Feasibility of Multi-Race Reporting for Cancer Incidence

Dave Stinchcomb¹, Mandi Yu², Steve Scoppa³, J. Todd Gibson³
June 18, 2015

¹ Westat.  ² NCI.  ³ IMS.
Topics

- How often are multiple races reported in cancer incidence data?  
  - How does this compare with Census data?

- What information about multiple races comes in to the registry?

- What data on multiple races are / will be available in EHR systems?

- How would multi-race data impact cancer rate and trend reporting?
Multiple races in SEER data

- SEER research dataset
  - Percent of records with more than one race
    - Compared with U.S. Census data
  - Breakdowns:
    - By registry
    - By age group
    - By year
    - By vital status
    - By reporting source
Since 2000, can enter more than one race:

- 2000: 2.4% entered two or more races
- 2010: 2.9% entered two or more races
### U.S Census multi-race reporting - SEER areas

<table>
<thead>
<tr>
<th></th>
<th>Census 2000</th>
<th>Census 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>US Total U.S.</td>
<td>2.4</td>
<td>2.9</td>
</tr>
<tr>
<td>US SEER 18 areas</td>
<td>3.5</td>
<td>3.9</td>
</tr>
<tr>
<td>AK Alaska</td>
<td>5.4</td>
<td>7.3</td>
</tr>
<tr>
<td>CA Greater California</td>
<td>4.6</td>
<td>4.9</td>
</tr>
<tr>
<td>CA Los Angeles</td>
<td>4.9</td>
<td>4.5</td>
</tr>
<tr>
<td>CA San Jose-Monterey</td>
<td>4.7</td>
<td>4.9</td>
</tr>
<tr>
<td>CA SF-Oakland</td>
<td>5.0</td>
<td>5.5</td>
</tr>
<tr>
<td>CT Connecticut</td>
<td>2.2</td>
<td>2.6</td>
</tr>
<tr>
<td>GA Atlanta</td>
<td>1.9</td>
<td>2.6</td>
</tr>
<tr>
<td>GA Greater Georgia</td>
<td>1.1</td>
<td>1.9</td>
</tr>
<tr>
<td>GA Rural Georgia</td>
<td>0.6</td>
<td>1.1</td>
</tr>
<tr>
<td>HI Hawaii</td>
<td>21.4</td>
<td>23.6</td>
</tr>
<tr>
<td>IA Iowa</td>
<td>1.1</td>
<td>1.8</td>
</tr>
<tr>
<td>KY Kentucky</td>
<td>1.1</td>
<td>1.7</td>
</tr>
<tr>
<td>LA Louisiana</td>
<td>1.1</td>
<td>1.6</td>
</tr>
<tr>
<td>MI Detroit</td>
<td>2.2</td>
<td>2.3</td>
</tr>
<tr>
<td>NJ New Jersey</td>
<td>2.5</td>
<td>2.7</td>
</tr>
<tr>
<td>NM New Mexico</td>
<td>3.6</td>
<td>3.7</td>
</tr>
<tr>
<td>UT Utah</td>
<td>2.1</td>
<td>2.7</td>
</tr>
<tr>
<td>WA Seattle</td>
<td>4.0</td>
<td>5.1</td>
</tr>
</tbody>
</table>
Multiple races in SEER incidence data

- Analysis of the NCI/SEER research dataset
- Asian subgroups collapsed and NHPI subgroups collapsed so they don’t count as different races
- 2009 to 2011:

<table>
<thead>
<tr>
<th>Race Count</th>
<th>Count</th>
<th>Percent</th>
</tr>
</thead>
<tbody>
<tr>
<td>One race</td>
<td>1,346,335</td>
<td>99.65%</td>
</tr>
<tr>
<td>2 or more</td>
<td>4,779</td>
<td>0.35%</td>
</tr>
<tr>
<td>Two races</td>
<td>4,199</td>
<td>0.31%</td>
</tr>
<tr>
<td>Three races</td>
<td>568</td>
<td>0.04%</td>
</tr>
<tr>
<td>Four races</td>
<td>12</td>
<td>0.00%</td>
</tr>
<tr>
<td>Five races</td>
<td>0</td>
<td>0.00%</td>
</tr>
<tr>
<td>Total</td>
<td>1,351,114</td>
<td>100.00%</td>
</tr>
</tbody>
</table>

