Multiple Primary - Standardized Incidence/Mortality Ratios (MP-SIR/SMR)

Sara Schonfeld and Rochelle Curtis

Radiation Epidemiology Branch, Division of Cancer Epidemiology and Genetics

NCI Analytic Tools SEERies

May 5 2022
Objectives

Introduce methods for conducting analyses within SEER*Stat MP-SIR/SMR

- Multiple primary malignancies (SIR)
- Standardized mortality ratios (SMR)

By the end of the webinar:

- Define SIR and SMR
- Generate SIRs and SMRs in SEER*Stat
- Interpret SIRs and SMRs generated in SEER*Stat
- Awareness of challenges and common pitfalls
Outline

1. Background: Cancer Survivors
2. Statistical Measures
3. SEER*Stat MP-SIR/SMR
4. Study Examples
5. Key Limitations and Strengths
Growing Population of Cancer Survivors

- Improved survival
- Aging population
Rising Number of Subsequent Malignancies

Subsequent malignancies → morbidity, mortality

Figure: Morton et al. Multiple Primary Cancers (chapter 60) in Cancer Epidemiology and Prevention 2018

SEER 9: (1975-2013)
Variation in subsequent cancer risks

- Treatment for first primary
- Age at first primary
- First primary type
- Time since first primary cancer
- Genetic susceptibility
- Shared risk factors
- Subsequent malignancy type
SEER*Stat MP-SIR/SMR: multiple primaries and mortality among cancer survivors

- Describe patterns of multiple primary cancers and mortality among survivors
  - Identify high-risk subsets
  - Quantify burden

- Etiologic clues - Hypothesis generating
  - Treatment
  - Age
  - Race
  - Latency
  - Reciprocal risks

Figure: Morton & Chanock. Nat Med 2011
Sample questions we seek to answer with MP-SIR/SMR

MP-SIR

Are cancer survivors at increased risk for developing new primary* malignancies compared with the general population?
- How do risks vary by patient-, tumor- and treatment- characteristics?

MP-SMR

Are cancer survivors at increased risk of death (overall excluding deaths due to initial cancer, cause-specific)?

*not recurrence or metastasis
New Malignancies Among Cancer Survivors: SEER Cancer Registries, 1973-2000 (Published 2006)

9 SEER cancer registries

• Population-based (10% of U.S.)
• High quality
• 2 million cancer survivors
• Nearly 30-year period (1973-2000)
• 185,000 subsequent cancers

Systematic, comprehensive evaluation subsequent cancer risk

• 50 adult and 18 childhood cancer types, including data on less common cancers
• Over 350 data tables risk by:
  • Time since initial diagnosis
  • Sex
  • Age
  • Treatment (RT vs no RT)
  • Cell type
• Key resource for methodologic considerations

Statistical Measures:

SIR, SMR, EAR
Standardized incidence/mortality ratios (SIR/SMR): Concept

- Indirect standardization for age, calendar-year, race, sex, and other factors
  - How many cases would I expect in my population if my population had the same rate of disease as some standard population?

- SEER*Stat
  - SIR: Compare to general population (SEER data)
  - SMR: Compare to US mortality data
SIR/SMR: Source of observed and expected cases

- Study population* followed for defined period of follow-up
- Person-years of follow-up can be stratified by age, sex, calendar year, race and other factors
  - Observed cases of subsequent cancers or deaths within strata
  - Expected cases of subsequent cancers or deaths
    - Stratum-specific SEER cancer incidence rates or US mortality rates multiplied by stratum-specific person-years
- Stratum-specific observed and expected values summed to generate totals

*Selected from SEER when using MP-SIR/SMR
Observed and Expected: Relative and absolute measures

Standardized incidence/mortality ratio
- Ratio: Observed/Expected
- Measure of relative risk
- Measure the strength of association

Excess absolute risk (EAR)
- Difference: \((\text{Observed-Expected})/\text{PYR}\)
- Often expressed per 10,000 PYR
- Measure of absolute risk, burden

