National Cancer Institute

U.S. DEPARTMENT
OF HEALTH AND
HUMAN SERVICES

National Institutes
of Health

Division of Cancer Control and Population Sciences

DCCPS-Funded Public Use Data Resources

http://cancercontrol.cancer.gov

As NCI's bridge to public health research, practice, and policy, the Division of Cancer Control and Population Sciences (DCCPS) plays a unique role in reducing the burden of cancer in America. DCCPS has the lead responsibility at NCI for supporting research in surveillance, epidemiology, health services, behavioral science, and cancer survivorship. The division also plays a central role within the Federal Government as a source of expertise and evidence on issues such as the quality of cancer care, the economic burden of cancer, geographic information systems, statistical methods, communication science, tobacco control, and the translation of research into practice. DCCPS staff members are innovators in creating resources for the public and the research community. Here we provide a brief summary of public use data resources available through DCCPS.

Surveillance

Finding Cancer Statistics

http://surveillance.cancer.gov/statistics/

Recently developed to facilitate the use of cancer data, Finding Cancer Statistics is a plain-language Web site that provides access to recent reports, datasets, and statistical tools for professionals and the general public. It includes definitions of commonly used statistics, descriptions of datasets and tools, and guides to their use.

Surveillance, Epidemiology, and End Results (SEER) Program

http://seer.cancer.gov/

SEER Web site

The redesigned SEER Web site is the preferred mechanism for distributing most of the SEER Program's products, such as those listed below. Included on the site are the latest SEER Cancer Statistics Review, with a search function, and the Cancer Stat Fact Sheets, where users can obtain printouts of the most recent cancer statistics for each type of cancer.

SEER*Stat

SEER*Stat is a statistical software system for the analysis of SEER and other population-based cancer databases. The system provides an easy-to-use Microsoft Windows desktop package for viewing individual cancer records and producing statistics to assess the impact of cancer on populations. In addition, county attributes data can be analyzed with the SEER*Stat software. The most common uses would be to create a list of the county attribute data (such as median income values by county) using the case listing session, and to calculate incidence and mortality rates by county attributes using rate sessions. Tutorials are available.

SEER*Prep

The SEER*Prep system allows users to prepare and format their own cancer incidence, mortality, population, and expected survival rate data for use with SEER*Stat.

Fast Stats

Fast Stats is an interactive system with Java interface allowing users access to millions of precalculated cancer statistics. There is a choice of outputs, including vertical bar charts, line charts, tabular reports, and delimited data files. The types of analyses are being expanded, along with the usability of graphing functions.

Cancer Stat Fact Sheets

Cancer Stat Fact Sheets are a collection of statistical summaries for a number of common cancer types. They provide a quick overview of frequently requested cancer statistics. Available statistics may include incidence, mortality, survival, stage, prevalence, and lifetime risk. The default display contains links to additional resources and definitions of statistical terms, and a link is provided to a printable format. The statistics are updated annually to coincide with the latest SEER data release.

▶▶▶

September 2010

HD*Calc

The Health Disparities Calculator (HD*Calc) is an extension of SEER*Stat designed to generate multiple summary measures to evaluate and monitor health disparities. Cross-sectional and trend data (for example, cancer rates, survival, stage at diagnosis) categorized by disparity groups (such as area-socioeconomic status, race/ethnicity, geographic areas) can be used with HD*Calc to generate four absolute and seven relative summary measures of disparity. The results are displayed as tables and charts, which may be exported for use in other applications.

Small Area Estimates for Cancer Risk Factors & Screening Behaviors

Model-based estimates for states, counties, and health service areas have been developed based on two surveys, the Behavioral Risk Factor Surveillance System (BRFSS) and the National Health Interview Survey (NHIS). The two surveys are combined using novel statistical methodology. This yields estimates for states that strengthen BRFSS results with the higher response rate and coverage of non-telephone households available from NHIS and bias-corrected estimates for counties by utilizing data from both surveys and information from other areas with similar characteristics. Small area estimates are available for current smoking prevalence (Age 18+); ever smoking prevalence (Age 18+); mammography prevalence within 2 Years (Age 40+); and Pap smear test prevalence within 3 Years (Age 18+).

Colorectal Cancer Mortality Projections

The NCI's Cancer Intervention and Surveillance Modeling Network (CISNET) developed this Web site (http://cisnet.cancer.gov/colorectal) to help cancer control planners, program staff, and policy makers consider the impact of risk factor reduction, increased early detection, and increased access to optimal treatment on future colorectal cancer (CRC) mortality rates. Create charts and graphs projecting future trends in CRC mortality based on results from simulation models. Results are presented as the percent decline in mortality and deaths per 100,000. Use this information to see how policy options to increase cancer prevention, screening, and access to state-of-the-science treatment can affect future mortality trends; help determine cancer control program priority areas for new intervention investments; and identify research questions and opportunities.

