Surveillance, Epidemiology and End Results (SEER) Program

Institute of Medicine
May 9, 2011
Brenda K. Edwards
Surveillance Research Program, NCI
Public Law 92-218, the National Cancer Act of 1971 directs NCI to

“collect, analyze, and disseminate all data useful in the prevention, diagnosis, and treatment of cancer...”
SEER: Monitoring Progress in Reducing the Cancer Burden

• Longitudinal surveillance
  – All sites, common or rare
  – All populations
    • by age, sex, race & ethnicity, geography
    – Evaluating the impact of public health interventions

• Unusual cancer patterns
  – Rapid changes in incidence
    • Relevance to etiology
    • Relevance to public health
US Population-Based Registries

• NCI SEER Program – Established 1973
  – Collects cancer incidence and survival data from population-based cancer registries covering approximately 28% of the U.S. population.
  – Collects data on patient demographics, primary tumor site, tumor morphology and stage at diagnosis, first course of treatment, and follow-up for vital status.

• NPCR – Established 1992
  – Collects cancer incidence data from 45 states, the District of Columbia, and Puerto Rico and is supporting registry development in the U.S. Pacific Island Jurisdictions, representing 96% of the U.S. population.
  – Collects data on the occurrence of cancer; the type, extent, and location of the cancer; and the type of initial treatment.
National Surveillance Partners

• **NCDB – Established 1989**
  - Hospital-based registries report data to the central registries.
  - 70% of all newly diagnosed cases of cancer in the U.S. are captured at the institutional level and reported to the NCDB.

• **NAACCR – Established 1989**
  - Establishes and maintains a consensus on standards for cancer registration.
  - Certifies registries that achieve a standard of high quality for producing an incidence rate.

• **ACS**
  - The largest non-government funder of cancer research.
  - Publishes *Cancer Facts & Figures* annually, based on registry data and NCI methodology for prediction of cases.
United States improved coverage for population-based cancer incidence

- **NAACCR 1998-2007** 84%
- **NAACCR 2003-2007** 93%
- **SEER 9: 1975-2007** (9.4%)
- **SEER 13: 1992-2007** (14%)
- **SEER 17: 2000-2007** (26%)
- **NAACCR: 2003-2007** (93%)
- **USCS: 2007 (12/10)** (99%)
## Aggregate Data Availability

<table>
<thead>
<tr>
<th>Source</th>
<th>Years Covered</th>
<th>US Population Coverage</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEER</td>
<td>1975 – 2008 (33 years)</td>
<td>10%</td>
</tr>
<tr>
<td>United States Cancer Statistics (NCI, CDC, NAACCR)</td>
<td>1998 – 2007 (10 years)</td>
<td>84%</td>
</tr>
<tr>
<td>United States Cancer Statistics (NCI, CDC, NAACCR)</td>
<td>2003 – 2007 (5 years)</td>
<td>97%</td>
</tr>
<tr>
<td>United States Cancer Statistics (NCI, CDC, NAACCR)</td>
<td>2004 – 2006 (3 years)</td>
<td>100%</td>
</tr>
</tbody>
</table>
Annual Report to the Nation on the Status Of Cancer, 1975-2007, Featuring Tumors of the Brain and Other Nervous System

Betsy A. Kohler, MPH, CTR; Elizabeth Ward, PhD; Bridget J. McCarthy, PhD; Maria J. Schymura, PhD; Lynn A. G. Ries, MS; Christie Eheman, PhD; Ahmdedin Jemal, DVM, PhD; Robert N. Anderson, PhD; Umed A. Ajani, MBBS, MPH; Brenda K. Edwards, PhD

Rates of new diagnoses and rates of death from all cancers combined declined significantly in the most recent time period for men and women overall and for most racial & ethnic US populations

- Incidence: - 0.8 % per year from 2003-2007
- Deaths: - 1.6% per year from 2003-2007
## U.S. Cancer Burden 2010 Estimates