- Risk in study population > referent population
- Risk in study population = referent population
- Risk in study population < referent population
Interpretation

- Consider magnitude and not only statistical significance
- Can have high SIR and low EAR when baseline risk is low (rare event)
- Moderate, low SIR but high EAR possible when baseline is high
SEER*Stat MP-SIR/SMR
SEER Solid Tumor and Hematologic Malignancy Rules

Purpose: Instructions for registrars

• Rules for differentiating between new primaries vs metastasis or recurrence
• Number of primaries to abstract and histology
• General and site-specific rules for selected sites

Impact for MP-SIR users

• Important to understand coding rules for outcome sites of interest
• Understand changes over time
  • Cases coded according to rules at time of diagnosis
  • May need to review multiple versions depending on time frame of your study

Resources

• https://seer.cancer.gov/tools/solidtumor/
• https://seer.cancer.gov/tools/heme/
Steps for Running MP-SIR

Our objective: walk through the considerations for setting up analysis

Steve Scoppa from IMS will give live tutorial next
1. Database selection

Suggested citation for the selected database:
### Overview of databases available for MP-SIR/SMR

| Characteristics | Incidence – SEER 8
| Previously SEER 9 | Incidence - SEER 12 (exc. AK)
| Previously SEER 13 (exc. AK) | Incidence - SEER 17 (exc. AK)
| Previously SEER 18 (exc. AK) |
|------------------|------------------|
| Registries       | San Francisco-Oakland SMSA Connecticut Hawaii Iowa New Mexico Seattle (Puget Sound) Utah Atlanta (Metropolitan) | SEER 8 plus San Jose-Monterey Los Angeles Rural Georgia Alaska Native Tumor Registry (AK)* | SEER 12 plus California excluding SF/SJM/LA Kentucky Louisiana New Jersey Greater Georgia |
| Proportion of US population covered | 8.3% | 12.2% | 26.5% |
| Total number of tumors | 4,765,822 | 5,021,215 | 8,712,246 |

*Excluded from MP-SIR/SMR

https://seer.cancer.gov/registries/data.html
1. Database selection

Limit to one database

- Different registries – different populations
- Different length follow-up
- Databases not intended to be combined – can distort results

Remember differences across the databases

SEER 8 : 1975- 2019

SEER 12 : 1992- 2019

SEER 17 : 2000- 2019
1. Database selection

**Trade-offs**

- Longer follow-up data vs larger set of registries
- Variation in variable availability
  - Example: Expanded versus more limited race variables
- Changes in classification over time
  - Example: WHO Classification of Lymphoid Neoplasms 2000
- Database selection depends on key research questions
Additional notes

- SIRs – select “incidence files for SIRs”
- SMRs – select “incidence files for SMRs”
- Research vs Research Plus
  - Research Plus includes additional variables such as:
    - Geography
    - Months in dates
    - Other demographic fields
    - Treatment information

See https://seer.cancer.gov/data/product-comparison.html
Registries ascertain radiation and chemotherapy given for first course of therapy

Available in Research Plus Data
- Based on medical record review
- Initial radiation – type*(received, no/unknown)
- Initial chemotherapy – yes, no/unknown

Interpretation
- Yes = Likely received
- No/unknown = no evidence found in medical records reviewed
- Cannot distinguish “no” from “unknown”

Implications
- Under-ascertainment: initial therapy received
- No ascertainment: treatments received after first course
- Lack information on chemotherapy agent and dose, radiation fields/doses

Observational data
- Cannot account for patient differences associated with receipt of treatment
- Caution comparing outcomes across treatment groups

* Broad categories (none/unknown, beam, radioisotopes, radioactive implants...)

