NAACCR Data Quality Indicators

NAACCR 2011-2012 Webinar Series
June 14, 2012

Q&A

• Please submit all questions concerning webinar content through the Q&A panel.

Reminder:

• If you have participants watching this webinar at your site, please collect their names and emails.
  – We will be distributing a Q&A document in about one week. This document will fully answer questions asked during the webinar and will contain any corrections that we may discover after the webinar.

2
Fabulous Prizes

Agenda

- **NAACCR Data Quality Reports**
  - Glenn Copeland, Director of the Michigan Cancer Surveillance Program, CINA Chair
- **Evaluation of NAACCR Survival Data**
  - Hannah K Weir, PhD, Division of Cancer Prevention and Control Centers for Disease Prevention and Control
  - Chris J Johnson, MS Cancer Data Registry of Idaho
- **Stage data profile**
  - Brad Wohler, Florida Cancer Data System, Manager, Statistical Analysis
- **Factors associated with unknown stage prostate cancer**
  - Maria Schymura, PhD, Director New York State Cancer Registry
NAACCR Data Quality Reports

Using NAACCR DQI Reports to Assess Submitted Call for Data

Objectives

• Explain Data Quality Indicators Report
  – What does the DQI include
  – Why they are generated
  – What they can tell you

• Review New DQI Analytical Summary
  – Introduced this year
  – Explanation of statistics and presentation
General Information

• Annual Call-for-Data submissions are analyzed
  – Assess submission for data problems
• NAACCR Certification
  – Determines Certification
• CINA Editorial
  – Inclusion in CINA Combined

Confidentiality

• IMS Receives the data submissions
  – Responsible for data file assessments
  – Designs and Produces DQI reports for NAACCR
  – Provides DQI to Certification and CINA Committees only

• Reports by registry are privileged
  – Available to committee members only
  – To be used to carry out committee duties
Provided to Submitting Registry

• Shared with each submitting registry
  – Provides summary data used by NAACCR committees
  – Delineates certification and inclusion measures
  – Offers tool for registry to review their data

DQI Contents

• Series of tables by year of diagnosis
• Incidence counts by year and by site
• Certification and inclusion criteria
• Field Specific tables of submitted variables by year
Inclusion Criteria Information

<table>
<thead>
<tr>
<th>Year of Diagnosis</th>
<th>Sex</th>
<th>Race</th>
<th>County</th>
<th>Age</th>
<th>Reporting Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unk</td>
<td>Valid</td>
<td>% Unk</td>
<td>Unk</td>
<td>Reporting Source</td>
</tr>
<tr>
<td>2005</td>
<td>11</td>
<td>53,934</td>
<td>0.02</td>
<td>1.201</td>
<td>52,744</td>
</tr>
<tr>
<td>2006</td>
<td>9</td>
<td>54,388</td>
<td>0.02</td>
<td>1.313</td>
<td>53,084</td>
</tr>
<tr>
<td>2007</td>
<td>32</td>
<td>55,537</td>
<td>0.06</td>
<td>1.402</td>
<td>55,167</td>
</tr>
<tr>
<td>2008</td>
<td>47</td>
<td>53,843</td>
<td>0.09</td>
<td>1.300</td>
<td>52,590</td>
</tr>
<tr>
<td>2009</td>
<td>41</td>
<td>53,489</td>
<td>0.08</td>
<td>1.105</td>
<td>52,425</td>
</tr>
</tbody>
</table>