Based on data from NAACCR 1995-2006 & CDC NCHS 1969-2007; NCI statistical models

### 1,529,560 Cases
- 222,520 Lung
- 207,090 Breast (female)
- 217,730 Prostate
- 142,570 Colon/rectum
- 70,530 Bladder
- 68,130 Melanoma
- 65,540 NHL
- 58,240 Kidney
- 44,670 Thyroid
- 43,470 Corpus uteri
- 43,140 Pancreas
- 43,050 Leukemia

### 569,490 Deaths
- 157,300 Lung
- 51,370 Colon/rectum
- 39,840 Breast (female)
- 36,800 Pancreas
- 32,050 Prostate
- 21,840 Leukemia
- 20,210 NHL
- 18,910 Liver & IHBD
- 14,680 Bladder
- 14,500 Esophagus
- 13,850 Ovary
- 13,140 Brain

A National Framework for Cancer Surveillance

Descriptions of the Cancer Burden Nationally, Regionally, and in States & Communities

Measurement of Race, Ethnicity, Socioeconomic Status & Culture; Costs; Individual, Social & Biologic Factors; and Provider Knowledge, Attitudes, & Practices


Who uses population-based cancer data?

- Researchers
- Clinicians
- Legislators
- Health planners
- Public health officials
- Communities
- Patients & their families
SEER Reporting Timeline

• In April 2011, SEER released cases diagnosed through 2008
• 3-year time lag in reporting represents major challenge in quickly identifying cases
• Ongoing efforts to reduce the time between diagnosis and reporting of new cases
• In the future, registries may be used to identify cases associated with specific exposure through data linkages
SEER’s Multi-Tiered Data Collection

Stage, Histology, Treatment
- SEER Extent of Disease
- Collaborative Stage v.2
- AJCC Manual 7th Edition
- ePath

Demographics
- Patient record
- IHS & other linkages
- Hispanic & API algorithms

Survival
- Active follow-up
- State Vital Records, MVA
- Passive follow-up: NDI, SSA, CMS
Data Dissemination

**Studios**
- Rapid Response Special Studies
- Patterns of Care Studies
- Investigator-initiated studies

**Online Resources**
- State Cancer Profiles
- Cancer Fact Sheets
- Fast Stats
- Cancer Statistics Review
- Cancer Query System
- Population Data

**Data Analysis Tools**
- SEER*Stat
- SEER*Stat Bridge
- SEER*Prep
- Joinpoint
- Delay Adjustment
- VA Case Adjustment
- CanSurv
- ComPrev
- HD*Calc
- DevCan
- ProjPrev

**Reports**
- Annual Report to the Nation (With CDC, ACS, NAACCR, others)
- SEER Cancer Statistics Review
- US Cancer Statistics (with CDC)
- Cancer Facts & Figures (with ACS)
- Cancer in North America (with NAACCR)
- Cancer in Five Continents (with WHO/IARC)
- Cancer Trends Progress Report
- Monographs

**Research Databases & Linkages**
- SEER*Stat
- Public-Use Data
- SEER-Medicare
- National Longitudinal Mortality Study
- Residual Tissue Repository
Cancer Incidence of All Sites 1992-2007

SEER(13) Incidence Delay-adjusted Trends

- Both Sexes:
  - -3.1% (1992-1994)
  - 0.4% (1994-1999)
  - -0.8% (1999-2007)

- Male trends:
  - -4.4% (1992-1995)
  - 0.1% (1995-2001)
  - -1.9% (2001-2005)
  - 0.3% (2005-2007)

- Female trends:
  - 0.8% (1992-1998)
  - -0.6% (1998-2007)

Rates are age-adjusted to the 2000 U.S. standard million population.
Source: Incidence data – NCI SEER Program; Report to the Nation, JNCI 2011.
SEER Observed Incidence, SEER Delay Adjusted Incidence and US Death Rates
Cancer of the Lung and Bronchus, by Race and Sex