SEER*Rx

 Understand what is included in “chemotherapy” recode (since 2005)

 Understand changes (since 2005)

 Access at: https://seer.cancer.gov/tools/seerrx/
### Standard rate files

<table>
<thead>
<tr>
<th>Incidence (SIR)</th>
<th>Mortality (SMR)</th>
</tr>
</thead>
</table>
| Site recode ICD-O-3/WHO 2008 (for SIRs)  
https://seer.cancer.gov/siterecode/  

*Rate file should come same database selected in Step 1*  

Mortality: Cause of death recode  

#### Incidence

- **Race (WU/B/O):** Unknown race included with white

#### Mortality

- **Race (W/B/O):** Use for analyses where you plan to stratify by race
Take time to know the rate file

2. Rate file Selection
Outcome classification: Defined in rate file

Rate file: stratify by population characteristics by which you plan to calculate SIRs/SMRs
- Example: Race, sex, calendar year, age
Create customized rates within SEER rate session, export to MP-SIR

Events customized by combinations of site, histology, behavior, etc…
- Example: Lung cancer incidence by histologic type
- Example: Thyroid cancer incidence by stage at diagnosis
- Example: Soft tissue sarcoma defined by histology and not only topography

Events based on other SEER recodes such as:
- AYA Site Recode
- International Classification of Childhood Cancer
- Lymphoma Subtype Recode

Alt. grouping population characteristics
Create customized rates within SEER rate session, export to MP-SIR

Events customized by combinations of site, histology, behavior, etc…
- Example: Lung cancer incidence by histologic type
- Example: Thyroid cancer incidence by stage at diagnosis
- Example: Soft tissue sarcoma defined by histology and not only topography

Events based on other SEER recodes such as:
- AYA Site Recode
- International Classification of Childhood Cancer
- Lymphoma Subtype Recode

Alt. grouping population characteristics

Reminder:
Use same database for rate file and study population

Observed and expected cases:
- Same source population*
- Same outcome classification
- Same stratification by population characteristics

* Applies only to SIR
Define study population (first primary cancer group)

- Select only malignant behavior (uncheck to include patients with in situ)
- Exclude death certificate only or autopsy cases
- Multiple Primary Selection: First primary only (sequence 0 or 1)
  - Selecting a population of individuals at the time of their first known malignancy when resident of SEER registry catchment area
    - Reason: Cannot account for treatments prior to the baseline cancer

Examples:
- Age at first primary cancer
- Specify type of first primary
  - Site, histology
- Sex
4. Parameters

- Latency exclusion (default 2 months)
- Reasons for different latency exclusions
  - Minimize surveillance bias
  - Minimum latency for treatment exposure (e.g., radiotherapy and subsequent neoplasms)
- Risk estimates by latency, attained age, attained calendar year
5. Events

Multiple vs Single Outcome Analysis (SIR only)

Multiple Outcome (default)

• Follow until earliest of end of study, death, date of last vital status
• Events include all subsequent (2nd, 3rd, …) primary cancers
• More complete estimate of risk among long-term survivors
  • Comparable to SEER incidence rates

Single Outcome

• Follow-up censored at second primary cancer diagnosis
• **Count 2nd primary cancers only**
  ➢ Preferable for evaluating risk of treatment related second cancers
Example: Subsequent cancer incidence after first primary thyroid cancer

1st: thyroid  2nd: colon  3rd: breast

Single outcome: Count second colon only

Multiple outcome: Count all subsequent cancers including colon, breast
6. Statistics

- SIR Tables
- Frequencies
- Case Listing

More to come during live tutorial...
Study examples
Risk of second primary papillary thyroid cancer among adult cancer survivors in the United States, 2000-2015

Rationale
• Known: Elevated risk of subsequent thyroid cancer after spectrum of first primary malignancies
• Factors contributing to risk among adult cancer survivors unclear

Objective
• Quantify risk of second papillary thyroid cancer (PTC) among adult cancer survivors by:
  • Stage of second PTC
  • Time since first primary (latency)
  • Treatment for first primary cancer

Population
• 3,175,216 ≥ 1-year adult survivors of non-thyroid malignancies
• First primary diagnosis: 2000-2014, followed through 2015 (SEER 18)
• Ages 20-84 at first primary diagnosis

Statistical analyses
• SIRs (second cancer)
• Poisson regression – tests of trends, heterogeneity (not in SEER*stat)

