

Examples of Available Datasets

Available Datasets for Harkness projects

All applicants are strongly encouraged to explore what U.S. datasets are available related to their proposed project. This document highlights some examples of important U.S. datasets for health services research. This is a partial list, and all applicants are strongly encouraged to explore further.

One option to explore is whether your proposed project can “piggyback” or collaborate with an already ongoing U.S. research project. The National Institutes of Health offers an excellent resource, “HSRProj,” that provides information on research-in-progress before results are published:

http://wwwcf.nlm.nih.gov/hsr_project/home_proj.cfm

While not exhaustive, HSRProj contains information on a large number of ongoing projects, funded by both public and private research agencies.

The National Institutes of Health also provides an excellent resource through the Health Services and Science Research Resources website, which contains information on publicly available datasets and research tools:

http://wwwcf.nlm.nih.gov/hsrr_search/index.cfm

This free portal allows users to search for datasets, instruments, and software, including: clinical records, discharge summaries, claims records, epidemiological surveys, health/ behavioral/ social surveys, disease registries, birth registries, and data about practitioners, programs, and facilities. The database also includes brief descriptions of each resource, links to PubMed for related articles, and URLs for additional information or access to the resources.

In addition, the Department of Health and Human Services operates a website where users can search and download publicly available datasets and tools:

http://www.data.gov/communities/node/81/data_tools

Users can filter the available datasets by category, including: Health Care Resources, Health Care Utilization, Cost, and Quality, and Pharmaceuticals and Other Regulated Products.

Finally, please note that Harkness Mentors may have access to data that is not publicly available, potentially offering applicants more options than are found through the NIH and HHS search engines.

Examples of Available Datasets

Examples of Available Datasets

Healthcare Cost and Utilization Project (HCUP) (Agency for Healthcare Research and Quality)

<http://www.ahrq.gov/data/hcup/>

A family of administrative, longitudinal databases, web-based products and software tools developed and maintained by the Agency for Healthcare Research and Quality (AHRQ) as part of a Federal-State-Industry partnership to build a standardized, multi-state health data system. HCUP is based on data collected by individual States and provided to AHRQ by the States. HCUP data are used for research on hospital utilization, access, charges, quality and outcomes. The data are used to describe national, regional and state level patterns of care for uncommon as well as common diseases, analyze hospital procedures, including those that are performed infrequently, and study the care of population sub-groups such as minorities, children, women, and the uninsured.

Medical Expenditure Panel Survey (MEPS) (Agency for Healthcare Research and Quality)

<http://www.meps.ahrq.gov/mepsweb/>

The Medical Expenditure Panel Survey (MEPS) is a nationally representative survey of health care use, expenditures, sources of payment and insurance coverage for the U.S. civilian non-institutionalized population, as well as a national survey of nursing homes and their residents. MEPS is designed to help understand how the recent growth in managed care and changes in the health care system have affected the kinds, amounts, and costs of health care that Americans use. MEPS consists of five component surveys:

The *Household Component* is a household interview of 10,500 families and 24,000 individuals in 190 communities across the U.S. The objective is to produce annual estimates for a variety of measures in health status, health insurance coverage, health care use and expenditures and sources of payment for health services.

The *Nursing Home Component* gathers information from 800 nursing homes and more than 5,000 residents nationwide on the characteristics of the facilities and services offered, individual resident expenditures and sources of payment, and resident characteristics. Data on the availability and use of community-based care prior to nursing home admission is also collected.

The *Medical Provider Component* covers 2,700 hospitals, nearly 2,000 physicians, and 300 home health care providers. Its purpose is to supplement information from respondents to the MEPS Household Component.

The *Insurance Component* consists of two surveys. The first survey interviews 9,200 employers, 300 union officials, and 400 insurers to obtain detailed information on the health insurance held by respondents to the MEPS Household Component. It also collects information about other health plans

Examples of Available Datasets

available to, but not chosen by, respondents. The second interviews managers at more than 20,000 establishments to obtain national and regional estimates of the availability of health insurance at the workplace.