White Male

Black Male

White Female

Black Female

Year of Diagnosis/Death

Year of Diagnosis/Death

Year of Diagnosis/Death

Year of Diagnosis/Death

Rate per 100,000

Rate per 100,000

Rate per 100,000

Rate per 100,000

SEER Incidence APCs
Delay Adj, 1984-08 = -1.9*
Observed, 1985-08 = -2.0*

SEER Incidence APCs
Delay Adj, 1990-08 = -1.7*
Observed, 1990-08 = -1.8*

SEER Incidence APCs
Delay Adj, 1990-08 = 0.5*
Observed, 1990-08 = 0.4

SEER Incidence APCs
Delay Adj, 1997-08 = 0.2
Observed, 1997-08 = 0.0

Source: SEER 9 areas and US Mortality Files (National Center for Health Statistics, CDC).
Rates are age-adjusted to the 2000 US Std Population (19 age groups - Census P25-1103).
Regression lines and APCs are calculated using the Joinpoint Regression Program Version 3.5, April 2011, National Cancer Institute.
The APC is the Annual Percent Change for the regression line segments. The APC shown on the graph is for the most recent trend.
* The APC is significantly different from zero (p < 0.05).
SEER Observed Incidence, SEER Delay Adjusted Incidence and US Death Rates
Cancer of the Lung and Bronchus, Ages 65 and Over, by Race and Sex

White Male
- SEER Incidence APCs
  - Delay Adj, 1992-08 = -1.2*
  - Observed, 1992-08 = -1.2*

Black Male
- SEER Incidence APCs
  - Delay Adj, 1997-08 = 1.2*
  - Observed, 1997-08 = 1.0*

White Female
- SEER Incidence APCs
  - Delay Adj, 1992-08 = -1.1*
  - Observed, 1992-08 = -1.1*

Black Female
- SEER Incidence APCs
  - Delay Adj, 1991-08 = 1.9*
  - Observed, 1991-08 = 1.8*

* Source: SEER 9 areas and US Mortality Files (National Center for Health Statistics, CDC).
Rates are age-adjusted to the 2000 US Std Population (10 age groups - Census P25-1103).
Regression lines and APCs are calculated using the Joinpoint Regression Program Version 3.5, April 2011, National Cancer Institute.
The APC is the Annual Percent Change for the regression line segments. The APC shown on the graph is for the most recent trend.
* The APC is significantly different from zero (p < 0.05).
Lung Cancer Incidence

SEER Incidence

From 2004-2008, the median age at diagnosis for cancer of the lung and bronchus was 71 years of age\textsuperscript{3}. Approximately 0.0% were diagnosed under age 20; 0.2% between 20 and 34; 1.6% between 35 and 44; 8.8% between 45 and 54; 20.9% between 55 and 64; 31.1% between 65 and 74; 29.0% between 75 and 84; and 8.3% 85+ years of age.

The age-adjusted incidence rate was 62.0 per 100,000 men and women per year. These rates are based on cases diagnosed in 2004-2008 from 17 SEER geographic areas.

<table>
<thead>
<tr>
<th>Race/Ethnicity</th>
<th>Male</th>
<th>Female</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Races</td>
<td>75.2 per 100,000 men</td>
<td>52.3 per 100,000 women</td>
</tr>
<tr>
<td>White</td>
<td>75.3 per 100,000 men</td>
<td>54.6 per 100,000 women</td>
</tr>
<tr>
<td>Black</td>
<td>99.8 per 100,000 men</td>
<td>54.7 per 100,000 women</td>
</tr>
<tr>
<td>Asian/Pacific Islander</td>
<td>53.2 per 100,000 men</td>
<td>28.5 per 100,000 women</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>51.2 per 100,000 men</td>
<td>39.5 per 100,000 women</td>
</tr>
<tr>
<td>Hispanic</td>
<td>39.6 per 100,000 men</td>
<td>24.5 per 100,000 women</td>
</tr>
</tbody>
</table>
Lung Cancer Survival

Survival & Stage

Survival can be calculated by different methods for different purposes. The survival statistics presented here are based on relative survival, which measures the survival of the cancer patients in comparison to the general population to estimate the effect of cancer. The overall 5-year relative survival for 2001-2007 from 17 SEER geographic areas was 15.6%. Five-year relative survival by race and sex was: 13.7% for white men; 18.3% for white women; 11.6% for black men; 14.5% for black women.

