

# Enhancing SEER Data: Linkages to Other Data Resources

Joan Warren, Ph.D.  
Applied Research Program  
National Cancer Institute



DCCPS New Grantee Workshop  
September 16, 2011

# Data that have been linked to SEER

The SEER data have been enhanced by linking to other data sources. These are the:

- SEER-Medicare database\*
- SEER- National Longitudinal Mortality Study (NLMS) database\*\*
- SEER-Medicare Health Outcomes Survey (MHOS) database\*

\*Data are available for release to investigators

\*\*Full access is available to SEER-NLMS via Special Sworn Status from Census Bureau

# What are the SEER-Medicare data?

- NCI takes people in the SEER data and matches them to Medicare's master enrollment file
- The data include:
  - 100% of patients in the SEER data who are found to be Medicare eligible (currently over 1.5 million cancer patients age 65+)
  - 5% random sample of people residing in the SEER areas who have not been diagnosed with cancer (control group)

## What is included in the SEER-Medicare data?

- The data include all SEER records for cases and claims for all Medicare covered health services for cases and controls
- SEER data include incidence, site, stage, initial treatment, demographics and vital status, date and cause of death
- Medicare claims are available for:
  - Short stay hospitalizations
  - Physician and lab services
  - Hospital outpatient claims
  - Home health and hospice bills
  - Part D prescription drug
- These data are longitudinal

# Why Link the SEER-Medicare Data?

The linked data can be used for a number of analyses that span the course of cancer control activities

Screening/ Detection	Diagnosis/ Treatment	Survivorship	Second Cancers	Terminal Care/ Death
Use of PSA testing: mammography, sigmoidoscopy/ colonoscopy  Impact of new technology or practice patterns on cancer detection/ incidence	Patterns of care Peri-operative complications Volume outcomes studies Extent of staging Comorbidities	Late effects of treatment Post- diagnostic surveillance Treatment of prevalent cancers	Rates of second primaries Relationship of second events to initial treatment and ongoing surveillance	Use of hospice End-of-life patterns of care Survival

← Health disparities, quality of care and cost of treatment →

# What is the National Longitudinal Mortality Survey (NLMS)?

- The NLMS combines socio-demographic data collected by the Current Population Surveys (CPS) with cause of death information from official state records
- Includes cohorts who took the CPS from 1979-2002
- CPS variables include race/ethnicity, marital status, education, income, employment status, occupation, household size, immigrant status, smoking (for certain cohorts), health status, and health insurance status
- The NLMS database includes approximately 3 million persons and cause of death information on over 250,000

## Linkage of the SEER and NLMS Data

- The NLMS data have been linked to SEER (piggybacking on the SEER-Medicare linkage)
- The SEER-NLMS data includes 35,839 people diagnosed with cancer through 2003
- A planned update of SEER-NLMS through 2009 will increase the number of cases in these datasets

# Potential Uses of SEER-NLMS data

These linked data can be used to assess:

- Socioeconomic determinants of mortality or incidence of major cancers
- Self-reported race and ethnic associations with mortality or incidence of major cancers
- Cancer mortality or incidence by employment status
- Educational or income disparities in cancer mortality, incidence, survival, stage, tumor characteristics
- Marital status and household structure associations with cancer mortality, incidence, survival, stage, and tumor characteristics
- Association of insurance status with cancer mortality, incidence, survival, stage, tumor characteristics



# What is the Medicare Health Outcomes Survey (MHOS)

- The MHOS is a questionnaire administered to a random sample of 1,000 beneficiaries from each managed care organization that participates in Medicare
- The MHOS survey includes 95 questions related to demographics, socio-economic status, health problems and Activities of Daily Living (ADLs). Includes the SF-36 with major focus on health related quality of life (HRQOL)
- The survey was first administered in May 1998 with follow-up of the same cohort in 2000
- A new cohort is selected each year for baseline measurement and a 2-year follow-up assessment

# What is SEER-MHOS ?

- Patients who are in the MHOS have been linked to SEER data (piggybacking on the SEER-Medicare linkage)
- Contains linked records for beneficiaries with both SEER and MHOS data (n=55,000) as well as those without a SEER reported cancer diagnosis but with MHOS data (n=800,000)
- Includes patient-reported outcomes measures for the years 1998-2005

## Potential Uses of the SEER-MHOS data

Researchers can use these data to:

- Evaluate HRQOL before/after a cancer diagnosis and compare cancer survivors with beneficiaries never diagnosed with cancer
- Evaluate HRQOL and receipt of initial cancer treatment, especially with respect to surgical interventions and radiation
- Assess the effects of different types of cancers on ADLs

## For more information

- SEER-Medicare database  
<http://healthservices.cancer.gov/seermedicare/>
- SEER- National Longitudinal Mortality Study (NLMS) database  
<http://surveillance.cancer.gov/disparities/nlms/>
- SEER-Medicare Health Outcomes Survey (MHOS) database  
<http://outcomes.cancer.gov/surveys/seer-mhos/>