NCI Surveillance Research Program

SEER, Standards for Collection of Race/Ethnicity, Measuring Health Disparities in Cancer Surveillance

IOM Committee on Future Directions for NHQR and NHDR

Feb 9, 2009
Brenda K. Edwards, PhD
Stephen Meersman, PhD
Barry Miller, DrPH

Surveillance Research Program, DCCPS
NCI
NCI Surveillance/SEER Outline

• National Surveillance Framework for Cancer

• NCI Surveillance, Epidemiology, and End Results (SEER) Program: 35+ years

• Cancer rates and trends by race/ethnicity

• Methods to improve reporting cancer statistics by race/ethnicity

• Social determinants of disease & health disparities

• Questions & discussion
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


Major Surveillance Partners

Federal Agencies

Cancer Incidence & Survival
- NCI SEER Program (1973)

Cancer Incidence
- CDC NPCR (1992)
- IHS (AI/AN data linkage)

Mortality
- CDC NCHS (1930s)

Population Denominators
- Bureau of the Census (1800s)

Professional & Private

- ACoS COC (1937): American College of Surgeons Commission on Cancer
- ACS (1950s): American Cancer Society
- IACR (1976): International Association of Cancer Registries
Monitoring the Impact of Cancer

• Important for ongoing surveillance
  – All sites, common or rare
  – All populations, by age, sex, race, geography
• Identifying unusual patterns
  – Rapid changes in incidence
    • Relevance to etiology
    • Relevance to public health
      – Planning
      – Evaluating the impact of public health interventions
SEER Geographic Coverage

Original SEER Registries

SEER Registries Added in 1992

SEER Registries Added in 2000

- Seattle/ Puget Sound
- Metropolitan Detroit
- Connecticut
- New Jersey
- Atlanta
- Los Angeles
- San Francisco/ Oakland
- San Jose/ Monterey
- CA
- UT
- NM
- IA
- LA

SEER: AK Native Tumor Registry Added in 1999

SEER: AZ American Indians added in 1980
Population Coverage by Race/Ethnicity (2005 est.)

Percentage of U.S. population

Original SEER 1973 on
First Expansion 1992 on
Second Expansion 2000 on

White 25%
Black 10%
AI/AN 25%
API 50%
Hispanic 20%

AI/AN: American Indian and Alaska Native
API: Asian and Pacific Islander
Racial/Ethnic Coverage in SEER

- **Total Population**: 26.2%
- **Latino**: 40.4%
- **Native Hawaiian / Pacific Islander**: 69.8%
- **Asian**: 53.3%
- **American Indian / Alaska Native**: 42.2%
- **Black**: 22.7%
- **White**: 23.4%
Characteristics of the SEER Population Compared With the Total United States Population
(Source: 2000 Census)

SEER areas included in this figure are the States of Connecticut, Hawaii, Iowa, Louisiana, New Jersey, New Mexico, Utah; multi-county areas of Atlanta & rural Georgia, Detroit, San Francisco-Oakland, Seattle-Puget Sound, San Jose-Monterey, Los Angeles county, remaining counties of California, and American Indians/Alaska Natives in Arizona and Alaska.
U.S. Cancer Registries

SEER\textsuperscript{1} Program (1973+)

NPCR\textsuperscript{2} (1995+)

NAACCR\textsuperscript{3} Certification
High Quality Data (5 years)\textsuperscript{4}

\textsuperscript{1}Surveillance, Epidemiology, and End Results, National Cancer Institute
\textsuperscript{2}National Program of Cancer Registries, Centers for Disease Control and Prevention
\textsuperscript{3}North American Association of Central Cancer Registries
\textsuperscript{4}Registries meeting NAACCR standards of data quality for combined 1995-99 data (2002 Report to the Nation; 53% of U.S. population)
United States
Improved coverage for population-based cancer incidence

NAACCR 2001-2005 80%
NAACCR 1996-2005 65%
SEER 1975-2005 10%

SEER 9: 1975-2005 (9.4%)
SEER 13: 1992-2005 (14%)
SEER 17: 2000-2005 (26%)
NAACCR: 2001-2005 (80%)
USCS: 2004 (10/07) (98%)
Treatment facilities identify cancer cases

Population-based cancer registries collect, verify & consolidate cancer cases and deaths

National organizations certify & analyze data for public health, research, planning & policy use