Schonfeld et al. Cancer Epidemiology 2020
SIRs by latency and stage of PTC

Schonfeld et al. Cancer Epidemiology 2020
SIRs by latency and stage of PTC

- SIRs elevated following broad range first primary cancers
- SIRs tend to decrease with time since first primary diagnosis
- SIRs generally similar for localized and regional/distant PTC
- No clear patterns with treatment (not shown)
Cause-specific mortality after classical HL (cHL)

Rationale

• Changes in treatment over time
  • ABVD replaced MOPP
  • Decreasing use radiotherapy
  • Comprehensive assessment of mortality after cHL lacking

Objective

• Characterize stage- and cause-specific mortality risks after diagnosis of classical Hodgkin lymphoma (cHL) in current treatment era

Population

• 24,985 adults diagnosed with cHL
• Diagnosis: 2000-2015 (SEER 18)
• Follow-up through 2016
• Restricted to patients treated with chemotherapy

Statistical analyses

• SMRs and EARs for >40 different causes of death

Dores et al. JCO 2020
## Leading causes of excess deaths after cHL

<table>
<thead>
<tr>
<th>Cause of death</th>
<th>Stages I/II</th>
<th></th>
<th>Stages III/IV</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SMR</td>
<td>EAR</td>
<td>SMR</td>
<td>EAR</td>
</tr>
<tr>
<td>All</td>
<td>3.3 (3.2-3.5)</td>
<td>103.1</td>
<td>6.0 (5.8-6.3)</td>
<td>310.3</td>
</tr>
<tr>
<td>All, excluding lymphoma</td>
<td>1.5 (1.4-1.6)</td>
<td>21.7</td>
<td>2.2 (2.0-2.4)</td>
<td>71.7</td>
</tr>
<tr>
<td>Heart disease</td>
<td>1.7 (1.5-2.0)</td>
<td>6.6</td>
<td>2.1 (1.8-2.5)</td>
<td>15.1</td>
</tr>
<tr>
<td>Interstitial lung disease</td>
<td>13.1 (9.2-17.9)</td>
<td>3.7</td>
<td>22.1 (16.6-28.8)</td>
<td>9.7</td>
</tr>
<tr>
<td>Infections</td>
<td>2.2 (1.6-2.8)</td>
<td>3.1</td>
<td>3.9 (3.1-4.9)</td>
<td>10.6</td>
</tr>
<tr>
<td>AEs due to medication/drug exposure</td>
<td>2.3 (1.7-3.2)</td>
<td>2.5</td>
<td>5.0 (3.7-6.7)</td>
<td>7.4</td>
</tr>
</tbody>
</table>

- Risk estimates increased with stage
- SMRs decrease with age
- EARs increase with age

Dores et al. JCO 2020
Summary of limitations and strengths
Be Mindful...

- Heightened medical surveillance
- Under-ascertainment of subsequent cancers in certain populations
- Under-ascertainment due to migration from registry catchment area
- Potential confounding due to age, stage, treatment and other differences when comparing SIRs
- Small cell size
- Multiple comparisons
- Misclassification of deaths on death certificates (SMR)

Curtis et al. 2006
SEER MP-SIR has great strengths!

- Large, high-quality, population-based registry data
- Key resource for studying patterns of multiple primaries and mortality among cancer survivors
- Patterns by tumor and patient characteristics, treatment
- Possible to export data for additional statistical analyses
Resources


Chapter 60: Multiple Primary Cancers

Migration issues for multiple primary analyses in SEER

- **Migration from SEER catchment area**
  - Under-ascertain subsequent malignancies among long-term survivors

- **Migration into SEER catchment area**
  - Previous malignancies ascertained outside of SEER catchment area (or prior to start of registry) not captured

- **Relocation between SEER registry catchment areas:**
  - Separate IDs in SEER (one per registry)
  - Sequence numbers correct if both registries have the person’s full cancer history.
Live Demonstration

Steve Scoppa

NCI Analytic Tools SEERies

May 5, 2022