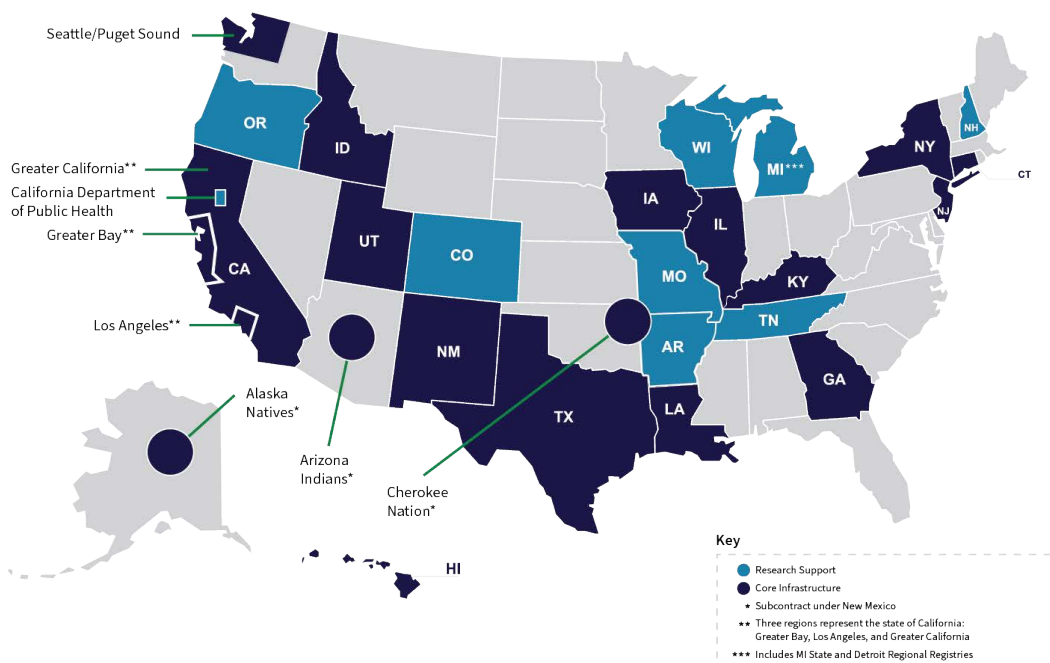# Surveillance, Epidemiology, and End Results (SEER)

Surveillance Research Program

SEER is supported by the National Cancer Institute's (NCI) 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?

Map of the United States showing SEER registries. Labels include: Seattle/Puget Sound, Greater California**, California Department of Public Health, Greater Bay**, Los Angeles**, Alaska Natives*, Arizona Indians*, Cherokee Nation*. State abbreviations shown: OR, ID, WI, MI***, NY, NH, CT, NJ, IA, IL, UT, CO, MO, KY, CA, NM, AR, TN, GA, TX, LA, HI.

Key:
- Research Support
- Core Infrastructure
- * Subcontract under New Mexico
- ** Three regions represent the state of California: Greater Bay, Los Angeles, and Greater California
- *** Includes MI State and Detroit Regional Registries

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.

Over 2 million people will be diagnosed with cancer in 2025, and the population of cancer survivors in the US will exceed 22 million by 2035. SEER collects and publishes cancer incidence and survival data from population-based cancer registries covering 45.9% 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 17 Core registries (**Greater California**; **Greater Bay**, *composed of San Francisco-Oakland and San Jose-Monterey*; **Los Angeles**; **Connecticut**; **Georgia**; **Hawaii**; **Idaho**; **Illinois**; **Iowa**; **Kentucky**; **Louisiana**; **New Jersey**; **New Mexico**, *including subcontracts with Alaska Natives, Arizona Indians, and Cherokee Nation [Oklahoma]*; **New York**; **Seattle-Puget Sound**; **Texas**; and **Utah**) and 11 Research Support registries (**Arkansas**, **California Department of Public Health**, **Colorado**, **Massachusetts**, **Michigan**, **Detroit**, **Missouri**, **New Hampshire**, **Oregon**, **Tennessee**, and **Wisconsin**).

