North American Association of Central Cancer Registries

Registry Certification – Past, Present and Future

Hannah K. Weir, Ph.D.
Centers for Disease Control and Prevention
and
Thomas C. Tucker, Ph.D.
University of Kentucky
NAACCR

- Umbrella organization includes
  - cancer registries, government agencies and professional associations throughout USA and Canada
- Consensus standards for cancer registries
- *Cancer In North American* (CINA)
- Certification
U.S. Cancer Registries

SEER\(^1\) Program
(1973+)

SEER 2000+/
NPCR 1995+

SEER

NPCR\(^2\) (1995+)

NPCR
NPCR1995+/S
EER 2000+

AK   HI
AK   HI
PR    AK   HI

NAACCR member registries

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

Canadian Cancer Registries

U.S. Cancer Registries

NAACCR member registries
NAACCR

- Umbrella organization includes
  - cancer registries, government agencies and professional associations throughout USA and Canada
- Consensus standards for cancer registries
- *Cancer In North American* (CINA)
- Certification
Registry Certification Goals

◆ To establish a process by which individual population-based cancer registries receive objective evaluation and confidential feedback regarding their performance in the areas of case completeness, completeness of information abstracted on critical variables, accuracy of data and timeliness
Goals (continued)

◆ To establish criteria for recognizing population-based cancer registries which achieve excellence in these areas

◆ To identify areas of strengths and weaknesses and to provide information for helping population-based cancer registries to plan future activities and to allocate resources
Elements of Certification

1. Completeness of case ascertainment
2. Completeness of critical information for incidence reporting
3. % death certificate-only cases
4. Duplicates in the database
5. Quality of the information (EDITS)
6. Timeliness
Cancer Registries Certified by NAACCR, United States, 1995 Incidence Data

Population-based registry

NAACCR-certified
Cancer Registries Certified by NAACCR, United States

- Population-based registry
- NAACCR-certified
Frequently Asked Questions and Comments

- Which cases are evaluated in certification?
- What is case completeness and how do you estimate it?
- Completeness should not be derived from small populations (e.g., unstable rates)
- Data used in estimating case completeness are not representative
  - Why use all US mortality rather than SEER area mortality?
  - SEER black populations do not represent rural black populations
- Why did my completeness estimates drop in recent years?
- Can I estimate my expected cases and why do expected case counts change each year?
Frequently Asked Questions and Comments

- Which cases are evaluated in certification?
- What is case completeness and how do you estimate it?
- Completeness should not be derived from small populations (e.g., unstable rates)
- Data used in estimating case completeness are not representative
  - Why use all US mortality rather than SEER area mortality?
  - SEER black populations do not represent rural black populations
- Why did my completeness estimates drop in recent years?
- Can I estimate my expected cases and why do expected case counts change each year?
Which cases are evaluated in Certification?

The NAACCR Call-for-Data includes all invasive cancers (including *in situ* bladder) and *in situ* breast.
...but only major invasive cancer sites are used to estimate case completeness

Major site groups include:
- Oral cavity and pharynx
- Esophagus
- Stomach
- Colon and rectum
- Liver
- Pancreas
- Lung and bronchus
- Melanoma (white only)
- Breast
- Cervix
- Corpus and uterus, NOS
- Ovary
- Bladder

Excludes
- Kidney and renal pelvis
- Brain and CNS
- Hodgkin disease
- Non Hodgkin lymphoma
- Multiple myeloma
- Leukemia
- Prostate
- Sites not listed
What is case completeness and how do you estimate it?

