Surveillance, Epidemiology, and End Results (SEER)

Surveillance Research Program

50TH ANNIVERSARY SEER
TURNING CANCER DATA INTO DISCOVERY

SEER is supported by the NCI's Surveillance Research Program (SRP), which provides national leadership in the science of cancer surveillance as well as analytical tools and methodological expertise in collecting, analyzing, interpreting, and disseminating reliable population-based statistics.
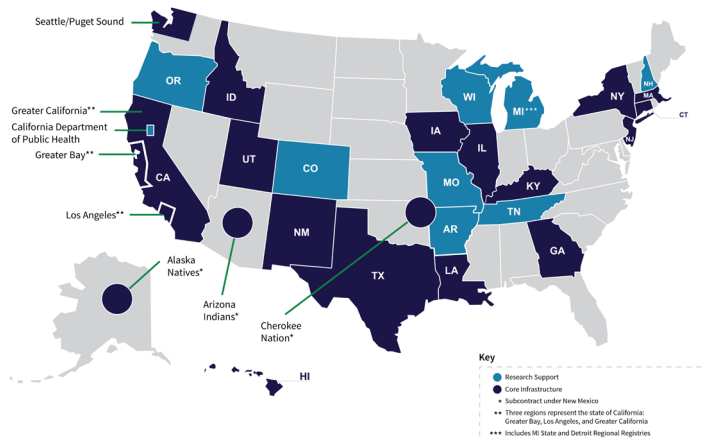
What Is SEER?

The SEER Program provides information on cancer statistics to help reduce the cancer burden among the US population. SEER has been funded by the NCI since 1973.

SEER collects and publishes cancer incidence and survival data from population-based cancer registries covering approximately 48% of the US population. These registries routinely collect data on patient demographics, primary tumor site, tumor morphology and stage at diagnosis, first course of treatment, and follow-up for vital status (survival). These data are collected on every cancer case reported from 22 US geographic areas.

There are 18 CORE registries (Greater California; Greater Bay, comprised of San Francisco-Oakland and San Jose-Monterey; Los Angeles; Connecticut; Georgia; Hawaii; Idaho; Illinois; Iowa; Kentucky; Louisiana; Massachusetts;

New Jersey; New Mexico, including subcontracts with Alaska Natives, Arizona Indians, and Cherokee Nation [Oklahoma]; New York; Seattle-Puget Sound; Texas; and Utah) and 10 Research Support registries (Arkansas, California Department of Public Health, Colorado, Michigan, Detroit, Missouri, New Hampshire, Oregon, Tennessee, and Wisconsin).

These areas are representative of the demographics of the entire US population. This broad coverage allows SEER to account for diverse populations throughout the US, including 44.4% of whites, 44.6% of Black people, 69.2% of Hispanics, 55.7% of American Indian/Alaska Natives, 72.2% of Asians, and 73.5% of Native Hawaiian/Pacific Islanders.

Who Uses SEER Data?

SEER data are used by thousands of researchers, clinicians, cancer registrars, public health officials, policymakers, community groups, and members of the public.

How Are SEER Data Used?

- Evaluate cancer prevention and screening programs and the quality of cancer care
- Document disparities by race/ethnicity, gender, geography, and other variables
- Demonstrate the effectiveness of public health interventions
- Guide the translation of research into health policy and practice

"The SEER Program is one of NCI's most important data collection and dissemination activities. In addition to providing essential information for tracking the Nation's progress against cancer, SEER data and data analysis tools provide researchers with unique opportunities to explore and explain cancer trends. The impact of SEER on science, policy, and practice reflects both the quality of the data collected and the creative expertise of the many scientists who use it."

Robert T. Croyle, PhD
Director, Division of Cancer Control and Population Sciences, NCI

As a research resource, SEER serves as a platform for studies that address emerging issues in the field of cancer and cancer-related care. SEER also plays a critical role in increasing the breadth of scientific research to include various factors such as health disparities, access to and quality of health care, geographic determinants, and other attributes.