Inclusion Criteria Information

<table>
<thead>
<tr>
<th>Year of Diagnosis</th>
<th>Sex</th>
<th>Race</th>
<th>County</th>
<th>Age</th>
<th>Reporting Source</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Unk</td>
<td>Valid</td>
<td>% Unk</td>
<td>Unk</td>
<td>Reporting Source</td>
</tr>
<tr>
<td>2005</td>
<td>0</td>
<td>34,369</td>
<td>0.00</td>
<td>373</td>
<td>34,016</td>
</tr>
<tr>
<td>2006</td>
<td>0</td>
<td>35,126</td>
<td>0.00</td>
<td>224</td>
<td>34,904</td>
</tr>
<tr>
<td>2007</td>
<td>0</td>
<td>38,884</td>
<td>0.00</td>
<td>440</td>
<td>38,435</td>
</tr>
<tr>
<td>2008</td>
<td>0</td>
<td>38,641</td>
<td>0.00</td>
<td>529</td>
<td>38,312</td>
</tr>
<tr>
<td>2009</td>
<td>1</td>
<td>35,465</td>
<td>0.00</td>
<td>360</td>
<td>35,106</td>
</tr>
</tbody>
</table>
Screening Item Details

- Code Distributions
- Illegal/Inappropriate
- IHS Link
- Cancer Sequence
- Pre 2004 Benign
- Blank and Unknown %
- Trends in Unknowns
- Edit Override Usage

Spot Incorrect – Nonstandard Coding

<table>
<thead>
<tr>
<th>Description</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>64,822</td>
<td>62,122</td>
<td>62,302</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>30</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>43</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>91</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>115</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>145</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>223</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>224</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>226</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>641</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>999</td>
<td>27</td>
<td>43</td>
<td>31</td>
</tr>
<tr>
<td>Inv. State unable to identify cnty</td>
<td>39</td>
<td>43</td>
<td>32</td>
</tr>
<tr>
<td>Unknown</td>
<td>1,212</td>
<td>1,092</td>
<td>1,056</td>
</tr>
<tr>
<td>Valid County</td>
<td>60,340</td>
<td>60,939</td>
<td>61,189</td>
</tr>
</tbody>
</table>
### Processing Assessments – IHS Link

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Record sent for linkage, IHS match</td>
<td>52,966</td>
<td>53,833</td>
<td>54,993</td>
<td>56,407</td>
<td>56,047</td>
<td>59,306</td>
</tr>
<tr>
<td>Record sent for linkage, no IHS match</td>
<td>80</td>
<td>83</td>
<td>105</td>
<td>105</td>
<td>111</td>
<td>106</td>
</tr>
<tr>
<td></td>
<td>52,886</td>
<td>55,750</td>
<td>54,888</td>
<td>56,302</td>
<td>57,935</td>
<td>58,200</td>
</tr>
<tr>
<td></td>
<td>59,809</td>
<td>61,149</td>
<td>61,390</td>
<td>61,430</td>
<td>62,185</td>
<td>62,321</td>
</tr>
<tr>
<td>Record sent for linkage, IHS match</td>
<td>118</td>
<td>109</td>
<td>146</td>
<td>143</td>
<td>137</td>
<td>148</td>
</tr>
<tr>
<td>Record sent for linkage, no IHS match</td>
<td>59,690</td>
<td>61,010</td>
<td>61,247</td>
<td>61,287</td>
<td>62,051</td>
<td>62,193</td>
</tr>
<tr>
<td></td>
<td>64,822</td>
<td>62,122</td>
<td>62,302</td>
<td>62,500</td>
<td>61,965</td>
<td>62,120</td>
</tr>
<tr>
<td>Record sent for linkage, IHS match</td>
<td>122</td>
<td>107</td>
<td>182</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Record sent for linkage, no IHS match</td>
<td>64,500</td>
<td>61,965</td>
<td>62,120</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>399,997</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### Data Quality Priorities - Derived Stage