Hospitals

Labs

Radiology Facilities

Cancer Incidence Surveillance - Working Together

National Cancer Institute

North American Association of Central Cancer Registries

Centers for Disease Control and Prevention
Rates are age-adjusted to the 2000 U.S. standard million population. Sources: Incidence data – NCI SEER Program; Mortality data – CDC NCHS NVSS
Male Lung & Bronchus Cancer
SEER Incidence (delay adjusted) & US Death Rates
1975-2005

Incidences

- Blacks
- Whites
- Asian/Pacific Islander
- Hispanic

Mortality

- Decline & Gap is Closing
- Black Males: - 2.7% (1994-2005)
- White Males: - 1.8% (1991-2005)

Rates are age-adjusted to the 2000 U.S. standard million population. Sources: Incidence data – NCI SEER Program; Mortality data – CDC NCHS NVSS
Female Lung & Bronchus Cancer
SEER Incidence (delay adjusted) & US Death Rates
1975-2005

Rates are age-adjusted to the 2000 U.S. standard million population. Sources: Incidence data – NCI SEER Program; Mortality data – CDC NCHS NVSS
Female Breast Cancer
SEER Delay-Adjusted Incidence Rates- 4JP
1975-2005

Recent sharp decrease for white females
-2.2% 1999-2005

Decrease not seen for black females
Rates stable since 1992

Rates are age-adjusted to the 2000 U.S. standard million population.
Source: Incidence data – NCI SEER Program
Prostate Cancer
SEER Delay-Adjusted Incidence Rates – 4 JP Males, 1975-2005

White men
+16.3% (1988-1992)
-12.5% (1992-1995)
+2.6% (1995-2001)
-4.5% (2001-2005)

Rates for black men much higher. Similar increase & decrease
Recent trends:
22.1% (1989-1992)
-5.9% (1992-1995) NS
-0.2% (1995-2002) NS
-4.8% (2002-2005)

Rates are age-adjusted to the 2000 U.S. standard million population.
Source: Incidence data – NCI SEER Program
National Cancer Surveillance - Data Sources

- Incidence & survival (SEER) (NAACCR, CDC-NPCR)
- Mortality (CDC/NCHS)
- Population (Census)
- SEER/Medicare
- Screening & health risk behaviors (CDC/BRFSS)
- Environmental exposures (USGS, EPA et al.)

Reports
Analyses
Maps
Projections
Annual Report to the Nation (ARN)

- Update on Cancer in the United States
- Special topic:
  - Lung Cancer
  - Tobacco Use & Control
- Collaboration since 1998
  - American Cancer Society (ACS)
  - National Cancer Institute (NCI)
  - Centers for Disease Control and Prevention (CDC)
  - North American Association of Central Cancer Registries (NAACCR)

Incidence Trends, All Cancers, by Race/Ethnicity and Sex, 1996-2005

* Asterisk indicates statistically significant trend (SEER and NPCR) from NAACCR
U.S. Mortality Trends, All Cancers, by Race/Ethnicity, Sex, 1996-2005†

- Hispanic: Males -1.9*, Females -1.1*
- AI/AN‡: Males -0.7, Females -0.2
- API: Males -2.0*, Females -1.1*
- Black: Males -2.5*, Females -1.2*
- White: Males -1.5*, Females -0.9*

* Asterisk indicates statistically significant trend
† Age-adjusted death rates, CDC, National Center for Health Statistics
‡ Rates for American Indians and Alaska Natives restricted to CHSDA Counties
Health Disparities Findings

- People from low income populations are often diagnosed at later stages with less favorable outcomes
- Need for better cancer prevention and early detection programs
Differences across Groups