Consumer Assessment of Health Plans (CAHPS) (Agency for Healthcare Research and Quality)

<https://www.cahps.ahrq.gov/default.asp>

Developed by a consortium of Harvard Medical School, RAND, and the Research Triangle Institute, and sponsored by the Agency for Healthcare Research and Quality (AHRQ), the CAHPS survey is designed to provide information that can help consumers and purchasers assess and choose among health plans. The CAHPS questionnaires are designed for three different types of target populations: commercially insured, Medicaid, and Medicare managed care. The CAHPS Survey includes a core group of standard items, a small group of supplementary items targeted towards specific populations, and a survey designed for those who disenroll from plans. Among the areas that are covered by the survey are: perceived quality of health care, perceived quality of health plan, administrative burden, enrollment/payment, utilization of health services, health status, and respondent characteristics. Information about specific plan features, such as access to specialists, quality of patient/physician communication and interaction, and coordination of care, is included, as well as questions targeted to persons with chronic conditions or disabilities, children, and Medicare and Medicaid beneficiaries.

Behavioral Risk Factor Surveillance System (Centers for Disease Control and Prevention)

<http://www.cdc.gov/brfss/>

The Behavioral Risk Factor Surveillance System (BRFSS) is one of several public health surveillance systems supported by the Centers for Disease Control and Prevention. The BRFSS collects uniform state-based data on preventive health practices and risk behaviors that are linked to chronic diseases, injuries, and preventable infectious diseases in the U.S. Data are collected through monthly telephone interviews conducted among a sample of each state's adult population. The survey consists of a standard set of questions asked by all states each year, an optional set of questions on specific topics which states may choose to include or not include on their questionnaires, and state-added questions which provide information on issues of local interest and emerging issues. The data are used to provide state and national information on access to preventive services such as breast, cervical, and colorectal cancer screenings and immunizations. Other health information includes high-risk behaviors, lack of physical exercise, obesity and data about smoking habits. Socio-demographic data, including age, education, income, and racial and ethnic background are also available.

Health Care Information System (Centers for Medicare & Medicaid Services)

http://www.cms.gov/NonIdentifiableDataFiles/02_HealthCareInformationSystem.asp

The Health Care Information System (HCIS) contains data from Medicare Part A (Inpatient, Skilled Nursing Facilities, Home Health Agencies (Part A & B) and Hospice) and Medicare Part B (outpatient)

Examples of Available Datasets

based on the type and state of the provider. The dataset includes discharges, total claim amount, total number of patients, total utilization days, and total visits.

Medicare Part B National Summary Data File (Centers for Medicare & Medicaid Services)

http://www.cms.gov/NonIdentifiableDataFiles/03_PartBNationalSummaryDataFile.asp

Medicare Part B covers doctors' services and outpatient hospital care for the Medicare population. The Medicare Part B National datasets contain data for a 5% sample of Medicare Part B claims submitted by physician/suppliers. The datasets are summarized by meaningful Health Care Common Procedure Coding/Current Procedural Terminology (HCPC/CPT) code ranges. Each dataset displays the allowed services, allowed charges and payment amounts by HCPC/CPT codes and prominent modifiers.

Surveillance, Epidemiology, and End Results (SEER) Registries (National Cancer Institute)

<http://seer.cancer.gov/>

Beginning in 1973, the Surveillance, Epidemiology, and End Results (SEER) Program of the National Cancer Institute (NCI) is a coordinated system of population-based cancer registries located across the United States. These registries monitor cancer trends and track data on cancer incidence, the extent of disease at diagnosis, therapy, and patient survival. The SEER Program currently collects cancer incidence and survival data from 18 geographic areas in the United States. here are nine states (New Mexico, Hawaii, Utah, Iowa, Connecticut, Greater California, Kentucky, Louisiana, New Jersey), five metropolitan areas (Metro Atlanta plus a sample of rural Georgia, the Greater Bay Area [San Francisco-Oakland and San-Jose Monterey], Los Angeles, Seattle, Detroit), and the Alaska Native Tumor Registry, which together represent approximately 26 percent of the U.S. population.

Employee Benefits Survey (Bureau of Labor Statistics)

<http://www.bls.gov/ncs/ebs/>

Produces comprehensive data on the incidence (the percentage of workers with access to and participation in employer provided benefit plans) and provisions of selected employee benefit plans.

Hospital Compare (Centers for Medicare & Medicaid Services)

https://www.cms.gov/HospitalQualityInits/11_HospitalCompare.asp

Hospital Compare is a consumer-oriented website that provides information on how well hospitals provide recommended care to their patients. On this site, the consumer can see the recommended care that an adult should get if being treated for a heart attack, heart failure, or pneumonia or having surgery. The performance rates for this website generally reflect care provided to all U.S. adults with the exception of the 30-Day Risk Adjusted Death and Readmission measures and the Hospital Outpatient Medical Imaging measures that only include data from Medicare beneficiaries. In March 2008, data from the Hospital

Last Updated May 3, 2011

Examples of Available Datasets

CAHPS (HCAHPS) survey, also known as the CAHPS Hospital Survey, was added to Hospital Compare. HCAHPS provides a standardized instrument and data collection methodology for measuring patient's perspectives on hospital care.