Surveillance, Epidemiology, and End Results (SEER)

These areas are representative of the demographics of the entire US population. This broad coverage allows SEER to account for populations throughout the US, including 39.6% of Whites, 43.5% of Blacks, 64.9% of Hispanics, 59.3% of American Indian/Alaska Natives, 68.2% of Asians, and 69.9% of Native Hawaiian/Pacific Islanders.

Who Uses SEER Data?

SEER data are used by **hundreds of thousands** of people each month, including researchers, clinicians, cancer registrars, public health officials, community groups, and members of the public.

How Are SEER Data Used?

- Evaluate cancer prevention and screening programs and the quality of cancer care

- Study trends in cancer data overall and across populations

- Demonstrate the effectiveness of public health interventions

> *"The SEER Úrogram tracks the Nation's progress against cancer and is one of NCI's most important research investments. SEER data and data analysis tools are used by thousands of researchers to explore and explain cancer incidence, mortality, and survival trends. **The exceptional data quality and the expertise of the scientists who use SEER help us to understand the burden of cancer in the US population**."*
>
> **Katrina A.B. Goddard, 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, like **increasing rates of early-onset cancer**, the effect of **healthy lifestyle choices on lowering cancer risk and increasing patient survival**, and more.

Examples of Cancer Statistics Resources

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

- *NCCR*Explorer*: incidence and survival statistics for ages 0 to 39 years using data from the National Childhood Cancer Registry

- *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

- *Cancer Stat Facts:* covers basic statistics of most types of cancer

- *State Cancer Profiles:* characterizes the cancer burden in a standardized manner at the state and county level to integrate surveillance into cancer control planning and understand how cancer affects different areas of the US; CDC and NCI co-sponsor this website

- *Did You Know? 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

- *Oral Cancer Survivor Calculator*: tool to estimate oral cancer survival based on characteristics of squamous cell carcinoma and overall health

- *CI*Rank*: ranked incidence and mortality rates by state, county, and special region, presented along with confidence intervals

- *Small Area Estimates for Cancer-Related Measures:* model-based estimates related to risk factors, screening, policies, and knowledge at the state and county levels.

Surveillance, Epidemiology, and End Results (SEER)

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), SEER-Medicaid, 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 Drivers Calculator (HD*Calc)*: generates multiple summary measures to evaluate and monitor health drivers

- *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

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 MoonshotSM. 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)

The SEER-Linked VTR Program uses the SEER public health cancer surveillance registries as honest brokers to supply researchers and other investigators with access to de-identified, linked data and archival, formalin-fixed, paraffin-embedded (FFPE) tissue obtained through clinical care.

Unlike a physical biorepository, which collects and stores tissue for future use in cancer research, a VTR uses honest brokers to fulfill requests for de-identified, linked tissue, whole slide images, pathology reports, and/or clinical data for secondary use in research and other projects.

The SEER-Linked VTR Program accepts requests for formalin-fixed, paraffin-embedded (FFPE) tissue obtained through clinical care for secondary use in research and other projects. While tissue blocks are not released to investigators, tissue, processed as de-identified unstained or stained slides, ribbons, or scrolls, is released to investigators. Other types of data resources that can be requested include de-identified pathology reports, clinical and demographic data, and/or whole slide images (WSIs).

Virtual Pooled Registry- Cancer Linkage System (VPR-CLS) Project

This system is a web-based, streamlined process for researchers to conduct minimal-risk linkage studies with the nationwide resource of Central Cancer Registries. A total of 45 registries are currently enrolled in the VPR. In collaboration with the North American Association of Central Cancer Registries (NAACCR) and Information Management Services (IMS), NCI created this system to quickly provide researchers with initial match counts and then minimize the complexity of the application process formerly necessary in state-by-state applications.

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.

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, and a blog describing SRP initiatives and collaborations, methods, and technologies, including those focused on SEER.

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