National Health Interview Survey Cancer Control Supplement

http://appliedresearch.cancer.gov/surveys/nhis/

The National Health Interview Survey (NHIS) is an annual nationwide in-person survey of approximately 40,000 households conducted by the National Center for Health Statistics and administered by the U.S. Census Bureau. In 1987, 1992, 2000, 2005, and 2010, Cancer Control Supplements (CCS) were administered as part of the NHIS to adults aged 18 and older. The 2003 and 2008 surveys also included smaller item sets specifically focused on cancer screening.

The CCS includes questions about cancer risk factors such as diet, physical activity, and tobacco use, as well as cancer screening behaviors, which are used to determine knowledge, attitudes, and practices concerning cancer-related health behaviors in the general population. Data from the CCS are used by DCCPS and others to monitor national levels, trends, and determinants for use of preventive services.

California Health Interview Survey

http://appliedresearch.cancer.gov/surveys/chis/

The California Health Interview Survey (CHIS) provides population-based, standardized local health-related estimates based on approximately 50,000 households throughout California. Fielded for the first time in 2001, CHIS is a telephone survey that provides information on a wide range of health indicators, including physical and mental health status, prevalence and management of chronic diseases, cancer screening, diet, physical activity, health insurance coverage, and access to and utilization of preventive and other health services.

CHIS is designed to be comparable to the NHIS, and one of the major strengths of CHIS is that it provides local-level estimates for key health and behavioral indicators comparable to the NHIS. Other unique qualities of CHIS are that it is administered in four different languages and collects sufficient data for analyses on certain sub-populations, such as American Indians, Asians, and Pacific Islanders. NCI uses CHIS data to learn more about cancer screening behaviors and cancer risk factors. CHIS is also used by local health departments, community health centers, health plans, and other entities throughout California to identify and address state and local health problems and needs.

National Health and Nutrition Examination Survey

http://riskfactor.cancer.gov/studies/nhanes/

The National Health and Nutrition Examination Survey (NHANES) is 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. NHANES is a major program of the National Center for Health Statistics (NCHS).

DCCPS has provided funds to support modules in NHANES that are critical to its risk factor monitoring mission, such as a food frequency questionnaire (FFQ) and a wearable physical activity monitor. Both of these modules were incorporated in NHANES from 2003–2006. Currently, DCCPS staff are collaborating with NCHS to develop a series of Web-based tutorials to encourage and support analyses of NHANES data. Additionally, DCCPS worked with NCHS to add a dietary screener to NHANES 2009–2010.

Tobacco Use Supplement to the Current Population Survey

http://riskfactor.cancer.gov/studies/tus-cps/

The Tobacco Use Supplement to the Current Population Survey (TUS-CPS) is an NCI-sponsored survey of tobacco use and policy information that has been administered as part of the U.S. Census Bureau's and the Bureau of Labor Statistics' (BLS) Current Population Survey. The TUS-CPS has been fielded since 1992, typically triennially and currently in 2010–2011. All the questionnaires and data are publicly available, with most recent data from 2006–2007. The TUS-CPS is a key source of national, state, and some local-level data on smoking, other tobacco use, policy, and other intervention information in U.S. households because it uses a large, nationally representative sample. It contains information on approximately 240,000 individuals within a given survey period. Researchers can use the data to monitor long-term progress in the control of tobacco use, conduct tobacco-related research, evaluate tobacco control programs, and examine tobacco use-related health disparities. The TUS-CPS is translated into Spanish, and the 2003 instrument was also translated into four Asian languages. Unique features are the ability to link TUS-CPS data to other social and economic U.S. Census Bureau and BLS data, as well as other sponsor-supported supplement data; cancer and other cause-specific mortality data through the National Longitudinal Mortality Study (http://surveillance.cancer.gov/disparities/nlms/); and occasional TUS-CPS 1-year prospective follow-up data on a subset of respondents (e.g., 2002–2003 and 2010–2011). In addition, the Robert Wood Johnson Foundation's ImpacTEEN program made extensive use of the TUS-CPS 15-year period data in its recent tobacco chart book, entitled Cigarette Smoking Prevalence and Policies in the 50 States: An Era of Change (available for download at http://www.impacteen.org/tobaccodata.htm).

Causes of Cancer

Genomic Datasets

http://epi.grants.cancer.gov/dac/

Investigators can apply for controlled access to datasets from cancer genome-wide association studies (GWAS), sequencing, and other genomic datasets. Additional datasets will be added as they become available.

