually during the spring semester. The national survey, conducted by CDC, provides data representative of high school students in public and private schools in the United States. The state and local surveys, conducted by departments of health and education, provide data representative of the state or local school district.

Substance Abuse and Mental Health Services Administration
National Survey on Drug Use & Health
https://nsduhweb.rti.org/respweb/homepage.cfm
This is the primary source of information on the prevalence, patterns, and consequences of alcohol, tobacco, and illegal drug use and abuse in the general U.S. civilian non-institutionalized population, aged 12 and older.

The Drug Abuse Warning Network (DAWN)
http://www.samhsa.gov/data/emergency-department-data-dawn
The Drug Abuse Warning Network (DAWN) is a public health surveillance system that monitors drug-related hospital emergency department visits in order to report on the impact of drug use, misuse, and abuse in metropolitan areas and across the nation. Relies on emergency department and medical examiner data.

Drug and Alcohol Services Information System (DASIS)
http://wwwdasis.samhsa.gov/dasis2/index.htm
Includes substance abuse treatment facilities data.

Centers for Medicare and Medicaid Services (CMS) (Formerly HCFA)
CMS Data Navigator: Research, Statistics, and Data & Systems
This site provides access to CMS computer data and systems; research studies and reports; statistics, series, and trends; and files for order.
Medicare Current Beneficiary Survey (MCBS)
A continuous, multipurpose survey of a nationally representative sample of aged, disabled, and institutionalized Medicare beneficiaries. MCBS produces two files annually, Access to Care and Cost and Use.

**Federal Interagency Forum on Aging-Related Statistics (Forum)**

Ongoing Federal Data Resources Relevant to the Study of the Aging
http://www.agingstats.gov/agingstatsdotnet/Main_Site/Data/Resources.aspx
The specific goals of the Forum are to improve both the quality and use of data on the aging population by:
- widening access to information on the aging population through periodic publications and other means;
- promoting communication among data producers, researchers, and public policymakers;
- coordinating the development and use of statistical databases among federal agencies;
- identifying information gaps and data inconsistencies;
- investigating questions of data quality;
- encouraging cross-national research and data collection on the aging population; and
- addressing concerns regarding collection, access, and dissemination of data.

**Data Sets: State and Local**

**Federal Government and National Organization Sources**

**Behavioral Risk Factor Surveillance System (BRFSS)**
http://www.cdc.gov/brfss/
States can add questions on a wide range of important health issues. For example, following the September 11, 2001, terrorist attack on the World Trade Center, New York, New Jersey, and Connecticut added questions to their Behavioral Risk Factor Surveys to measure the psychological effects of this traumatic event. BRFSS data can also be analyzed to examine smaller geographic areas within states.

**Foodborne Illness Surveillance, Response, and Data Systems**
http://www.cdc.gov/foodborneburden/surveillance-systems.html
Each surveillance system described plays a role in detecting and preventing foodborne disease and outbreaks.

**Child Stats – Federal Interagency Forum on Child and Family Statistics**
http://www.childstats.gov/
This web site provides federal and state statistics and reports on children and their families.

**FEDSTATS – Federal government**
http://www.fedstats.gov/
Provides access to official statistics collected and published by more than 100 federal agencies, much of it at the state level.

**State & County QuickFacts – U.S. Census Bureau**
http://quickfacts.census.gov/qfd/
QuickFacts tables are summary profiles showing frequently requested data items from various Census Bureau programs. Profiles are available at the national, state, and county level.

**State and Local Area Integrated Telephone Survey - CDC**
http://www.cdc.gov/nchs/slaits.htm
State or Territorial Health Departments - CDC
http://www.cdc.gov/mmwr/international/relres.html

State Profiles – Children in the States Factsheets – 2-14: Children’s Defense Fund

State Health Facts Online - Kaiser Family Foundation
http://www.statehealthfacts.kff.org/

**State and Local Data Sources**
http://phpartners.org/health_stats.html#State Health Data

Many states have their own systems for reporting data. Usually these systems include births, deaths, marriages, and divorces. A core public health function is statewide injury data collection and analysis. This data collection provides data for mortality statistics found on most state public health websites. Check individual state public health department websites to find what is provided. Often local statistics are available by census tract, community, county, and region. Cities and regions are also providing internet access to data.