- Cancer burden differs across racial/ethnic groups
- Increasing diversity in US
  - ~30% population = Asian or Latino
<table>
<thead>
<tr>
<th>Code</th>
<th>Race Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>01</td>
<td>White</td>
</tr>
<tr>
<td>02</td>
<td>Black</td>
</tr>
<tr>
<td>03</td>
<td>American Indian, Aleutian, Alaskan Native or Eskimo (includes all indigenous populations of the western hemisphere)</td>
</tr>
<tr>
<td>04</td>
<td>Chinese</td>
</tr>
<tr>
<td>05</td>
<td>Japanese</td>
</tr>
<tr>
<td>06</td>
<td>Filipino</td>
</tr>
<tr>
<td>07</td>
<td>Hawaiian</td>
</tr>
<tr>
<td>08</td>
<td>Korean (Effective with 1/1/1988 dx)</td>
</tr>
<tr>
<td>09</td>
<td>Asian Indian, Pakistani (Effective with 1/1/1988 dx)</td>
</tr>
<tr>
<td>10</td>
<td>Vietnamese (Effective with 1/1/1988 dx)</td>
</tr>
<tr>
<td>11</td>
<td>Laotian (Effective with 1/1/1988 dx)</td>
</tr>
<tr>
<td>12</td>
<td>Hmong (Effective with 1/1/1988 dx)</td>
</tr>
<tr>
<td>13</td>
<td>Kampuchean (including Khmer and Cambodian) (Effective with 1/1/1988 dx)</td>
</tr>
<tr>
<td>14</td>
<td>Thai (Effective with 1/1/1994 dx)</td>
</tr>
<tr>
<td>20</td>
<td>Micronesian, NOS (Effective with 1/1/1991)</td>
</tr>
<tr>
<td>21</td>
<td>Chamorran (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>22</td>
<td>Guamanian, NOS (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>25</td>
<td>Polynesian, NOS (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>26</td>
<td>Tahitian (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>27</td>
<td>Samoan (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>28</td>
<td>Tongan (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>30</td>
<td>Melanesian, NOS (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>31</td>
<td>Fiji Islander (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>32</td>
<td>New Guinean (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>96</td>
<td>Other Asian, including Asian, NOS and Oriental, NOS (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>97</td>
<td>Pacific Islander, NOS (Effective with 1/1/1991 dx)</td>
</tr>
<tr>
<td>98</td>
<td>Other</td>
</tr>
<tr>
<td>99</td>
<td>Unknown</td>
</tr>
<tr>
<td>Code</td>
<td>Description</td>
</tr>
<tr>
<td>------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>0</td>
<td>Non-Spanish/Non-Hispanic</td>
</tr>
<tr>
<td>1</td>
<td>Mexican (includes Chicano)</td>
</tr>
<tr>
<td>2</td>
<td>Puerto Rican</td>
</tr>
<tr>
<td>3</td>
<td>Cuban</td>
</tr>
<tr>
<td>4</td>
<td>South or Central American (except Brazil)</td>
</tr>
<tr>
<td>5</td>
<td>Other specified Spanish/Hispanic origin (includes European; excludes Dominican Republic)</td>
</tr>
<tr>
<td>6</td>
<td>Spanish, NOS; Hispanic, NOS; Latino, NOS There is evidence, other than surname or maiden name, that the person is Hispanic but he/she cannot be assigned to any of the categories</td>
</tr>
<tr>
<td>7</td>
<td>Spanish surname only (effective with diagnosis on or after 1/1/1994) The only evidence of the person’s Hispanic origin is the surname or maiden name and there is no contrary evidence that the patient is not Hispanic.</td>
</tr>
<tr>
<td>8</td>
<td>Dominican Republic (effective with diagnosis on or after 1/1/2005)</td>
</tr>
<tr>
<td>9</td>
<td>Unknown whether Spanish/Hispanic or not</td>
</tr>
</tbody>
</table>
IMPROVING CLASSIFICATION OF HISPANIC PATIENTS IN SEER
NHIA DERIVED HISPANIC ORIGIN
SEER Program Coding and Staging Manual 2007
(NAACCR Item #: 191)

The NAACCR Hispanic Identification Algorithm (NHIA) is a computerized algorithm that uses a combination of variables to directly or indirectly classify cases as Hispanic for analytic purposes. The computer program that is run to derive Hispanic origin will automatically assign the code for this data item. The algorithm must be run for all cases.

0  Non-Hispanic
1  Mexican, by birthplace or other specific identifier
2  Puerto Rican, by birthplace or other specific identifier
3  Cuban, by birthplace or other specific identifier
4  South or Central American (except Brazil), by birthplace or other specific identifier
5  Other specified Spanish/Hispanic origin (includes European; excludes Dominican Republic), by birthplace or other specific identifier
6  Spanish, NOS; Hispanic, NOS; Latino, NOS
7  NHIA surname match only
8  Dominican Republic
Blank  Algorithm has not been run
The Indian Health Service Linkage (IHS) reports the results of linking of the registry database with the Indian Health Service patient registration database.

The IHS linkage identifies American Indians who were misclassified as non-Indian in the registry.

The computer linkage program will automatically assign the code for this data item. SEER requires the IHS Link for cases diagnosed January 1, 1988 and forward. IHS link may be submitted for cases diagnosed in earlier years. The field will be blank unless an attempt was made to link the case with the records from the Indian Health Service.