- SEER Overall: **0.35%**
- Census 2010 for SEER areas: **3.9%**
Multiple races in SEER – by registry

<table>
<thead>
<tr>
<th></th>
<th>SEER Incidence 2009-2011</th>
<th>Census 2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>US SEER 18 areas</td>
<td>0.35%</td>
<td>3.9</td>
</tr>
<tr>
<td>AK Alaska Natives*</td>
<td>1.15%</td>
<td>7.3</td>
</tr>
<tr>
<td>CA Greater California</td>
<td>0.10%</td>
<td>4.9</td>
</tr>
<tr>
<td>CA Los Angeles</td>
<td>0.09%</td>
<td>4.5</td>
</tr>
<tr>
<td>CA San Jose-Monterey</td>
<td>0.08%</td>
<td>4.9</td>
</tr>
<tr>
<td>CA SF-Oakland</td>
<td>0.11%</td>
<td>5.5</td>
</tr>
<tr>
<td>CT Connecticut</td>
<td>0.04%</td>
<td>2.6</td>
</tr>
<tr>
<td>GA Atlanta</td>
<td>0.05%</td>
<td>2.6</td>
</tr>
<tr>
<td>GA Greater Georgia</td>
<td>0.02%</td>
<td>1.9</td>
</tr>
<tr>
<td>GA Rural Georgia</td>
<td>0.13%</td>
<td>1.1</td>
</tr>
<tr>
<td>HI Hawaii</td>
<td><strong>10.83%</strong></td>
<td><strong>23.6</strong></td>
</tr>
<tr>
<td>IA Iowa</td>
<td>0.02%</td>
<td>1.8</td>
</tr>
<tr>
<td>KY Kentucky</td>
<td>0.01%</td>
<td>1.7</td>
</tr>
<tr>
<td>LA Louisiana</td>
<td>0.06%</td>
<td>1.6</td>
</tr>
<tr>
<td>MI Detroit</td>
<td>0.34%</td>
<td>2.3</td>
</tr>
<tr>
<td>NJ New Jersey</td>
<td>0.47%</td>
<td>2.7</td>
</tr>
<tr>
<td>NM New Mexico</td>
<td>0.20%</td>
<td>3.7</td>
</tr>
<tr>
<td>UT Utah</td>
<td>0.04%</td>
<td>2.7</td>
</tr>
<tr>
<td>WA Seattle</td>
<td>0.79%</td>
<td>5.1</td>
</tr>
</tbody>
</table>

* Alaska SEER incidence is for Alaska Natives only, Census is for all Alaska residents
Multiple races in SEER – by age group

- There is much more multi-race reporting in younger age groups in Census data than in SEER.

![Graph showing multi-race reporting by age group and ethnicity between Census 2010 and SEER Incidence 2009-2011.](image)
Multiple races in SEER – by year

- Percent multi-race reporting in SEER 2000-2011
  - Increased after Census 2000, stable since 2004
Multiple races in SEER – by vital status

- Percent multi-race reporting higher for deceased cases
  - Increasing since 2003 (availability on death certificates?)
Multiple races in SEER – by reporting source

- Percent multi-race reporting is highest for radiation treatment, medical oncology, and outpatient treatment centers

<table>
<thead>
<tr>
<th>Reporting Source</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Radiation treatment or medical oncology center</td>
<td>1.44%</td>
</tr>
<tr>
<td>Other hospital outpatient unit or surgery center</td>
<td>1.42%</td>
</tr>
<tr>
<td>Autopsy only</td>
<td>0.98%</td>
</tr>
<tr>
<td>Laboratory only (hospital or private)</td>
<td>0.74%</td>
</tr>
<tr>
<td>Death certificate only</td>
<td>0.60%</td>
</tr>
<tr>
<td>Hospital inpatient/outpatient or clinic</td>
<td>0.31%</td>
</tr>
<tr>
<td>Nursing/convalescent home/hospice</td>
<td>0.18%</td>
</tr>
<tr>
<td>Physicians office/private medical practitioner</td>
<td>0.15%</td>
</tr>
</tbody>
</table>
Multi-race info in registry source data

- Hospital records – usually report single race

- Doctor’s offices – generally do not report race

- Multiple races specified in source documents from eight SEER registries using SEER*DMS – 2012 cases:
  - Overall 0.6% of cases have 2 or more races
  - Hawaii: 12.5% of cases have 2 or more races
  - In general, race values are only coded to multiple races if:
    - There is an abstract with multiple race codes
    - There are terms like “biracial” in the supporting text
Multi-race info in registry source data (cont.)