SEER Online Resources

The SEER website (https://seer.cancer.gov) offers resources geared toward different audiences. These include cancer statistics summarized in a variety of formats, as well as SEER data and associated software.

Summaries of Cancer Statistics

- *SEER Cancer Statistics Review (CSR)* — annual report of the most recent cancer incidence, mortality, survival, prevalence, and lifetime risk statistics

- *Annual Report to the Nation on the Status of Cancer* — annual update on rates of new cancer cases and deaths as well as trends in the most common cancers

- *SEER Bibliography* — searchable database of SEER publications produced by registry and program staff

- *SEER Monographs* — special journal issues and reports on special topics

- *Fact Sheets* — cover specific cancer types

- *State Cancer Profiles* — characterize the cancer burden in a standardized manner to motivate action, integrate surveillance into cancer control planning, characterize areas and demographic groups, and expose health disparities; CDC and NCI co-sponsor this website

- *SEER*Explorer* — interactive, easy access to wide range of SEER cancer statistics

- *Fast Stats* — build your own tables and graphs of key SEER and US cancer statistics

- *Videos* — highlight key topics and trends in cancer statistics

- *Geographic Information Systems* — portal enabling interactive mapping and visualization of cancer-related geospatial data

- *Cancer Trends Progress Report* — summarizes our nation's progress against cancer in relation to Healthy People targets

- *Know Your Chances* — interactive risk charts to put probabilities (risk) of dying from cancer and other diseases in context

- *Prevalence and Cost of Cancer Care* — cancer prevalence and costs estimated and projected by tumor site through 2020

- *Cancer Query Systems* — provide access to cancer statistics in online databases

- *CI*Rank* — confidence intervals for ranks of age-adjusted US cancer incidence and mortality rates as well as confidence intervals for ranks of age-adjusted US mortality rates for other causes of death

- *Small Area Estimates for Cancer-Related Measures* — access to NCI small area estimation projects; the measures pertain to risk factors (e.g., smoking prevalence), screening, policies, and knowledge

SEER Data and Software

The SEER research data include SEER incidence and population data associated by age, sex, race, year of diagnosis, and geographic areas. Options for accessing and requesting the data, plus statistical software and documentation, are available at https://seer.cancer.gov.

Datasets

- *SEER Incidence Data* — associated by age, sex, race, year of diagnosis, and geographic areas
- *SEER-linked Datasets* — SEER-Medicare, SEER-Medicare Health Outcomes Survey (SEER-MHOS), and SEER-Consumer Assessment of Healthcare Providers and Systems (SEER-CAHPS)
- *Specialized Datasets* — apply for access to variables not included in the standard SEER research data
- *Other Datasets*
  - US Mortality — the data include all causes of death, not just cancer
  - US Populations — county population estimates used to calculate cancer incidence and mortality rates
  - Standard Populations — age distributions used as weights to create age-adjusted statistics

Software

- *SEER*Stat* — for analyzing SEER data and other cancer-related databases; powerful tool for studying the impact of cancer on a population
- *SEER*Prep* — converts user-supplied ASCII text data files to SEER*Stat database format
- *SEER Data Management System (SEER*DMS)* — supports all core cancer registry functions: importing data, editing, linkage, consolidation, and reporting
- *Health Disparities Calculator (HD*Calc)* — generates multiple summary measures to evaluate and monitor health disparities

- *ComPrev* — calculates complete prevalence estimates, based on limited-duration prevalence statistics calculated on SEER cancer data
- *DevCan* — computes the probability of being diagnosed or dying of cancer
- *Joinpoint* — software for analyzing trends

Additional Resources

Visit the SEER website (https://seer.cancer.gov) to view resources and training content for cancer registrars, funding opportunities, SEER news items, explanations of cancer statistics and surveillance, SRP's Twitter handle, and a blog describing SRP initiatives and collaborations, methods, and technologies, including those focused on SEER.