<table>
<thead>
<tr>
<th>Description</th>
<th>2003</th>
<th>2004</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>IS</td>
<td>61,393</td>
<td>61,430</td>
<td>62,188</td>
<td>62,341</td>
<td>64,622</td>
<td>62,122</td>
<td>62,302</td>
</tr>
<tr>
<td>L</td>
<td>0</td>
<td>7,765</td>
<td>8,070</td>
<td>7,774</td>
<td>7,885</td>
<td>8,059</td>
<td>8,503</td>
</tr>
<tr>
<td>RE</td>
<td>0</td>
<td>3,452</td>
<td>3,665</td>
<td>3,839</td>
<td>4,008</td>
<td>3,840</td>
<td>3,886</td>
</tr>
<tr>
<td>RN</td>
<td>0</td>
<td>3,444</td>
<td>3,450</td>
<td>3,702</td>
<td>3,624</td>
<td>3,842</td>
<td>3,884</td>
</tr>
<tr>
<td>RE+RN</td>
<td>0</td>
<td>2,028</td>
<td>2,150</td>
<td>2,316</td>
<td>2,282</td>
<td>2,261</td>
<td>2,300</td>
</tr>
<tr>
<td>RNOS</td>
<td>0</td>
<td>83</td>
<td>101</td>
<td>246</td>
<td>556</td>
<td>527</td>
<td>525</td>
</tr>
<tr>
<td>D</td>
<td>0</td>
<td>9,646</td>
<td>11,031</td>
<td>11,607</td>
<td>12,924</td>
<td>11,960</td>
<td>12,099</td>
</tr>
<tr>
<td>NA</td>
<td>0</td>
<td>1,443</td>
<td>1,502</td>
<td>1,410</td>
<td>1,461</td>
<td>1,428</td>
<td>1,536</td>
</tr>
<tr>
<td>U</td>
<td>0</td>
<td>13,409</td>
<td>11,132</td>
<td>9,468</td>
<td>6,580</td>
<td>6,489</td>
<td>5,499</td>
</tr>
<tr>
<td>Blank</td>
<td>61,393</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>
Issues

• Registry Specific
• Lacks Comparisons
• Missing effects of other factors
  – Population changes

Needed Something Better

• Statistical relevance
• Rates and proportions
• Easy to compare across registries
CINA Submission Summary Report

- Summary of total records used in CINA.
- “Fit For Use” Criteria
- Frequency distributions and bar charts
- Compare counts across submissions
- Box and whisker plots.

Cases Received/Cases Included in CINA

<table>
<thead>
<tr>
<th>Category</th>
<th>Cases Included in CINA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total Case Records Received (1995-2009)</td>
<td>892,907</td>
</tr>
<tr>
<td>Case Exclusions for 1995-2009 (In Order of Exclusion)</td>
<td>617,235</td>
</tr>
<tr>
<td>Invalid Year, State Code, County Code = 998</td>
<td>19,547</td>
</tr>
<tr>
<td>In-situ, Benign, Borderline Malignant (1995-2009)</td>
<td>77,975</td>
</tr>
<tr>
<td>Year of Diagnosis not 2005-2009</td>
<td>519,359</td>
</tr>
<tr>
<td>Invalid Site, Missing Age, Non-Male/Female Cases</td>
<td>354</td>
</tr>
<tr>
<td>Total Cases Excluded</td>
<td>617,235</td>
</tr>
<tr>
<td>Cases Included for CINA 2005-2009</td>
<td>275,672</td>
</tr>
<tr>
<td>Malignant Cases (Inc. In situ Bladder)</td>
<td>265,899</td>
</tr>
<tr>
<td>In Situ Breast Cases</td>
<td>9,773</td>
</tr>
</tbody>
</table>
Data Quality Inclusion Criteria