<table>
<thead>
<tr>
<th>Code</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>Record sent for linkage, no IHS match</td>
</tr>
<tr>
<td>1</td>
<td>Record sent for linkage, IHS match</td>
</tr>
<tr>
<td>Blank</td>
<td>Record not sent for linkage or linkage result pending</td>
</tr>
</tbody>
</table>
Contract Health Service Delivery Areas (CHSDA) and States Included in 2007 Annual Report to Nation
Top 10 AIAN Cancer Incidence Rates – Men
Age-adjusted Rates, 1999-2004, U.S. (CHSDA counties)

AIAN All cancers rate = 406.9
NHW All cancers rate = 553.5
**NAPIIA stands for NAACCR Asian and Pacific Islander Identification Algorithm. Race-NAPIIA recodes** some single-race cases with a Race 1 code of 96 to a more specific Asian race category, based on an algorithm that makes use of the birthplace and name fields (first, last, and maiden names). For single-race cases with Race 1 other than 96, it returns Race 1.

The NAACCR Asian and Pacific Islander Identification Algorithm (NAPIIA) is a computerized algorithm that uses a combination of variables to directly or indirectly classify cases as Asian or Pacific Islander for analytic purposes. The computer program will automatically assign the code for this data item.

In Version 1 of the algorithm, **birth place** can be used to indirectly assign a specific race to one of 8 Asian race groups (Chinese, Japanese, Vietnamese, Korean, Asian Indian, Filipino, Thai, and Cambodian), and **names** can be used to indirectly assign a specific race to one of 7 Asian groups (Chinese, Japanese, Vietnamese, Korean, Asian Indian, Filipino, and Hmong).

Subsequent versions of NAPIIA may incorporate Pacific Islanders and may potentially incorporate name lists for Thai, Cambodian, and Laotians.
NAACCR Semantic Interoperability Work Group
Race and Ethnicity Crosswalks

• Harmonization: compare NAACCR race and ethnicity categories and codes to those used by other health-related standard-setting organizations

• Compare categories: NCI’s cancer Biomedical Informatics Grid (caBIG) ™, the US Census Bureau, Health Level Seven (HL-7), and the National Center for Health Statistics (NCHS)

|$|$ Current NAACCR classification schemes, with slight modification best meet the needs of the NAACCR community. (The modification will provide separate codes for Asian Indians and Pakistanis while maintaining an Asian Indian/Pakistani, NOS category.)
Survival by Race/ethnicity

- The SEER survival monograph (chapter 31) shows survival by detailed race/ethnicity (white, white non-Hispanic, white Hispanic, black, American Indian/Alaska Native, Asian (Asian Indian, Chinese, Filipino, Korean, Japanese, Vietnamese, other), Pacific Islander (Hawaiian, other):
  - The new Cancer Statistics Review (about 4/15/2009) will contain survival by detailed race/ethnicity. This information will be updated annually.
  - Due to lack of U.S. lifetables by detailed race/ethnicity, cause-specific survival rates are used for comparisons.

Source: Race/ethnicity chapter of the SEER Survival monograph:

Source: Race/ethnicity chapter of the SEER Survival monograph:
SEER Ecological Data Set

- **Geo-coding** of address
  - Includes county (public) & tract (restricted) data bases for analysis
    - County level attributes include:
      * age (% <18; % 65+)
      * crowding (% households with >1 person / room)
      * educational attainment (% <9th grade; % <HS; % at least bachelors)
      * employment (% unemployment; % white collar)
      * foreign born
      * **language isolation**
      * median income
      * migration (% same house; % moved, same county; % moved, different county, same state; % moved, different state; % moved, outside the U.S.)
      * poverty (% persons below poverty; % families below poverty; % persons below 150%; % persons below 200%)
      * urban
Incidences rates by county-level poverty measure - Men

Poverty categories: <10% below poverty (most affluent), 10-19%, 20+% (least affluent).
Incidence rates by county-level poverty measure - Women

Poverty categories: <10% below poverty (most affluent), 10-19%, 20+% (least affluent).
Cancer Risk Factors

Data from 2007 Ann Rep Nation

Percent

Percentages of U.S. AIAN and U.S. NHW for various risk factors:

- Current smokers aged 18+ yrs
- Obese (self-report) aged 18+ yrs
- No leisure physical activity aged 18+ yrs
- 5+ drinks/day in past yr aged 18+ yrs

Data from 2007 Ann Rep Nation
Access to Health Care

- U.S. AIAN
- U.S. NHW

Percent

<table>
<thead>
<tr>
<th>Category</th>
<th>U.S. AIAN</th>
<th>U.S. NHW</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ages &lt;65 years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ages 65+ years</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Males</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Females</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