**Data - International**

United Nations Statistics Division
http://unstats.un.org/unsd/

The Statistics Division compiles statistics from many international sources and produces global updates in specialized fields of statistics. Free access is provided to country specific population data.

World Bank Group Data and Statistics
http://data.worldbank.org/

Provides data derived, either directly or indirectly, from official statistical systems organized and financed by national governments. Click on Data by Topic to find several options, including health.

Global Health Observatory (GHO)
http://www.who.int/gho/en/

WHO’s gateway to health-related statistics from around the world. GHO covers global health priorities such as the health-related Millennium Development Goals, women and health, mortality and burden of disease, disease outbreaks, health equity, and health systems.

Pan American Health Organization (PAHO)
http://www.paho.org/

Organisation for Economic Co-operation and Development (OECD)
http://www.oecd.org/statistics/

**Statistics - Federal**

Many federal agencies provide access to reports, tables, and figures that present results of statistical analysis of health data

**National Library of Medicine**

MedlinePlus - Health Statistics
This site provides access to web sites that contain compiled health statistics, often on a topic or by an audience.

**U.S. Department of Health and Human Services**

Healthy People 2020 Database
Interactive database system contains the most recent monitoring data for tracking Healthy People 2020. Has national and state level data. Users can construct tables for specific objectives, or objectives identified by focus areas, data source, or by select population.

**Centers for Disease Control and Prevention**

CDC Wonder
Allows users to query dozens of numeric data sets on CDC’s mainframe and other computers, via "fill-in-the blank" request screens. Public-use data sets about mortality, cancer incidence, hospital discharges, AIDS, behavioral risk factors, diabetes, and many other topics are available for query. The requested data can be readily summarized and analyzed.

Data Summary Descriptions
A list of the more than 30 data sets that can be queried by CDC Wonder

**NCHS Health E-Stats**
NCHS gathers compiled statistics from its printed publications at this web site.

Prevalence Data from the BRFSS (Behavioral Risk Factor Surveillance System)
[http://www.cdc.gov/brfss/data_tools.htm](http://www.cdc.gov/brfss/data_tools.htm)
Users can query the BRFSS data sets for statistical information.

Youth Risk Behavior Surveillance System (YRBSS)
[http://www.cdc.gov/HealthyYouth/yrbs/index.htm](http://www.cdc.gov/HealthyYouth/yrbs/index.htm)
The Youth Risk Behavior Surveillance System (YRBSS) monitors health-risk behaviors that contribute to the leading causes of death and disability among youth and adults, including behaviors that contribute to unintentional injuries and violence, tobacco use, alcohol and other drug use, unhealthy dietary behaviors, and physical activity.

**Agency for Healthcare Research and Quality**

HCUPnet
On-line query system based on data from the Healthcare Cost and Utilization Project (HCUP) that provides access to health statistics and information on hospital inpatient and emergency department utilization.

HIV & AIDS Research and Data
Data tools, resources, and research findings on HIV/AIDS cost and use.
Medical Expenditure Panel Survey (MEPS)
http://meps.ahrq.gov/mepsweb/
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 health care and health insurance coverage.

Centers for Medicare and Medicaid Services (CMS)
Research, Statistics, Data & Systems
This site provides access to CMS computer data and systems; research studies and reports; statistics, series, and trends; and files for order.

Substance Abuse and Mental Health Services Administration (SAMHSA)
On-Line Analysis and Public Use Files
http://www.oas.samhsa.gov/SAMHDA.htm
System provides ready access to substance abuse and mental health research data. The site also includes links to detailed analysis provided in data tables by Topic and Data Type.