<table>
<thead>
<tr>
<th>Indicator</th>
<th>2005</th>
<th>2006</th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completeness of Case Ascertainment</td>
<td>106.5</td>
<td>104.6</td>
<td>106.3</td>
<td>100.9</td>
<td>97.4</td>
</tr>
<tr>
<td>Missing Age</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Missing Sex</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
<td>0.0</td>
</tr>
<tr>
<td>Missing Race</td>
<td>1.7</td>
<td>1.8</td>
<td>2.1</td>
<td>2.2</td>
<td>1.9</td>
</tr>
<tr>
<td>Missing County</td>
<td>0.1</td>
<td>0.2</td>
<td>0.9</td>
<td>1.2</td>
<td>0.6</td>
</tr>
<tr>
<td>Death Certificate Cases Only (DCO)</td>
<td>1.2</td>
<td>1.7</td>
<td>1.1</td>
<td>1.7</td>
<td>1.2</td>
</tr>
<tr>
<td>Passing Edits</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
<td>100.0</td>
</tr>
<tr>
<td>Duplicate case reports per 1,000 records:</td>
<td>0.80</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Meets the Inclusion Criteria for Combined Volume? Yes

Call to Call Comparison - Cases

<table>
<thead>
<tr>
<th>Year</th>
<th>CINA 2011</th>
<th>CINA 2012</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>2004</td>
<td>61,234</td>
<td>NA</td>
<td>N/A</td>
</tr>
<tr>
<td>2005</td>
<td>61,468</td>
<td>61,599</td>
<td>0.21%</td>
</tr>
<tr>
<td>2006</td>
<td>63,213</td>
<td>63,408</td>
<td>0.30%</td>
</tr>
<tr>
<td>2007</td>
<td>64,213</td>
<td>64,521</td>
<td>0.48%</td>
</tr>
<tr>
<td>2008</td>
<td>63,837</td>
<td>65,321</td>
<td>2.32%</td>
</tr>
<tr>
<td>2009</td>
<td>NA</td>
<td>64,135</td>
<td>N/A</td>
</tr>
<tr>
<td>Total</td>
<td>313,985</td>
<td>318,979</td>
<td>1.60%</td>
</tr>
</tbody>
</table>
Call to Call Comparison - Race

<table>
<thead>
<tr>
<th>Counts by Race*/Ethnicity</th>
<th>CINA 2011</th>
<th>CINA 2012</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>All Races</td>
<td>313,965</td>
<td>318,979</td>
<td>1.60%</td>
</tr>
<tr>
<td>White</td>
<td>261,943</td>
<td>265,536</td>
<td>1.37%</td>
</tr>
<tr>
<td>Black</td>
<td>41,373</td>
<td>42,290</td>
<td>2.22%</td>
</tr>
<tr>
<td>Asian or Pacific Islander</td>
<td>6,002</td>
<td>6,353</td>
<td>5.85%</td>
</tr>
<tr>
<td>Am. Indian/Alaska Native</td>
<td>183</td>
<td>194</td>
<td>6.01%</td>
</tr>
<tr>
<td>Hispanic</td>
<td>15,268</td>
<td>16,042</td>
<td>5.07%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case Counts by Gender</th>
<th>CINA 2011</th>
<th>CINA 2012</th>
<th>% Change</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>160,250</td>
<td>162,335</td>
<td>1.30%</td>
</tr>
<tr>
<td>Female</td>
<td>153,715</td>
<td>156,644</td>
<td>1.91%</td>
</tr>
<tr>
<td>Total</td>
<td>313,965</td>
<td>318,979</td>
<td>1.60%</td>
</tr>
</tbody>
</table>

Relative Rates – Box and Whisker Plots

- Intended to provide a quick comparative look
- Displays the distribution of rates for all registries
  - Identifies the Median
  - Identifies the interquartile range
  - Shows maximum values
  - Identifies registry rate within the overall distribution
- Displays rates by race/ethnicity by sex
  - All cancers, lung, colorectal, breast, prostate
<table>
<thead>
<tr>
<th>Issues or problems:</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jim Hofferkamp, CTR</td>
</tr>
<tr>
<td>NAACCR, Inc.</td>
</tr>
<tr>
<td>Phone: (217) 698-0800 ext 5</td>
</tr>
<tr>
<td>Fax: (217) 698-0188</td>
</tr>
<tr>
<td><a href="mailto:jhofferkamp@naaccr.org">jhofferkamp@naaccr.org</a></td>
</tr>
</tbody>
</table>