% Persons lacking health care coverage

% Persons lacking regular source of care (aged 18-64 years)

Data from 2007 Ann Rep Nation
Methods for Measuring Health Disparities

- NCI funded research (on-going) to examine methods for measuring health disparities
  - 2 NCI cancer surveillance monographs
  - Series of papers published that empirically compare multiple measures
  - Health Disparities Calculator (HD*Calc) software application (Beta Version)

Rationale for this work:

Broad consensus on importance of social disparities in health
- No consensus on how to measure and monitor progress toward the goal of eliminating health disparities
  - Difficult to communicate extent of disparities
  - Hinder ability of public health organizations to monitor progress toward HP2010 (and beyond) cancer objectives
Methods for Measuring Cancer Disparities: A Review Using Data Relevant to Healthy People

2010 Cancer-Related Objectives


• Despite the increased attention to social disparities in health, no clear framework exists to define and measure health disparities.

• Monograph raises some conceptual issues and reviews different methodological approaches.

• Explores 22 potential absolute and relative measures.

The monograph was published in 2005 and written under contract from the Surveillance Research Program and the Applied Research Program of the NCI.
Selected Comparisons of Measures of Health Disparities: A Review Using Databases Relevant to Healthy People 2010 Cancer-Related Objectives


• 22 separate analyses in 10 case studies
  – 9 (41%) of analyses revealed situations where trend in disparity is difficult to determine without *a priori* judgment about what dimensions of disparity are important (e.g. relative vs. absolute disparity, weighting by population size, reference group, etc.)

• Empirically compared summary measures
  (assessments of: SES, race / ethnic, and geographic disparities)

• Detailed look at the performance of these measures

The monograph was published in 2007 and written under contract from the Surveillance Research Program and the Applied Research Program of the NCI.
Trends in lung cancer incidence among males by race and trends in overall absolute and relative racial disparity, 1990-2001

Health Disparities Calculator (HD*Calc)
Beta Version

• Software application designed to process population based health data

• Created as an extension of SEER*Stat that can calculate multiple measures of cancer-related health disparities from SEER data

• Application also able to “process” other public health data

• Graphic and tabular review of 9 absolute and relative disparities measures
Research on Disparities: Effect of Immigration Status

- Cancer patterns differ between immigrants and persons in country of origin
- Immigrant populations in the US are increasing rapidly
  - In California, comprising
    - ~80% of Asians
    - ~60% of Latinos
- SEER data is being used to create a resource to study cancer incidence trends in specific Asian and Latino immigrant groups
The Emerging U.S. Cancer Surveillance Program

Explain cancer trends
Explain cancer disparities
Assess technologies and Rxs
Provide leads for new research

National Cancer Registry System

+Linked data sets
+Rapid Response Studies

Medicare
NHIS
CHIS

<table>
<thead>
<tr>
<th></th>
<th>Total U.S. Population</th>
<th>SEER 9 of U.S</th>
<th>SEER 13 of U.S</th>
<th>SEER 17 of U.S</th>
</tr>
</thead>
<tbody>
<tr>
<td>White</td>
<td>240,081,991</td>
<td>21,130,038</td>
<td>30,409,546</td>
<td>60,109,773</td>
</tr>
<tr>
<td>Black</td>
<td>39,067,463</td>
<td>3,516,120</td>
<td>4,642,456</td>
<td>8,809,439</td>
</tr>
<tr>
<td>American Indian/Alaska Native</td>
<td>3,161,493</td>
<td>399,367</td>
<td>649,084</td>
<td>1,038,247</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>14,196,114</td>
<td>2,899,989</td>
<td>4,842,382</td>
<td>7,459,487</td>
</tr>
<tr>
<td>Hispanic*</td>
<td>42,864,088</td>
<td>3,291,289</td>
<td>8,715,209</td>
<td>16,802,783</td>
</tr>
<tr>
<td>Non-Hispanic White**</td>
<td>200,154,257</td>
<td>18,130,480</td>
<td>22,235,116</td>
<td>44,363,559</td>
</tr>
</tbody>
</table>

*   Hispanic ethnicity may be of any race (e.g., White Hispanic, Black Hispanic,...)
**  Non-Hispanic White is a subset of the White population.

By the end of the 2005 diagnosis year, the database of 13 SEER and 4 expansion registries (plus Arizona Indians) contained information on over 7 million cases diagnosed since 1973.
New cases added in the most recent data year (not including Arizona Indians) numbered over 370,000.