... is the extent to which all expected incident cases of invasive cancer occurring in a defined population are included in the cancer registry database

Observed / Expected
Start with an assumption.....
Rearrange the equation...

\[ I_{\text{Pop}} = \left( \frac{M_{\text{Pop}}}{M_{\text{US}}} \cdot I_{\text{SEER}} \right) \]
Plug in the data….

\[
2 \text{ 19} = \sum_{j=1}^{19} \sum_{I=1}^{I} \left( M_{\text{Pop}} \times \left( I_{\text{SEER}} / M_{\text{US}} \right) \right)
\]

- \( I_{\text{Pop}} \) = ?? age-adjusted incidence rate in population (1 yr)
- \( M_{\text{Pop}} \) = age-adjusted death rate in the population (2 or 3 yrs)
- \( I_{\text{SEER}} \) = age-adjusted incidence rate in 11 SEER Programs (5 yrs)
- \( M_{\text{US}} \) = age-adjusted death rate in United States (5 yrs)
- \( i \) = 19 major invasive sites
- \( j \) = gender (male, female)
Using the mortality experience in your population...

\[
\text{Expected } I_{\text{Pop}} = \sum_{j=1}^{19} \sum_{i=1}^{219} (M_{\text{Pop}} \times \frac{I_{\text{SEER}}}{M_{\text{US}}})
\]

- \(I_{\text{Pop}}\) = ?? age-adjusted incidence rate in population (1 yr)
- \(M_{\text{Pop}}\) = age-adjusted death rate in the population (2 or 3 yrs)
- \(I_{\text{SEER}}\) = age-adjusted incidence rate in 11 SEER Programs (5 yrs)
- \(M_{\text{US}}\) = age-adjusted death rate in United States (5 yrs)
- \(i\) = 19 major invasive sites
- \(j\) = gender (male, female)
The I/M ratios from...

\[
2\ 19
\]

Expected \( I_{\text{Pop}} \) = \( \sum \sum \left( M_{\text{Pop}} \times \left( I_{\text{SEER}} / M_{\text{US}} \right) \right)_{j=1}^{19} \)

\( I_{\text{Pop}} \) = ?? age-adjusted incidence rate in population (1 yr)
\( M_{\text{Pop}} \) = age-adjusted death rate in the population (2 or 3 yrs)
\( I_{\text{SEER}} \) = age-adjusted incidence rate in 11 SEER Programs (5 yrs)
\( M_{\text{US}} \) = age-adjusted death rate in United States (5 yrs)
i = 19 major invasive sites
j = gender (male, female)
Repeat for each major cancer site...

\[ \text{Expected } I_{\text{Pop}} = \sum_{j=1}^{19} \sum_{i=1}^{2} (M_{\text{Pop}} \times (I_{\text{SEER}} / M_{\text{US}}) ) \]

- \( I_{\text{Pop}} \) = ?? age-adjusted incidence rate in population (1 yr)
- \( M_{\text{Pop}} \) = age-adjusted death rate in the population (2 or 3 yrs)
- \( I_{\text{SEER}} \) = age-adjusted incidence rate in 11 SEER Programs (5 yrs)
- \( M_{\text{US}} \) = age-adjusted death rate in United States (5 yrs)
- \( i \) = 19 major invasive sites
- \( j \) = gender (male, female)
... And for males and females...

\[
\text{Expected } I_{\text{Pop}} = \sum_{j=1}^{2} \sum_{l=1}^{19} \left( M_{\text{Pop}} \times \left( I_{\text{SEER}} / M_{\text{US}} \right) \right)
\]

- \( I_{\text{Pop}} \) = ?? age-adjusted incidence rate in population (1 yr)
- \( M_{\text{Pop}} \) = age-adjusted death rate in the population (2 or 3 yrs)
- \( I_{\text{SEER}} \) = age-adjusted incidence rate in 11 SEER Programs (5 yrs)
- \( M_{\text{US}} \) = age-adjusted death rate in United States (5 yrs)
- \( i \) = 19 major invasive sites
- \( j \) = gender (male, female)
... then adjust the completeness estimate for background mortality
... then adjust the completeness estimate for background mortality
Assuming.....

\[ \text{Age-adjusted rate} \]

\[ \text{80\% incidence,} \]

\[ \text{20\% case fatality} \]
Adjust mortality up if US mortality is higher
Adjust mortality down if US mortality is lower
Example: lung and bronchus, white males, Kentucky