Please submit questions through the Q&A Panel

**QUESTIONS?**
Evaluation of NAACCR Survival Data
June 14, 2012

Chris J Johnson, MS
Cancer Data Registry of Idaho
Boise, ID

Hannah K Weir, PhD
Division of Cancer Prevention and Control
Centers for Disease Prevention and Control
Atlanta, GA

And the NAACCR Survival Analysis Workgroup (SAWG)

<table>
<thead>
<tr>
<th>Name</th>
<th>State/Province or Agency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Deb Hurley</td>
<td>SC (co-chair)</td>
</tr>
<tr>
<td>Chris Johnson</td>
<td>ID (co-chair)</td>
</tr>
<tr>
<td>Glenn Copeland</td>
<td>MI</td>
</tr>
<tr>
<td>Lamy Ellison</td>
<td>Stat Can</td>
</tr>
<tr>
<td>Monique N. Hernandez</td>
<td>FL</td>
</tr>
<tr>
<td>Bin Huang</td>
<td>KY</td>
</tr>
<tr>
<td>Angela Mariotto</td>
<td>NCI</td>
</tr>
<tr>
<td>Zoran Miladinovic</td>
<td>Stat Can</td>
</tr>
<tr>
<td>Cyllene Morris</td>
<td>CA</td>
</tr>
<tr>
<td>Xiaoling Neu</td>
<td>NJ</td>
</tr>
<tr>
<td>Arti Parikh-Patel</td>
<td>CA</td>
</tr>
<tr>
<td>Paulo S. Pinheiro, MD</td>
<td>NV</td>
</tr>
<tr>
<td>Trevor Thompson</td>
<td>CDC</td>
</tr>
<tr>
<td>Donna Turner</td>
<td>MB</td>
</tr>
<tr>
<td>Baokai Qiao</td>
<td>NY</td>
</tr>
<tr>
<td>Zhenguo Qiu</td>
<td>AB</td>
</tr>
<tr>
<td>Kevin Ward</td>
<td>GA</td>
</tr>
<tr>
<td>Hannah Weir</td>
<td>CDC</td>
</tr>
<tr>
<td>Reda Wilson</td>
<td>CDC</td>
</tr>
<tr>
<td>Brad Wohler</td>
<td>FL</td>
</tr>
<tr>
<td>Kevin Zhang</td>
<td>MACRO</td>
</tr>
</tbody>
</table>
Overview

• What is population-based survival and how is it used?
• Data evaluation
• Putting it all together
• Next steps

What is Population-Based Survival

• Measures survival achieved in the population regardless of age, race, stage of disease, access to health care, etc.
• Can be used to:
  • Target and monitor cancer control and health policy initiatives
  • Evaluate the effectiveness of healthcare delivery (measure of cancer system performance)
Innovative Uses of Survival Data

- Compare survival by geographic area, race, ethnicity, SES, etc.
- Estimate the number of avoidable deaths within a specified time period if there were no disparities
- Estimate the population “cure” fraction
- Estimate “current” survival using period analysis

EUROCARE: Survival of Cancer Patients in Europe
http://www.eurocare.it/

Types of Population-based Survival

- Observed survival:
  ... how many individuals diagnosed with cancer at diagnosis are alive after xx years?
  ... endpoint is death from any cause

- Cause-specific survival:
  ... how many individuals diagnosed with cancer have not died of cancer after xx years?
  ... endpoint is death from cancer only

Relative survival:
... compares the survival experience of individuals with cancer to individuals without cancer (of the same age, race, gender, etc.)*
... measure excess mortality among cancer patients
... endpoint is death from any cause

* Uses life tables

Both Cause Specific and Relative are a way of comparing survival of people who have cancer with those who don’t—they shows how much cancer shortens life
Advantages and Disadvantage of Relative vs. Cause Specific Survival