<table>
<thead>
<tr>
<th>Stage at Diagnosis</th>
<th>Stage Distribution (%)</th>
<th>5-year Relative Survival (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Localized (confined to primary site)</td>
<td>15</td>
<td>52.0</td>
</tr>
<tr>
<td>Regional (spread to regional lymphnodes)</td>
<td>22</td>
<td>24.2</td>
</tr>
<tr>
<td>Distant (cancer has metastasized)</td>
<td>56</td>
<td>3.6</td>
</tr>
<tr>
<td>Unknown (unstaged)</td>
<td>7</td>
<td>8.1</td>
</tr>
</tbody>
</table>

The stage distribution is based on Summary Stage 2000. (See Fast Stats for more detailed statistics)
Cancer Staging: Timelines

Derived Staging, based on EOD or CS:

- **SEER Historic Stage**: 1973+
- **SEER Summary Stage 77**: 1988+
- **SEER AJCC 3rd ed.**: 1988+
- **AJCC 5th ed.**: 1998+
- **SEER Summary Stage 2000**: 1998+
- **AJCC 6th ed.**: 2004
- **AJCC 7th ed.**: 2010

**AJCC**: American Joint Committee on Cancer
**EOD**: Extent of Disease
**CS**: Collaborative Stage
Cancer Prevalence

• A statistic of primary interest in public health
• Identifies the level of burden of disease or health-related events on the population and health care system
• A function of both the incidence of the disease and survival
Cancer Prevalence (cont’d)

**Limited-Duration Prevalence** = the proportion of people alive on a certain day who had a diagnosis of the disease within the past \( x \) years.

**Complete Prevalence** = the proportion of people alive on a certain day who previously had a diagnosis of the disease, regardless of how long ago the diagnosis was, or if the patient is still under treatment or is considered cured.
Cancer Prevalence and Cost of Care Projections

Graph by Cancer Site & Phase of Care

Cancer Site
- Lung

Gender
- Male and Female

Age
- All Ages

End Year
- 2020

Incidence and Survival Assumptions
- Incidence: Constant (2003-05 average rate)
- Survival: Constant (2005 rate)

Cost Assumptions
- 0% rise in medical costs per year (2010 dollars)

Graph
- Cost of Cancer Care
- Prevalence

Cost of Cancer Care by Phase of Care, Lung, All Ages, Male and Female, in 2010 Dollars

Phase of Care
- Initial
- Continuing
- Last

Cost ($ per Year (in millions))

0 2,000 4,000 6,000 8,000

2010
2020
Registries in Epidemiologic Research

- Ecological studies; hypothesis generation
- Case-control studies
- Notification of research opportunities
  - E.g. Passive recruitment for cohort studies
- Recruitment for trials?
- GIS – census or environmental databases
- Record linkage – prospective studies
- Case-validation
Registries in Epidemiologic Research

• Population mobility can be an issue:
  – Studies on less mobile populations have used registries
    • Agricultural Health Study
    • NIH AARP-Diet & Health Study
    • California Teachers Study
  – Studies on younger, more mobile populations have used medical records
    • Nurses Health Study
• Lung cancer is a disease of the elderly, a less mobile population
SEER & Post-Marketing Surveillance

- More drugs/products have postmarketing surveillance requirements
- Central cancer registries provide population-based data in a standard format
- Rapid response surveillance studies have a quicker turn-around
SEER & Post-Marketing Surveillance

- Quantify cancer incidence and prevalence
- Identify cases for further study of risk factors
- Provide sufficient cases to distinguish background rates
- Identify “missed” cancer cases
- Confirm cancer diagnoses
Strengths and Limitations of Post-Marketing Surveillance using SEER Rates

• Provides rates for the general population
  – To compare with adverse events would need person years at risk

• Does not account for other differences in the patients compared to the general population
  – Other risk factors, etc.

• Provides background to detect a signal for further investigation, does not replace the need for signal evaluation and validation
Other Data Sources

- SEER-Medicare
- Cancer Research Network
- Other HMO databases
  - NIH HMO Research Network “Collaboratory”
- National surveys