Research Data List

Tennessee Hospital Discharge Data

Inpatient & Outpatient Data (2003-2007)
How many appendectomies were performed in Tennessee hospitals? What was the average charge? How long was the average patient hospitalized? These questions and many other can be answered using data from the Hospital Discharge Data System or HDDS. HDDS receives information from UB-92 (HCFA-1450) forms on all inpatient discharges and other selected patient visits from Tennessee hospitals. Each form contains information on patient diagnoses, procedures performed on the patient, charges for services provided, and selected patient demographics. HDDS data can be used to learn about hospitalizations from hundreds of causes. Patients can be compared by sex and age. Different hospitals can be compared for usage and charges. Patterns of hospitalization can be compared in different parts of the State and among different segments of the population.

American Time Use Survey Data (2003-2007)

- The American Time Use Survey (ATUS) provides nationally representative estimates of how, where, and with whom Americans spend their time, and is the only federal survey providing data on the full range of nonmarket activities, from childcare to volunteering.
- ATUS data files are used by researchers to study a broad range of issues; the data files include information collected from over 85,000 interviews conducted from 2003 to 2008.
- ATUS data files can be linked to data files from the Current Population Survey (CPS). This expands the context in which time-use data can be analyzed and saves taxpayer money because fewer questions must be asked in the ATUS interview.


The Behavioral Risk Factor Surveillance System (BRFSS) is a state-based computer-assisted telephone interviewing effort conducted in cooperation with the Centers for Disease Control and Prevention. Questions are constructed to determine
the behaviors of individuals that will affect their risk of developing chronic diseases that may lead to premature mortality and morbidity. The data collected helps to identify high risk populations that can be targeted for intervention programs. The data can also be used to track changes over time of prevalence of risk factor behaviors and related diseases, and can assess the impact of health promotion and prevention intervention programs. Currently, every state in the country, the District of Columbia, and three U.S. territories are members of this surveillance system. The Tennessee Department of Health has been participating in this system on a continuing basis since 1984, surveying adults from randomly selected households throughout the state every month.

**NHIS Cancer Control Supplement (CCS)**

The NHIS CCS (also known as the Cancer Control Module) is administered every five years and focuses on issues pertaining to knowledge, attitudes, and practices in cancer-related health behaviors, screening, and risk assessment.

The CCS covers a variety of topics including, but not limited to:

- Diet and Nutrition,
- Physical Activity,
- Cancer Screening,
- Knowledge and Use of the Human Papillomavirus (HPV) Vaccine,
- Sun Avoidance,
- Tobacco Use and Control,
- Genetic Testing,
- Family History, and
- Cancer Risk Assessment.

**The Implicit Association Test (IAT) Data (2005)**

The IAT assesses strengths of associations between concepts by observing response latencies in computer-administered categorization tasks. In an initial block of trials, exemplars of two contrasted concepts (e.g., face images for the races Black and White) appear on a screen and subjects rapidly classify them by pressing one of two keys (for example, an e key for Black and i for White). Next, exemplars of another pair of contrasted concepts (for example, words representing positive and negative valence) are also classified using the same two keys. In a first combined task, exemplars of all four categories are classified, with each assigned to the same key as in the initial two blocks (e.g., e for Black or positive and I for White or negative). In a second combined
task, a complementary pairing is used (i.e., e for White or positive and i for Black or negative). In most implementations, respondents are obliged to correct errors before proceeding, and latencies are measured to the occurrence of the correct response. The difference in average latency between the two combined tasks provides the basis for the IAT measure. For example, faster responses for the {Black_positive/White_negative} task than for the {White_positive/Black_negative} task indicate a stronger association of Black than of White with positive valence.

The National Long Term Care Survey (1982-2004)

The NLTCS has completed six waves, nominally at five-year intervals, 1982, 1984, 1989, 1994, 1999, and 2004. The NLTCS is a nationally-representative sample both of the community and of institutional populations and is longitudinal in that sample persons age in once they reach 65 years of age and stay in the survey until they either die or are lost to follow-up. At each wave, a screener questionnaire is administered to the sample which divides the sample into three parts: the non-disabled (frequently called screen-outs), the disabled but living in the community, and the disabled living in an institution. About 5,000 people die between waves and are replaced by a sample of about that size of people who have become age 65 since the prior wave. Because of budget considerations it usually has not been possible to continue the entire non-disabled sample into the next wave. Instead a sample of the non-disabled is drawn to keep the total sample size for a wave at about 20,000. The sampling fraction for prior non-disabled participants has been frequently adjusted to compensate for low representation in some survey components, in particular the 75+ component, and the sample has been supplemented to increase the size of the 95+ component.

The NLTCS is a very data-rich resource with many components, including disability measures, medical conditions, attained education levels, and income. Numerous papers have used it as a source of data addressing a wide variety of topics related to aging and disability. The NLTCS website currently categorizes the varied types of publications under the headings of Active Life Expectancy, Activities of Daily Living, Aging, Assistive Devices, Caregiver Income, Cognitive Functioning, Disability Trends, Disease, Ethnicity, Family Support, Gender, Institutionalization, Instrumental Activities of Daily Living, Insurance, Mathematical Modeling, Medical Providers, Medicare and Medicaid, Military Service, Mortality, Paid Caregiver, and Unpaid Caregiver.

Health and Retirement Study (HRS)
The HRS is a national longitudinal study based on core biennial interviews of nearly 20,000 individuals representing the US population over age 50. Hispanics and African-Americans were over-sampled. Every six years, the HRS enrolls a new birth cohort in order to maintain a steady-state representation of the over 50 US population. Baseline interviews with existing birth cohorts were conducted in 1992, 1993, 1998, and 2004, with plans to enroll the next 6-year cohort in 2010. Participants are followed through the life course with biennial surveys and supplemental data collections. Data are de-identified and made publicly available at no charge to users.

**Surveillance, Epidemiology and End Results (SEER)**

The Surveillance, Epidemiology and End Results (SEER) Program, a premier source for cancer statistics in the United States. SEER collects information on incidence, survival, and prevalence from specific geographic areas representing 26 percent of the US population and compiles reports on all of these plus cancer mortality for the entire US.

Use SEER data (1973-2006) to address multiple issues:

- Examine stage at diagnosis by race/ethnicity.
- Calculate survival by stage at diagnosis, age at diagnosis, and grade or size of tumor.
- Determine trends and incidence rates of cancers at various sites over time.

**Medical Expenditure Panel Survey (1996-2007)**

MEPS currently has two major components for which data are released: the Household Component and the Insurance Component. The Household Component data are based on questionnaires fielded to individual household members and their medical providers. The Insurance Component estimates come from a survey of employers conducted to collect health insurance plan information. (In 1996 only, MEPS also included a Nursing Home Component that gathered information from a sample of nursing homes and residents. These data are only available through the AHRQ Data Center or through one of the Census Bureau's Research Data Centers.)

**The American Community Survey**

The American Community Survey (ACS) is an ongoing survey that provides data every year -- giving communities the current information they need to plan investments and services. Information from the survey generates data that help
determine how more than $400 billion in federal and state funds are distributed each year.

To help communities, state governments, and federal programs, we ask about:

- age
- sex
- race
- family and relationships
- income and benefits
- health insurance
- education
- veteran status
- disabilities
- where you work and how you get there
- where you live and how much you pay for some essentials

All this detail is combined into statistics that are used to help decide everything from school lunch programs to new hospitals.