<table>
<thead>
<tr>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Relative</strong></td>
<td>Relies on fact of death not cause of death</td>
</tr>
<tr>
<td></td>
<td>Life tables may not be available for all populations</td>
</tr>
<tr>
<td><strong>Cause Specific</strong></td>
<td>Not limited to populations with life tables</td>
</tr>
<tr>
<td></td>
<td>Death Certificates may not be reliable (e.g., may be coded to site of mets or recurrence)</td>
</tr>
</tbody>
</table>

Overview

- What is population-based survival and how is it used?
- Data evaluation
- Putting it all together
- Next steps
Data

- CINA (1995-2008) 2010 data submission
- First year requested follow-up data
- Excluded Canadian data due to coding of vital status variable
- Registries
  - SEER: CA (LA, SF), Detroit, HI, IA, KY, LA, NJ, NM, UT, Seattle
  - NPCR: remaining states
    - 2 NPCR state cancer registries not included

Data Elements

- Patient Demographics
  - date of birth
  - sex
  - race/ethnicity
  - name
  - SS#
- Tumor Record
  - site
  - histology
  - behavior
  - stage
  - date of diagnosis
  - type of reporting source
- Incidence
- Follow-Up
  - date of last follow-up
  - vital status
  - cause of death
  - follow-up source central
- Alive
- Death
### Evaluation Criteria

#### CONCORD

#### EUROCare

#### C-SPAN (Cancer Survival and Prevalence Analytic Network in Canada)
- C-SPAN Data Quality Assessment Protocol for Survival Analysis

---

#### Evaluation Criteria

<table>
<thead>
<tr>
<th>% Sex, Age or Race Unknown</th>
<th>% Multiple Primaries</th>
</tr>
</thead>
<tbody>
<tr>
<td>% DCO/Autopsy</td>
<td>% Alive with 0 Survival Time</td>
</tr>
<tr>
<td>% Vital status Unknown</td>
<td>% Death within 1 Month of Diagnosis</td>
</tr>
<tr>
<td>% Edi Errors</td>
<td>% Dead 0 Survival Time not reported by DCO/Autopsy</td>
</tr>
<tr>
<td>% MV</td>
<td>% Missing Cause of Death</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The Foundation for Population-Based Survival

The validity of population-based survival comparisons is clearly dependent on the validity of the incidence data. Berrino, 2003

Factors that Impact Incidence

- NAACCR Certification
  - Completeness of case ascertainment
  - DCO/autopsy
  - Missing critical information (age, sex, race)
  - Edits
  - Duplicates
Data Quality Indicators

Certification Status of NAACCR US Cancer Registries for 1995 Data

Gold and Silver Level Certification Status of NAACCR U.S. Cancer Registries for 2008 Data

http://www.naaccr.org/Certification/WhoisCertified.aspx
Factors that Impact Incidence

- NAACCR Certification
  - Completeness of case ascertainment
  - DCO/autopsy
  - Missing critical information (age, sex, race)
  - Edits
  - Duplicates

- Population Coverage
  - 1995 - 19 US registries NAACCR Certified
  - 2008 - 53 US registries Certified

Factors that Impact Incidence

- NAACCR Certification

- Completeness of Case Ascertainment
  - Clinical vs. Microscopically Verified (MV)
% Type Diagnostic Confirmation
SEER (1992-2008)

% Type Diagnostic Confirmation
NPCR (1995-2008)
Case Completeness and % MV

Follow-Up

Incidence

Alive

Death
Demographic Variables

- Variable: Name (last, first), Sex, Date of birth, Social Security No (SS#)
- Critical for enhancing race/ethnicity, follow-up information through linkage
- Results from Melissa Jim – IHS linkage project