Quick Statistics - - State Profiles
http://wwwdasis.samhsa.gov/webt/NewMapv1.htm
Get state level data from Treatment Epidode Data Set (TEDS) which provides information on the demographic and substance abuse characteristics of the 1.9 million annual admissions to treatment for abuse of alcohol and drugs in facilities that report to individual state administrative data systems. Also get statistics from N-SSATS, an annual survey designed to collect data on the location, characteristics, and use of alcohol and drug abuse treatment facilities and services throughout the 50 states, the District of Columbia, and other U.S. jurisdictions. These resources are good examples of systems that uses state level administrative data.

National Addiction & HIV Data Archive Program
http://www.icpsr.umich.edu/icpsrweb/NAHDAP/
Acquires, preserves and disseminates data relevant to drug addiction and HIV research

Environmental Protection Agency (EPA)
EPA Data Finder
http://www.epa.gov/datafinder/
Provides access to EPA's data sources organized by subject as well as links to EPA's Public Data JSON file, EPA's Public Excel file, and EPA's Environmental Dataset Gateway.

Health Resources and Services Administration (HRSA)
HRSA Data Warehouse
http://datawarehouse.hrsa.gov/
The enterprise repository for HRSA data. The data are updated frequently and are integrated with external information and data sources, enabling users to gather relevant and meaningful information about health care programs and the populations they serve.

Statistical Information by Subject
Geographic Information Systems (GIS)
GIS data sets can be useful for many purposes. They are used primarily to display data geographically. In public health, applications for GIS are becoming more accessible as geo-coded health data and environmental exposure data increasingly become available, and new and easier-to-use GIS software is developed. The U.S. Agency for Toxic Substances Disease Registry (ATSDR) uses GIS to monitor the health of persons living near hazardous waste sites, and to identify areas of potential concern resulting from accidental release of chemicals in the environment.

Chronic Disease GIS Exchange  
An online forum for showcasing maps of heart disease, stroke, and other chronic diseases (e.g., cancer, diabetes, and asthma) as well as learning Geographic Information Systems (GIS) techniques and accessing a wide range of GIS-related resources. The site encourages the use of GIS to document the burden of chronic diseases, inform and guide policy and program decisions, build and strengthen partnerships, and facilitate collaboration among health department programs, especially within chronic disease units.

Web-based Injury Statistics Query and Reporting System (WISQARS™) Interactive Statistics and Mapping System - CDC  
CDC’s WISQARS™ (Web-based Injury Statistics Query and Reporting System) is an interactive, online database that provides fatal and nonfatal injury, violent death, and cost of injury data from a variety of trusted sources. Researchers, the media, public health professionals, and the public can use WISQARS™ data to learn more about the public health and economic burden associated with unintentional and violence-related injury in the United States.

Geographic Analysis Tool for Health & Environmental Research (GATHER) - Agency for Toxic Substances and Disease Registry (ATSDR)  
[http://serc.carleton.edu/resources/15704.html](http://serc.carleton.edu/resources/15704.html)
GATHER is an online spatial data access system that provides members of the public health community access to spatial data that is pertinent to the analysis and exploration of public health issues.

EnviroMapper Storefront – U.S. Environmental Protection Agency (EPA)  
[http://www.epa.gov/emefdata/em4ef.home](http://www.epa.gov/emefdata/em4ef.home)
A single point of access to select U.S. EPA environmental data

Dartmouth Atlas of Health Care – Center for the Evaluative Clinical Sciences at Dartmouth Medical School  
[http://www.dartmouthatlas.org/](http://www.dartmouthatlas.org/)
The Atlas project focuses on how medical resources are distributed and used in the United States.