- Death certificate data –
  - Single race historically
  - Multiple-race reporting introduced in 2003:
    - Adopted by 40+ states with the 2013 data:

- Supplemental data on race:
  - Department of Motor Vehicle records
  - Birth records
  - Voter registration files
NAACCR Data Item Consolidation Manual

- **Race 1 [160]**
  - Known over unknown
  - More specific over less specific
  - Most frequent
  - Manual review (can consult other data sources)

- **Race 2-5 [161-164]**
  - Known over unknown
  - Unknown Race 2-5 if Race 1 is unknown
  - First record reported
    - Rational: “not a critical variable and may not warrant the time required for manual review”
  - Manual review (can consult other data sources)
EHR support for multiple races

- The Cancer Reporting CDA supports an unlimited number of races from an EHR system.
  - Appears in the `/ClinicalDocument/recordTarget/patientRole/patient document`
  - Uses the CDC's Race and Ethnicity coding vocabulary
  - An example of a valid two-race entry:
    - `<raceCode code="2058-6" codeSystem="2.16.840.1.113883.6.238" codeSystemName="Race and Ethnicity - CDC" displayName="African American"/>
    - `<sdtc:raceCode code="2106-3" codeSystem="2.16.840.1.113883.6.238" codeSystemName="Race and Ethnicity - CDC" displayName="White"/>
Implications for cancer rate reporting

- Incidence rates are commonly reported by race
Possible multi-race categories

- Bridged races
  - White
  - Black
  - Asian and Pacific Islander (API)
  - American Indian, Alaska Native (AIAN)
  - Other

- Multi-race reporting
  - White alone
  - Black alone
  - API alone
  - AIAN alone
  - Some other race alone
  - 2 or more races

- These groups will be smaller than the bridged race groups
- More cell suppression in areas with small minority populations
Implications for trend reporting

- Trends in cancer incidence used to track disparities
Discussion / conclusions

- SEER data results
  - Multi-race reporting currently much lower than Census

- Multiple race in source records
  - Future: lobby hospitals and physicians to collect multiple races with admitting paperwork

- Support in electronic medical record systems

- Alternatives to race reporting & tracking disparities?
  - SES (individual, neighborhood)
  - Foreign born
  - Hispanic ethnicity alone
  - Behavioral (smoking, screening)
  - Haplotype: haploid genotype
Special acknowledgements

- Co-authors:
  - Mandi Yu, NCI
  - Steve Scoppa, IMS
  - Todd Gibson, IMS

- Race reporting in medical records
  - Mary Mesnard, Westat
  - Linda Coyle, IMS
  - Folks from eight SEER registries (AK, CT, Detroit, GA, HI, IA, NJ, UT)

- EHR reporting capabilities:
  - Eric Durbin, Kentucky Cancer Registry
  - York Dobyns, Kentucky Cancer Registry
  - Tonya Brandenburg, Kentucky Cancer Registry
Thank You.

DavidStinchcomb@westat.com
Collapsing Asian and NHPI subgroups

- **Asian includes**
  - Chinese
  - Japanese
  - Filipino
  - Korean (1988+)
  - Vietnamese (1988+)
  - Laotian (1988+)
  - Hmong (1988+)
  - Kampuchean (1988+)
  - Thai (1994+)
  - Asian Indian or Pakistani, NOS (1988+)
  - Asian Indian (2010+)
  - Pakistani (2010+)
  - Other Asian (1991+)

- **NHPI includes**
  - Hawaiian
  - Micronesian, NOS (1991+)
  - Chamorran (1991+)
  - Guamanian, NOS (1991+)
  - Polynesian, NOS (1991+)
  - Tahitian (1991+)
  - Samoan (1991+)
  - Tongan (1991+)
  - Melanesian, NOS (1991+)
  - Fiji Islander (1991+)
  - New Guinean (1991+)
  - Pacific Islander, NOS (1991+)