% Missing - Linkage Variables

<table>
<thead>
<tr>
<th></th>
<th>SS#</th>
<th>Birth Date</th>
<th>Sex</th>
<th>Last Name</th>
<th>First Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>SEER - range</td>
<td>0.00-3.93</td>
<td>0.00-0.09</td>
<td>0.00-0.02</td>
<td>0.00-&lt;0.00</td>
<td>0.00-&lt;0.00</td>
</tr>
<tr>
<td>- No. states w/missing</td>
<td>7/10</td>
<td>4/10</td>
<td>3/10</td>
<td>1/10</td>
<td>3/10</td>
</tr>
<tr>
<td>NPCR - range</td>
<td>0.00-2.58</td>
<td>0.00-0.07</td>
<td>0.00-0.03</td>
<td>0.00-&lt;0.02</td>
<td>0.00-&lt;0.02</td>
</tr>
<tr>
<td>- No. states w/missing</td>
<td>30/41</td>
<td>21/41</td>
<td>22/41</td>
<td>9/10</td>
<td>15/41</td>
</tr>
</tbody>
</table>

Source: M Jim, IHS linkage data, variable years of diagnosis
Follow-Up Variables:
Inter-Field and Intra-Record Edits

Data Variables and Edits

- Date of last contact
- Vital status
- Cause of death
- ICD revision number
- Follow-up source central
- Types of reporting source

- All NPCR and SEER registries reported <1% edit errors for any individual edits
Vital Status

- All NPCR and SEER registries reported <1% missing vital status information

Follow-Up Requirements

**Alive Status**
- SEER Program requires all SEER registries to follow alive patients
  - 95% patients have last contact date within 18 months of the annual date of submission
- NPCR registries are not required to follow patients

**Death Status**
- All Registries conduct death clearance with state DC
- SEER and NPCR provide support for registries to link with the National Death Index and the Social Security Death Index
**Events in Follow-UP**

- Incidence
- Alive
- Death

**Immediately Lost to Follow UP Alive**

- SEER 11 database (not CINA)
- 1992-2006
- Information obtained from SEER survival session
- Alive with “0” survival time
- Contribute no follow-up information
- Survival time could be 0-<1 months
- <1% survival time = 0 months (range 0.1- 0.3%)
The Importance of Death Ascertainment


**OBJECTIVE:** designed to measure the impact of variation in patient follow-up on survival statistics.

**METHODS:** SEER data used to construct datasets simulated scenarios of complete (SEER), incomplete, and no follow-up (NPCR) of alive patients; and complete and incomplete death ascertainment.

**CONCLUSIONS:**
- Complete death ascertainment important for producing accurate cancer survival statistics, and
- Ascertainment of deaths only should generally be sufficient for survival analysis.
Full Dates vs. Partial Dates

- Date of Birth: Age at diagnosis needed for Life Tables
- Date of diagnosis: Survival interval
- Date of last contact

SEER Program uses month and year
Example: Patient diagnosed April 2000 and dies May 2000. Survival interval could be 1 – 60 days

NAACCR / NPCR uses month, day and year

Survival Interval
Full Dates vs. Partial Dates

Woods LM, Rachet B, Ellis L, Coleman MP
Full dates (day, month, year) should be used in population-based cancer survival studies
Day of Diagnosis (2004-2008)

Day of Death among Decedents (2004-2008)
Cause of Death among Decedents
SEER 1995-2008

- Unknown/missing/invalid COD
- State DC not available or state DC available but no COD
- Non cancer death
- In situ, benign or unknown behavior neoplasm
- All Malignant Cancers

Cause of Death among Decedents
NPCR 1995-2008

- Unknown
- DC not available or DC available but no COD
- Non Cancer
- In situ, benign or unknown behavior neoplasm
- All Malignant Cancers
Overview

• What is population-based survival and how is it used?
• Data evaluation
• Putting it all together
• Next steps
Events in Follow-UP