Cancer Mortality Maps & Graphs – National Cancer Institute (NCI), NIH  
Provides interactive maps, graphs (which are accessible to the blind and visually-impaired), text, tables and figures showing geographic patterns and time trends of cancer death rates for the time period 1950-1994 for more than 40 cancers.

**Hospital and Health Care Records**  
National Hospital Care Survey  
Tracks the latest trends affecting hospitals and health care organizations and factors that influence the use of health care resources, the quality of health care, and disparities in health care services provided to population subgroups in the United States.
Medical Expenditure Panel Survey [MEPS]
http://meps.ahrq.gov/

HCUP
http://www.ahrq.gov/research/data/hcup/index.html

**Mortality and Morbidity Data**

**National Vital Statistics System**
http://www.cdc.gov/nchs/nvss/about_nvss.htm
The National Vital Statistics System is the oldest and most successful example of inter-governmental data sharing in Public Health and the shared relationships, standards, and procedures form the mechanism by which NCHS collects and disseminates the Nation's official vital statistics.

**Mortality Data from the National Vital Statistics System**
http://www.cdc.gov/nchs/deaths.htm
Mortality data from the National Vital Statistics System (NVSS) are a fundamental source of demographic, geographic, and cause-of-death information. This is one of the few sources of health-related data that are comparable for small geographic areas and are available for a long time period in the United States. The data are also used to present the characteristics of those dying in the United States, to determine life expectancy, and to compare mortality trends with other countries.

**National Notifiable Disease Surveillance System (NNDSS)**
http://www.cdc.gov/nndss/
The National Notifiable Diseases Surveillance System (NNDSS) is a nationwide collaboration that enables all levels of public health—local, state, territorial, federal, and international—to share notifiable disease-related health information.

**MMWR Search**
http://www.cdc.gov/mmwr/mmwrsrch.htm
Often called “the voice of CDC,” the MMWR series is the agency’s primary vehicle for scientific publication of timely, reliable, authoritative, accurate, objective, and useful public health information and recommendations.

---

**Using Data**

**Tools for Building Reports and Tables**

DataFerrett: Federal Electronic Research and Review Extraction Tool
http://dataferrett.census.gov/
DataFerrett is a data analysis and extraction tool to customize federal, state, and local data to suit your requirements.

Epi Info
http://www.cdc.gov/epiinfo/
Epi Info fills a need for a simple tool that allows the rapid creation of data collection instruments and data analysis, visualization, and reporting using epidemiologic methods.

**Building Your Own Tables**
WISQARS (Web-based Injury Statistics Query and Reporting System)
http://www.cdc.gov/ncipc/wisqars/
An interactive database system that provides customized reports of injury related data.

Resources to Help You Learn More
Research Data Assistance Center (ResDAC)
http://www.resdac.org/
Provides technical assistance to researchers interested in using Medicare and/or Medicaid data.

State Health Access Data Assistance Center (SHADAC)
http://www.shadac.org/
SHADAC, funded by the Robert Wood Johnson Foundation, helps states monitor rates of health insurance coverage and to understand factors associated with uninsurance. SHADAC provides targeted policy analysis and technical assistance to states that are conducting their own health insurance surveys and/or using data from national surveys.

TOXNET and Beyond: Using the National Library of Medicine Environmental Health and Toxicology Portal
This is a training manual to familiarize participants with reliable online environmental health and toxicology information from the National Library of Medicine® (NLM) and other reliable sources. Skills and knowledge acquired in this training class will enable participants to access, utilize, and refer others to environmental health and toxicology information.

Selected US Government Publications
http://www.cdc.gov/nchs/hus.htm
The full-text of this annual report on national trends in health statistics is available on the Internet.

Vital and Health Statistics Series, NCHS
Series 1 through Series 24 (commonly called the Rainbow Series) are available in html.

Morbidity and Mortality Weekly Report (MMWR)
http://www.cdc.gov/mmwr/index.html
The data in the weekly MMWR are provisional, based on weekly reports to CDC by state health departments.