New Grantee Workshop
September 16, 2011

SEER Data and Analytic Tools

Kathy Cronin

Data Analysis & Interpretation Branch

Surveillance Research Program
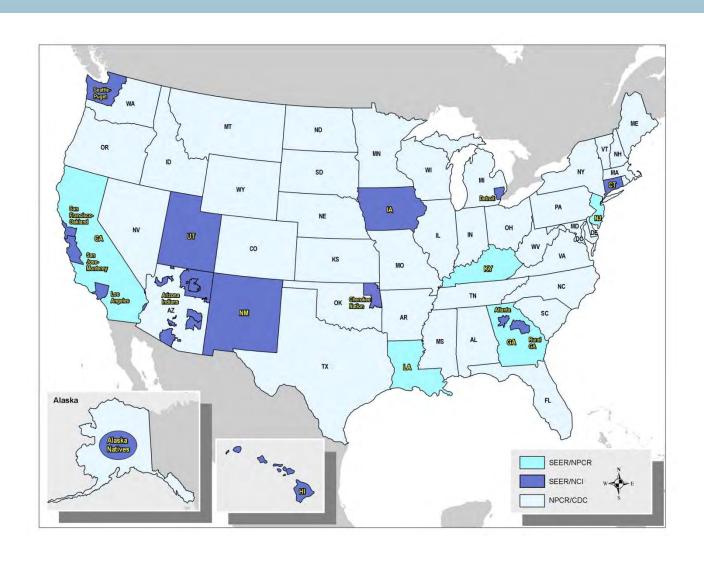
What Is SEER?

➢ NCI's **Surveillance, Epidemiology, and End Results (SEER) Program** gathers information on cancer incidence and survival in the United States.

➢ SEER collects and publishes incidence and survival data from population-based cancer registries that cover about 28% of the US population

➢ SEER Registries also collect data on demographics, primary tumor site, morphology and stage at diagnosis, and follow-up for vital status.

➢ SEER data are used by researchers, clinicians, public health officials, legislators, policymakers, community groups, and the public**.**
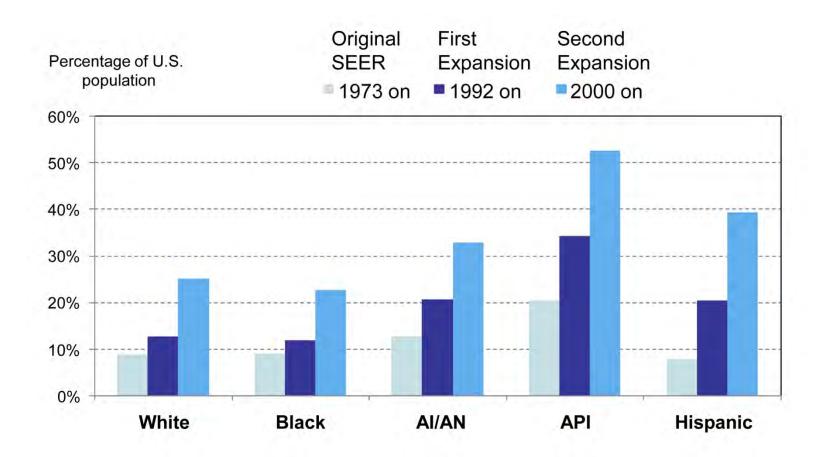
U.S. Map with SEER Registries Shown

Evolution of SEER Over Time

➢ SEER 9 covering years 1975+

  ▪ San Francisco-Oakland, Connecticut, Detroit, Hawaii, Iowa, New Mexico, Seattle, Utah, Atlanta

➢ SEER 13 covering years 1992+

  ▪ SEER 9 plus San Jose-Monterey, Los Angeles, Rural Georgia, Alaska Natives

➢ SEER 17 covering years 2000+

  ▪ SEER 13 plus California (excluding SF – O/SJM/LA), Kentucky, Louisiana, New Jersey

Population Coverage by Race/Ethnicity (2005 est.)

AI/AN: American Indian and Alaska Native
API: Asian and Pacific Islander

SEER Data Quality

- Quality improvement has been an integral part of SEER Program activities since its inception

- Data quality is monitored and improved through rigorous quality control studies and various data assessments

- Study results improve documentation and provide a focus for educational programs

Where Are SEER Statistics Reported?

➢ Cancer Statistics Review – Most recent cancer incidence, mortality, survival, prevalence and lifetime risk statistics

➢ Annual Report to the Nation

➢ Fact Sheet – Summary of key statistics by cancer site

➢ Fast Stats – Quick way to generate tables and graphs

➢ State Cancer Profiles

http://www.seer.cancer.gov/statistics/

Size of SEER: Rare Cancers, Cancer Heterogeneity

- The four most common cancers (lung, colorectal, breast and prostate) comprise approximately 50% of the cancer burden

- However, >50 other cancers exist
  - Lead to considerable morbidity and mortality

- Beyond primary cancer sites, ~300 anatomic subsites and ~500 histologic subtypes