Incidence

Alive

Death

Follow-Up Requirements

Alive Status
- SEER Program requires all SEER registries to follow alive patients
  - 95% patients have last contact date within 18 months of the annual date of submission
- NPCR registries are not required to follow patients

Death Status
- All Registries conduct death clearance with state DC
- SEER and NPCR provide support for registries to link with the National Death Index and the Social Security Death Index
Follow-Up Requirements

Alive Status
- SEER Program requires all SEER registries to follow alive patients
  - 95% patients have last contact date within 18 months of the annual date of submission
- NPCCR registries are not required to follow patients
  - impute follow-up date to be the end of study (e.g., 12/31/08)

Death Status
- All Registries conduct death clearance with state DC
- SEER and NPCCR provide support for registries to link with the National Death Index and the Social Security Death Index

60-Month Observed Survival 2003-2007 Cases Followed Through 2008 Female Breast Cancer
Lung & Bronchus Cancer

Liver & Bile Duct Cancer
Overview

- What is population-based survival and how is it used?
- Data evaluation
- Putting it all together
- Next steps
What to do with Multiple Primaries in Survival

- **Background**: Historic use of first cancers only in survival
- **Objective**:
  - Compare first cancers vs. all cancers
  - Evaluate the impact of SEER and IACR MP rules on survival
- **Methods and Materials**: SEER data, SEER MP rules and IACR MP rules

Multiple Primaries (MP) Available for Analysis (2004-2008)

- Only 1
- 1st of 2 +
- 2nd +
- NOS
% MP (all sites, both sexes) by IACR and SEER Rules: SEER 11 (1995-2008)

% Multiple Primaries (1995-2008) by Years of Operation among Statewide Population-based Cancer Registries
5 Yr. Survival Female Breast Cancer

5 Yr. Survival Urinary Bladder, Males
What to do with Multiple Primaries in Survival

• **Background:** Historic use of first cancers only in survival

• **Objective:**
  – Compare first cancers vs. all cancers
  – Evaluate the impact of SEER and IACR MP rules on survival

• **Methods and Materials:** SEER data, SEER MP rules and IACR MP rules

• **Results:**
  – First cancers only excludes a large and increasing number of cancers
  – First cancer only survival higher than survival using all primaries (SEER or IACR MP rules)
  – Using all cancers, survival with SEER MP lower than IACR MP for female breast and urinary bladder (males) cancer

• **Conclusion:**
  • NAACCR registries should include all primary cancers in comparative survival studies using IACR MP rules

---

Overview

• What is population-based survival and how is it used?

• Data evaluation

• Putting it all together

• Next steps
Next Steps

- Deceased with 0 survival time (and not a DCO/AO case)
  - E.g., Physician only reporting source, follow up source central (State or NDI). These events are included in analysis whereas DCO/AO cases are excluded
- Immortal cases
- Survival using full dates - SEER*Stat enhancement
- State specific life tables – available in 2012
- Participation in CONCORD Study

Christopher J. Johnson, MPH
Cancer Data Registry of Idaho
cjohnson@teamiha.org
208-489-1380

Hannah K. Weir, PhD
Division of Cancer Prevention and Control
Centers for Disease Control and Prevention
hbw4@cdc.gov
770 488-3006

The findings and conclusions in this presentation are those of the presenter and do not necessarily represent the official position of the Centers for Disease Control and Prevention.
BRAD WOHLER, FLORIDA CANCER DATA SYSTEM, MANAGER, STATISTICAL ANALYSIS

STAGE DATA PROFILE

MARCIA SCHYMURA, PHD, DIRECTOR NEW YORK STATE CANCER REGISTRY

FACTORS ASSOCIATED WITH UNKNOWN STAGE PROSTATE CANCER
Please submit all questions through the Q&A panel

QUESTIONS?

Coming up!

• 7/12/12
  – ICD-10-CM and Cancer Surveillance
• 8/2/12
  – Collecting Cancer Data: Hematopoietics

And the winners of the fabulous prizes are....