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 (NIH) offers an online 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 (HHS) operates a website where users can search and download publicly available datasets and tools:

http://www.healthdata.gov/

Users can filter the available datasets by category, including costs, utilization, insurance type, and quality and safety.

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

Healthcare Cost and Utilization Project (HCUP) (Agency for Healthcare Research and Quality)
https://www.hcup-us.ahrq.gov/databases.jsp

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 two major components:

The *Household Component* is a household interview of over 12,000 families and 30,000 individuals 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. Data is also collected from a sample of providers (physicians, hospitals, home health agencies, and pharmacies) who provided medical care to MEPS Household Component respondents.

The *Insurance Component* is a survey of over nearly 40,000 employers on the number and types of private health insurance plans offered, benefits associated with these plans, annual premiums, annual contributions by employers and employees, eligibility requirements, and employer characteristics.

Examples of Available Datasets

Consumer Assessment of Health Plans (CAHPS) (Agency for Healthcare Research and Quality)
https://cahps.ahrq.gov

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 toward 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 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/Research-Statistics-Data-and-Systems/Files-for-Order/NonIdentifiableDataFiles/HealthCareInformationSystem.html

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

Examples of Available Datasets

(outpatient) 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/Research-Statistics-Data-and-Systems/Files-for-Order/NonIdentifiableDataFiles/PartBNationalSummaryDataFile.html

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. There 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)
http://www.medicare.gov/hospitalcompare/

Hospital Compare is a consumer-oriented website that provides information on how well hospitals provide recommended care to their patients. Hospital Compare allows users to select

multiple hospitals and directly compare performance measure information related to heart attack, heart failure, pneumonia, surgery, and other conditions.

Nursing Home Compare (Centers for Medicare & Medicaid Services)
http://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 15,000 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 sores, percent of residents with urinary incontinence, and more; 3) summary information about nursing homes during their last state inspection; 4) and information on the number of registered nurses, licensed practical or vocational nurses, and nursing assistants in each nursing home.

National Bureau of Economic Research (NBER)
http://nber.org/data/

NBER is a private, non-profit, non-partisan organization dedicated to conducting economic research and to disseminating research findings among academics, public policy makers, and business professionals. NBER-affiliated researchers study a wide range of topics and employ many different methods in their work. Key focus areas include developing new statistical measurements, estimating quantitative models of economic behavior, and analyzing the effects of public policies. NBER's publicly available datasets cover:

- Healthcare Data -- Hospitals, Providers, Drugs, and Devices
- Demographic and Vital Statistics Databases
- Individual Data:
    - National Health and Nutrition Examination Survey (NHANES)
    - National Ambulatory Medical Care Survey (NAMCS)
    - National Health Interview Survey (NHIS)

Examples of Available Datasets

Recent Commonwealth Fund Surveys (data is available for Harkness Fellows)

The Commonwealth Fund conducts surveys of health care consumers, health professionals, and health system leaders. The following are select examples of Fund surveys that may be of interest to Harkness Fellows. For more information on these and other Commonwealth Fund Surveys please visit the [Survey Data Center](#) where you can filter by series and search for key terms.

Commonwealth Fund International Health Policy Survey (annual)

- [2016 Commonwealth Fund International Health Policy Survey of Adults](#)
- [2015 Commonwealth Fund International Survey of Primary Care Physicians](#)
- [2014 Commonwealth Fund International Health Policy Survey of Older Adults](#)

Since 1998, the International Health Policy Surveys have been conducted annually in 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 2016 survey was of adults in 11 countries. Questions focused on people's experiences with their country's health care system in terms of access, quality, and affordability, as well as on self-reported health and well-being. The survey sample included 5,248 adults in Australia; 4,547 in Canada; 1,103 in France; 1,000 in Germany; 1,227 in the Netherlands; 1,000 in New Zealand; 1,093 in Norway; 7,124 in Sweden; 1,520 in Switzerland; 1,000 in the U.K; and 2,001 in the U.S.

Commonwealth Fund Biennial Health Insurance Survey

- [2016 Biennial Health Insurance Survey](#)
- [2014 Biennial Health Insurance Survey](#)
- [2012 Biennial Health Insurance Survey](#)

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 2016 survey included nationally representative sample of 6,005 adults ages 19 and older living in the continental United States.

Examples of Available Datasets

Commonwealth Fund Affordable Care Act Tracking Surveys (2013-2016)

- [Affordable Care Act Tracking Survey, February–April 2016](#)
- [Affordable Care Act Tracking Survey, March-May 2015](#)
- [Affordable Care Act Tracking Survey, April-June 2014](#)
- [Affordable Care Act Tracking Survey, July-September 2013](#)

Since its launch in 2013, The Commonwealth Fund's Affordable Care Act Tracking Survey has provided policymakers and the media with insights into the experiences of people with coverage through the marketplaces and Medicaid. The 2016 survey, the fourth in a series, included 4,802 adults.

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

- [2011 Survey of Public Views of the Health System](#)
- [2008 Survey of Public Views of the Health System](#)
- [2006 Survey of Public Views of the Health System](#)

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 Health Insurance Tracking Survey of U.S. Adults (2011)

- [2011 Tracking Survey of U.S. Adults](#)

The 2011 Commonwealth Fund Health Insurance Tracking Survey of U.S. Adults examined health insurance and health care experiences among a nationally representative sample of adults aged 19 to 64. The survey sample was drawn from a probability-based online panel that includes cell-phone only and low-income households typically difficult to reach using traditional telephone surveys and random digit dialing sampling. The survey had a sample size of 2,134, including 977 low-income adults who have incomes below 250% of the federal poverty level.

Commonwealth Fund Health Insurance Tracking Survey of Young Adults (2009, 2011)

- [2011 Tracking Survey of Young Adults](#)
- [2009 Tracking Survey of Young Adults](#)

Examples of Available Datasets

The 2011 Commonwealth Fund Health Insurance Tracking Survey of Young Adults examined health care insurance experiences and views among adults aged 19 to 29. The survey had a sample size of 1,863.

Commonwealth Fund National Survey of Federally Qualified Health Centers (2009, 2013)
- 2013 National Survey of Federally Qualified Health Centers
- 2009 National Survey of Federally Qualified Health Centers

The 2009 and 2013 Commonwealth Fund National Surveys of Federally Qualified Health Centers examined access, care coordination, IT capacity, performance reporting, quality, and medical home status at Federally-Qualified Health Centers. The 2013 survey included 679 executive directors or clinical directors at federally qualified health centers.