- SEER database is a resource with adequate numbers to evaluate this detail

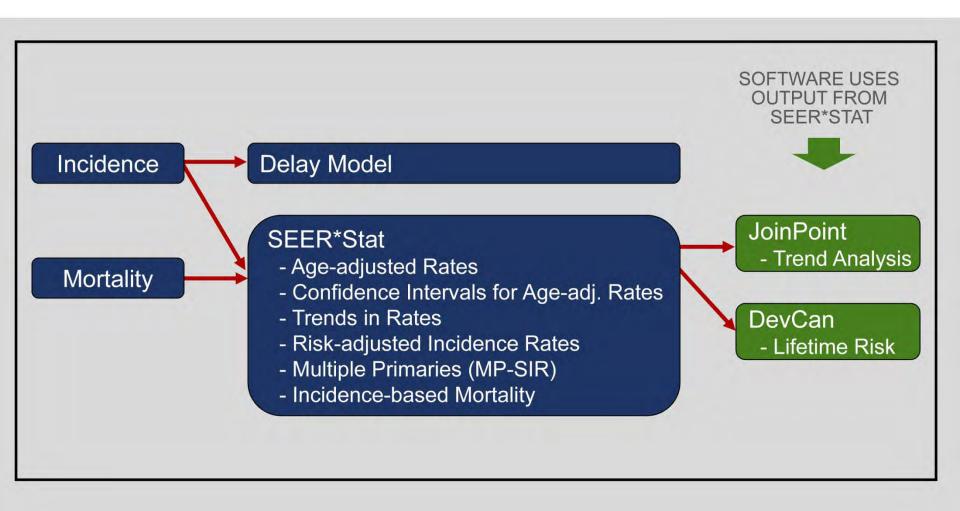# Methods and Software for Population Based Cancer Statistics

SEER*STAT

➤ Includes SEER database, US mortality files, and population data from the Census

➤ Individual level records with all the detail possible from publicly available files

➤ Analyses can run from simple to complex

➤ Six session types
  - Frequencies
  - Rates
  - Survival
  - Prevalence (Limited Duration)
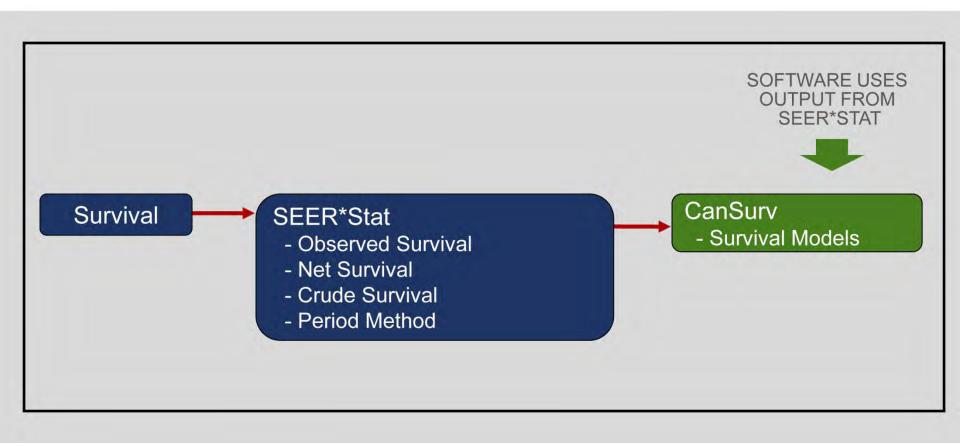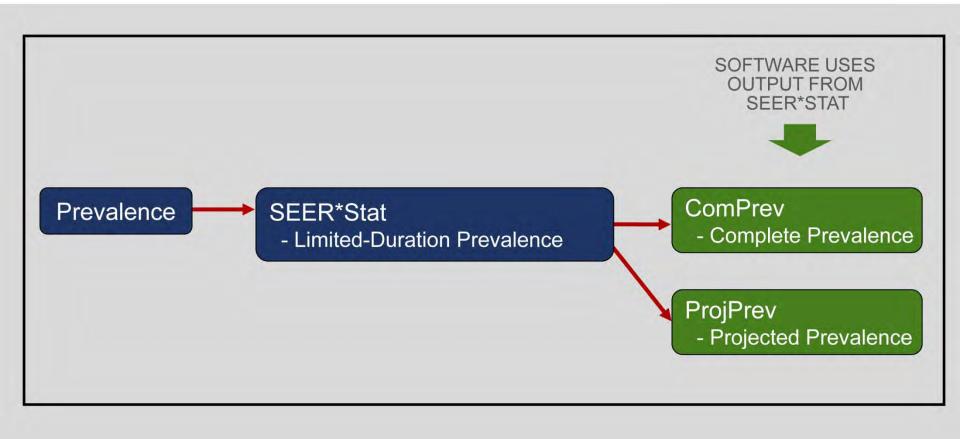 - Multiple Primaries - Standardized Incidence Ratio (MPSIR)
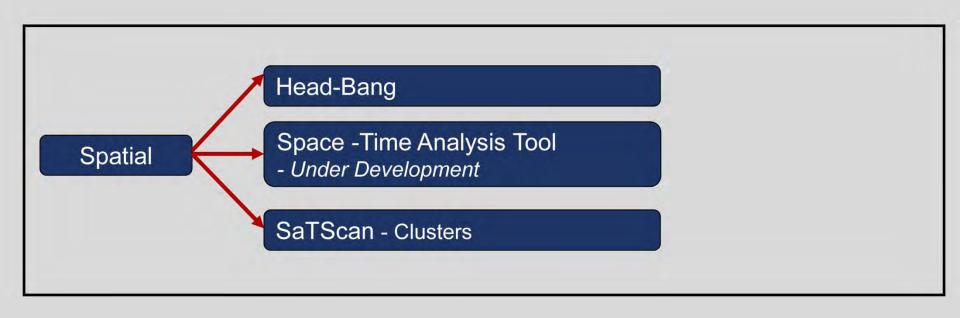 - Case-Listing

Survival

Prevalence

Mapping and Spatial Analysis

Health Disparities Calculator (HD*Calc)

http://www.seer.cancer.gov/hdcalc/index.html

➢ The Health Disparities Calculator (HD*Calc) is statistical software designed to generate multiple summary measures to evaluate and monitor health disparities (HD).

➢ HD*Calc was created as an extension of SEER*STAT that allows the user to import SEER data or other population-based health data and calculate any of eleven disparity measurements.