SEER Research Initiatives

SEER-focused initiatives include those described below. Visit SRP's Research Areas page (https://surveillance.cancer.gov/research) to learn more.

Department of Energy (DOE) Collaboration — NCI is partnering with the DOE on pilot efforts that will impact future cancer research and guide advances in scientific computing. The objectives are to enhance the SEER data to better characterize each cancer patient and provide information to support understanding of the care trajectory. A major component of this Pilot process is working to develop NLP and machine learning tools to automate the capture of data to overcome challenges in precision oncology at the molecular, patient, and population levels.

Natural Language Processing (NLP) pilot studies — These studies are evaluating NLP tools focusing on specific data elements (e.g., biomarkers, recurrence) to extract information automatically from pathology reports and other clinical documents. They are part of broader goals to enhance cancer surveillance.

The SEER Program and DOE's Oak Ridge National Laboratory are co-leading the population-level pilot (Pilot 3) as an important component of the NCI's Cancer Moonshot[SM]. This large pilot will address the growing cancer surveillance challenges in capturing essential information for understanding the effectiveness of cancer diagnosis and treatment. The objective is to deliver an infrastructure that will support the development of algorithms, informatics tools, and predictive models to enable an enhanced and expanded national cancer surveillance program.

SEER-linked Virtual Tissue Repository (VTR) Pilot Study — This effort was initiated in 2015 with seven SEER registries to identify the best practices for a scaled Virtual Biorepository, including multiple SEER registries. The initial pilot focuses on two important clinical outcomes representing extreme survivors: breast cancer patients with localized node-negative disease who died within 2 years of diagnosis and long-term (> 5 years) survivors of pancreatic adenocarcinoma. For these cases and matched controls, the pilot will obtain custom annotations of detailed treatment data and other relevant clinical information such as body mass index and biomarker levels for pancreas and female breast cancer cases who may have biospecimens available. The goal is to match cases with exceptional survival and controls with more typical survival. Whole exome sequencing for the pancreatic specimens (supported by a partnership with PanCan) will be linked with the clinical annotation and made available to the larger research community.

SEER Rapid Response Surveillance Studies (RRSS) — These studies address issues in cancer control by collecting biological materials and information from medical record and central cancer registry data reviews. They also focus on making data collection more efficient.

Cancer Intervention and Surveillance Modeling Network (CISNET) — This consortium uses simulation modeling to examine the impact of prevention, screening, and treatment on cancer incidence and mortality. These models can help us understand current trends, project future trends, and determine optimal cancer control strategies. SEER data are integral to CISNET activities, e.g., the breast cancer-specific models are designed to match incidence and mortality rates observed in SEER.

Electronic Pathology (E-path) Software — SRP has been working to expand and enhance e-path reporting from its 300+ network of pathology labs, which report in near-real time to their associated SEER registries. This effort supports important activities at the registry level, such as assisting researchers in recruiting patients to clinical trials and other studies.

SEER Treatment Data Expansion — SRP is spearheading efforts to capture more complete treatment information on traditional and orally administered cancer therapies through the linkage of data from important collaborations with external partners. These include pharmacy organizations for receiving orally administered treatment information as well as claims data from oncology practices to capture longitudinal information on initial and subsequent therapies provided in the oncology practice setting.

SEER Quality Improvement (QI) — The QI team leads a variety of efforts to maintain and improve the quality of SEER data. Visit https://seer.cancer.gov/qi/ to learn more.

Contact Information

SEER Program
Surveillance Research Program
Division of Cancer Control and Population Sciences, National Cancer Institute

9609 Medical Center Drive, MSC 9765
Bethesda, MD 20892 (US Mail)
Rockville, MD 20850 (Express Mail)

Telephone: 240-276-6740
Fax: 240-276-7908
Email: seerweb@imsweb.com