Cancer Family Registries (CFRs)

http://epi.grants.cancer.gov/CFR/

The Breast Cancer Family Registry (B-CFR) and Colon Cancer Family Registry (C-CFR) are international research infrastructures for investigators interested in conducting population and clinic-based interdisciplinary studies on the genetic and molecular epidemiology of breast and colon cancer and their behavioral implications. A central goal of the CFRs is the translation of this research to the clinical and prevention setting for the benefit of Registry participants and the general public. The CFRs have information and biospecimens contributed by families across the spectrum of risk for these cancers and from population-based or relative controls.

Cancer Genetics Network

http://epi.grants.cancer.gov/CGN/

The Cancer Genetics Network (CGN) is a resource for investigators conducting research on the genetic basis of human cancer susceptibility; integration of this information into medical practice; and behavioral, ethical, and public health issues associated with human genetics. The CGN can provide a wide variety of research services and specialized expertise to assist investigators with approved studies. Prospective investigators can freely query the CGN core database to learn more about the aggregate characteristics of participants and discover how the CGN may be used for research purposes.

Geographic Information System for Breast Cancer Studies on Long Island

http://li-gis.cancer.gov/default.html

The Epidemiology and Genetics Research Program has developed a Geographic Information System for Breast Cancer Studies on Long Island (LI GIS). The LI GIS provides researchers a unique tool with which to investigate potential relationships between environmental exposures and risk for breast cancer. It potentially can be used for research on other types of cancer and other diseases.

Quality of Care

SEER–Medicare Datasets

http://healthservices.cancer.gov/seermedicare/

The SEER-Medicare data reflect the linkage of two large population-based sources of data that provide detailed information about Medicare beneficiaries with cancer. The data come from the SEER program of cancer registries that collect clinical, demographic, and cause of death information for persons with cancer and the Medicare claims for covered health care services from the time of a person's Medicare eligibility until death.

These combined datasets can be used for an array of studies, including

- Assessing patterns of care for persons with cancer
- Use of tests and procedures during the period prior to and following a cancer diagnosis
- Determining costs of cancer treatment

This linkage of the SEER data with Medicare claims is an important part of cancer health services and outcomes research.

Behavioral Research

Health Information National Trends Survey

http://hints.cancer.gov/

The Health Information National Trends Survey (HINTS) is a nationally representative, biennial survey of approximately 8,000 randomly selected adults. Data are available—from 2003, 2005, and 2007—to allow data and results users to gain insight into people's knowledge about cancer, the communication channels through which they obtain health information, their attitudes, knowledge, and perceptions of cancer, and their cancer-related behaviors. The survey

- Provides updates on changing patterns, needs, and information opportunities
- Identifies changing communications trends and practices
- Assesses cancer information access and usage
- Provides information about how cancer risks are perceived

The HINTS Web site expands access to HINTS data for multiple audiences, using tables, charts, population estimates, reports, and technical history information on every question in the survey. New tools within the site were developed with extensive input from Federal and private partners in cancer communication research and practice.

Survivorship

Cancer Survivor Prevalence Data

http://survivorship.cancer.gov/prevalence/index.html

To better understand the demographics of the U.S. population of cancer survivors, the Office of Cancer Survivorship (OCS) and the Surveillance Research Program worked together to develop survivorship prevalence estimates based on the Surveillance, Epidemiology, and End Results (SEER) registry database, which represents five states (Connecticut, Hawaii, Iowa, New Mexico, and Utah) and four standard metropolitan statistical areas (Detroit, Atlanta, San Francisco–Oakland, and Seattle–Puget Sound).

In an effort to make survivorship prevalence estimates accessible and easy to understand, OCS displays graphical representations of the data on a Web page dedicated to the topic. Population-based estimates for the composition of current cancer survivors are derived by gender, cancer site, age, race, and time since diagnosis. The data are updated annually.

Dissemination and Diffusion

Cancer Control P.L.A.N.E.T.

http://cancercontrolplanet.cancer.gov/

Cancer Control P.L.A.N.E.T. (Plan, Link, Act, Network with Evidence-Based Tools) is a Web portal that provides easy access to data and research-based resources that can help state and local cancer control program planners and staff and cancer prevention and control researchers to design, implement, and evaluate evidence-based cancer control programs.

Cancer Trends Progress Report

http://progressreport.cancer.gov/

The Cancer Trends Progress Report summarizes our nation's progress against cancer in relation to the Healthy People 2010 targets developed by the U.S. Department of Health and Human Services. First issued in 2001 as the Cancer Progress Report, the online report is released every other year. The information presented in this report—intended for policy makers, researchers, clinicians, and public health service providers—is gathered through a collaborative effort with other key agencies and groups, such as the Centers for Disease Control and Prevention and the American Cancer Society.