Nursing Home Compare (Centers for Medicare & Medicaid Services)

www.medicare.gov/nhcompare/

Nursing Home Compare allows Medicare beneficiaries and their caregivers to access comparison information about nursing homes. It contains information on every Medicare and Medicaid-certified nursing home in the country, including over 17,000 nationwide. Nursing Home Compare includes: 1) Nursing home characteristics such as number of beds, type of ownership and whether or not the nursing home participates in Medicare, Medicaid or both; 2) Resident characteristics including percent of residents with pressure sore, percent of residents with urinary incontinence and more; 3) Summary information about nursing homes during their last State inspection; 4) Information on the number of registered nurses, licensed practical or vocational nurses, and nursing assistants in each nursing home.

Recent Commonwealth Fund Surveys (data is available for Fellows)

Commonwealth Fund International Health Policy Survey (annual)

<http://www.commonwealthfund.org/Content/Surveys/2010/Nov/2010-International-Survey.aspx>

The International Health Policy Surveys are conducted annually in the Australia, Canada, France (since 2008), Germany (since 2005), the Netherlands (since 2006), Norway (since 2009), Sweden (since 2009), Switzerland (since 2010), the U.K., and the U.S. Past surveys have focused on the views and experiences of the general population, “sicker” adults, primary care doctors, elderly adults, and hospital executives. The 2010 survey focused on the experiences of adults in eleven countries. Topics included overall views of the health system, out-of-pocket spending and insurance complexity, care coordination, medical errors, and the doctor-patient relationship. The survey sample included 1,009 adults in Australia, 3,003 in Canada, 1,407 in Germany, 1,557 in the Netherlands, 1,000 in New Zealand, 1,434 in the United Kingdom, and 2,500 in the United States.

Commonwealth Fund Biennial Health Insurance Survey

<http://www.commonwealthfund.org/Content/Surveys/2011/Mar/2010-Biennial-Health-Insurance-Survey.aspx>

The biennial health insurance surveys provide important trend information on insurance coverage, access barriers, and satisfaction with and confidence in the health care system. In addition, the survey examines the impact of medical bill burdens and debt on family finances, and poor quality coverage on health. The 2010 survey included 4,005 U.S. adults.

Examples of Available Datasets

Commonwealth Fund Survey of Public Views of the U.S. Health Care System (2006, 2008, 2011)

<http://www.commonwealthfund.org/Content/Surveys/2011/Apr/Survey-of-Public-Views.aspx>

The 2011 Survey of Public Views of the Health System examined U.S. adults' health care experiences and views. Topics included access to care, cost barriers, care coordination, patient safety, and efficiency, as well as views on policies to reform the health system. The 2011 survey included 1,011 adults.

Commonwealth Fund Survey of Young Adults (2009)

<http://www.commonwealthfund.org/Content/Surveys/2009/Dec/The-Commonwealth-Fund-Survey-of-Young-Adults.aspx>

The 2009 Commonwealth Fund Survey of Young Adults examined health care insurance experiences and views among adults aged 19 to 29. The survey had a sample size of 2,002.

Commonwealth Fund National Survey of Federally Qualified Health Centers (2009)

<http://www.commonwealthfund.org/Content/Surveys/2010/May/The-2009-Commonwealth-Fund-National-Survey-of-Federally-Qualified-Health-Centers.aspx>

The 2009 Commonwealth Fund National Survey of Federally Qualified Health Centers examined access, care coordination, IT capacity, performance reporting, quality, and medical home status at Federally-Qualified Health Centers. The survey included 795 executive directors or clinical directors at federally qualified health centers.

Commonwealth Fund/Modern Healthcare Health Care Opinion Leaders Survey (quarterly)

<http://www.commonwealthfund.org/Content/Surveys/2011/Feb/Views-on-Congressional-Priorities-2011.aspx>

The Commonwealth Fund/Modern Healthcare Health Care Opinion Leaders Survey measures the views of a panel of health care leaders and experts on health care policies and priorities. The January 2011 survey had a sample size of 203.