US
- M_{5\text{ yrs}} = 76.6

Kentucky
- M_{2\text{ yrs}} = 117.2
- M_{5\text{ yrs}} = 114.0

Mortality rate difference:
76.6 - 114.0 = -37.4 deaths per 100,000

20% due to case fatality: (0.2) (-37.4) = -7.5

Adj M_{KY} = 117.2 - 7.5 = \sim 110
Lung and bronchus cancer, white males, Kentucky

RR = 1.02

SEER/US: 119
KY expected: 142
KY expected and adjusted: 111
KY observed: 117

M = 110

M = 117
Estimate overall race proportional completeness

<table>
<thead>
<tr>
<th>Kentucky</th>
<th>White</th>
<th>Black</th>
<th>Race Proportional</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>% complete</td>
<td>Population</td>
<td>% complete</td>
</tr>
<tr>
<td>White</td>
<td>100.9%</td>
<td>3,709,328</td>
<td>96.9%</td>
</tr>
<tr>
<td>Black</td>
<td>100.6%</td>
<td>312,451</td>
<td></td>
</tr>
<tr>
<td>Race Proportional</td>
<td>100.6%</td>
<td>4,021,779</td>
<td></td>
</tr>
</tbody>
</table>
Frequently Asked Questions and Comments

- Which cases are evaluated in certification?
- What is case completeness and how do you estimate it?
- Completeness should not be derived from small populations (e.g. unstable rates)
- Data used in estimating case completeness are not representative
  - Why use all US mortality rather than SEER area mortality?
  - SEER black populations do not represent rural black populations
- Why did my completeness estimates drop in recent years?
- Can I estimate my expected cases and why do expected case counts change each year?
Completeness estimates derived from small populations can be unstable.

<table>
<thead>
<tr>
<th></th>
<th>White</th>
<th></th>
<th>Black</th>
<th></th>
<th>Race Proportional</th>
</tr>
</thead>
<tbody>
<tr>
<td>% complete</td>
<td>Population</td>
<td>% complete</td>
<td>Population</td>
<td>% complete</td>
<td>Population</td>
</tr>
<tr>
<td>93%</td>
<td>800,000</td>
<td>40%</td>
<td>5,000</td>
<td>92%</td>
<td>805,000</td>
</tr>
</tbody>
</table>
Completeness estimates derived from small populations can be unstable

<table>
<thead>
<tr>
<th>White</th>
<th>Black</th>
<th>Race Proportional</th>
</tr>
</thead>
<tbody>
<tr>
<td>% complete</td>
<td>Population</td>
<td>% complete</td>
</tr>
<tr>
<td>94%</td>
<td>700,000</td>
<td>165%</td>
</tr>
</tbody>
</table>
Completeness estimates derived from small populations can be unstable.
Completeness estimates based on 2 and 3 yrs. of mortality data

<table>
<thead>
<tr>
<th>Population Size</th>
<th>2 yr. 95-96</th>
<th>3 yr. 94-96</th>
</tr>
</thead>
<tbody>
<tr>
<td>Small</td>
<td>81.7</td>
<td>93.8</td>
</tr>
<tr>
<td>Small</td>
<td>88.5</td>
<td>90.6</td>
</tr>
<tr>
<td>Small</td>
<td>80.7</td>
<td>83.8</td>
</tr>
<tr>
<td>Medium</td>
<td>89.6</td>
<td>87.5</td>
</tr>
<tr>
<td>Medium</td>
<td>108.7</td>
<td>108.8</td>
</tr>
<tr>
<td>Large</td>
<td>96.6</td>
<td>96.5</td>
</tr>
</tbody>
</table>
Frequently Asked Questions and Comments

- Which cases are evaluated in certification?
- What is case completeness and how do you estimate it?
- Completeness should not be derived from small populations (e.g., unstable rates)
- Data used in estimating case completeness are not representative
  - Why use all US mortality rather than SEER area mortality?
  - SEER black populations do not represent rural black populations
- Why did my completeness estimates drop in recent years?
- Can I estimate my expected cases and why do expected case counts change each year?
Why do you use US mortality rather than mortality from SEER areas?

![Bar chart showing age-adjusted rates with RR = 2 and RR = 1.6]

SEER Area    SEER / US
SEER black populations do not represent rural black populations

Major site groups include:
- Oral cavity and pharynx
- Esophagus
- Stomach
- Colon and rectum
- Liver
- Pancreas
- Lung and bronchus
- Melanoma (white only)
- Breast
- Cervix
- Corpus and uterus, NOS
- Ovary
- Bladder

- Kidney and renal pelvis
- Brain and CNS
- Hodgkin disease
- Non Hodgkin lymphoma
- Multiple myeloma
- Leukemia

Excludes
- Prostate
- Sites not listed
Frequently Asked Questions and Comments

◆ Which cases are evaluated in certification?
◆ What is case completeness and how do you estimate it?
◆ Completeness should not be derived from small populations (e.g. unstable rates)
◆ Data used in estimating case completeness are not representative
  - Why use all US mortality rather than SEER area mortality?
  - SEER black populations do not represent rural black populations
◆ Why did my completeness estimates drop in recent years?
◆ Can I estimate my expected cases and why do expected case counts change each year?
Why did my completeness estimates drop in recent years?

<table>
<thead>
<tr>
<th>Completeness estimates for....</th>
<th>1999</th>
<th>2000</th>
</tr>
</thead>
<tbody>
<tr>
<td>State A</td>
<td>98.5</td>
<td>96.6</td>
</tr>
<tr>
<td>State B</td>
<td>91.1</td>
<td>88.7</td>
</tr>
<tr>
<td>State C</td>
<td>95.8</td>
<td>92.7</td>
</tr>
<tr>
<td>State D</td>
<td>98.3</td>
<td>92.4</td>
</tr>
<tr>
<td>State E</td>
<td>91.1</td>
<td>88.7</td>
</tr>
</tbody>
</table>
... because of the newly revised population estimates following the 2000 census

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Breast</td>
<td>4.93</td>
<td>5.18</td>
</tr>
<tr>
<td>Lung</td>
<td>1.23</td>
<td>1.25</td>
</tr>
<tr>
<td>Colon and rectum</td>
<td>2.55</td>
<td>2.64</td>
</tr>
</tbody>
</table>
Consider what happens to your completeness estimates over time

<table>
<thead>
<tr>
<th>Completeness estimates for dx year 1999 data as reported in .....</th>
<th>2001</th>
<th>2003</th>
</tr>
</thead>
<tbody>
<tr>
<td>State A</td>
<td>80.2%</td>
<td>91.4%</td>
</tr>
<tr>
<td>State B</td>
<td>85.3%</td>
<td>87.9%</td>
</tr>
<tr>
<td>State C</td>
<td>79.8%</td>
<td>94.5%</td>
</tr>
</tbody>
</table>
Frequently Asked Questions and Comments

- Which cases are evaluated in certification?
- What is case completeness and how do you estimate it?
- Completeness should not be derived from small populations (e.g., unstable rates).
- Data used in estimating case completeness are not representative.
  - Why use all US mortality rather than SEER area mortality?
  - SEER black populations do not represent rural black populations.
- Why did my completeness estimates drop in recent years?
- Can I estimate my expected cases and why do expected case counts change each year?
Can I estimate the completeness of my data before I submit it?

Worksheet for Completeness of Case Ascertainment

Source: NAACCR website (http://www.naaccr.org), search for “case completeness worksheet”
Expected Cases

Observed cases = 21,000

Completeness estimate = 92.95%

Expected case = 21,000 / .9295 = 22,593
Limitations

- NAACCR method for estimating the completeness of case ascertainment is intended to give an estimate of overall completeness
- Certification evaluates selected data elements necessary for reporting incidence data
Future Work

- NAACCR High Quality Data workgroup continues to refine the case completeness method