

# 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:

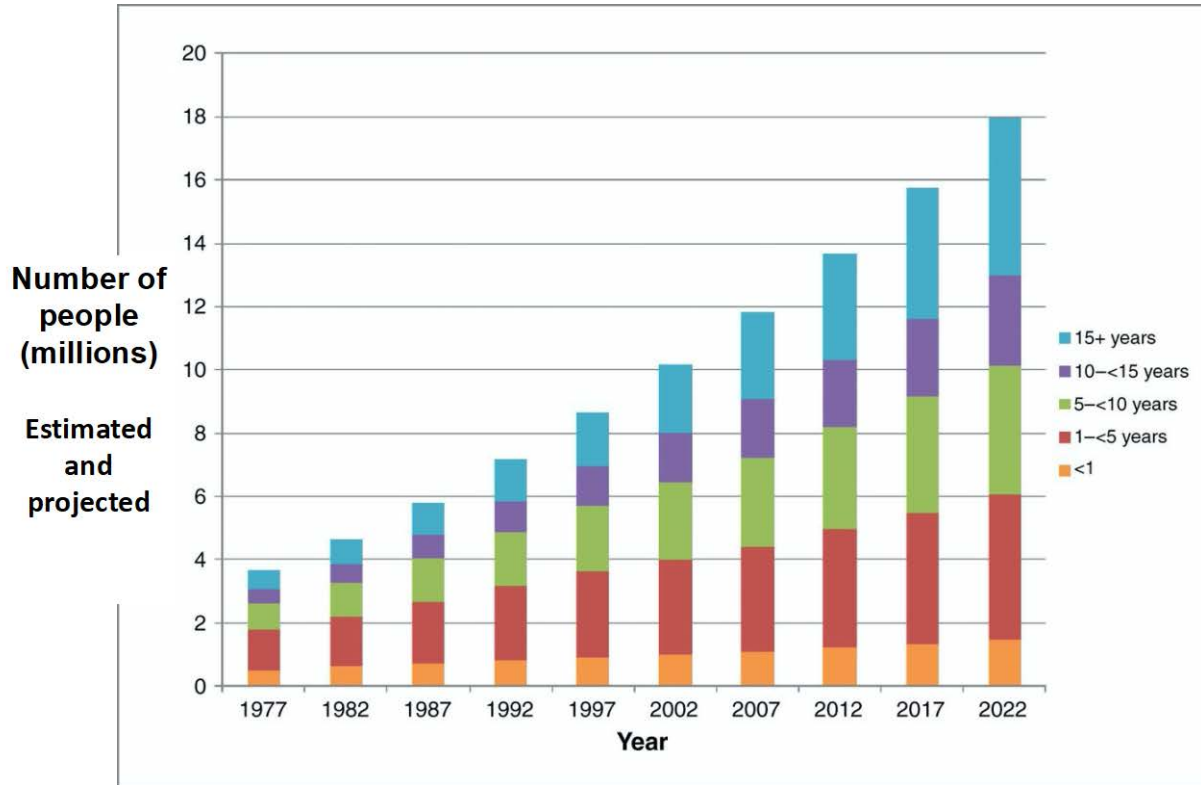
- Define SIR and SMR
- Generate SIRs and SMRs in SEER\*Stat
- Interpret SIRs and SMRs generated in SEER\*Stat
- Awareness of challenges and common pitfalls

## Outline

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

# Background

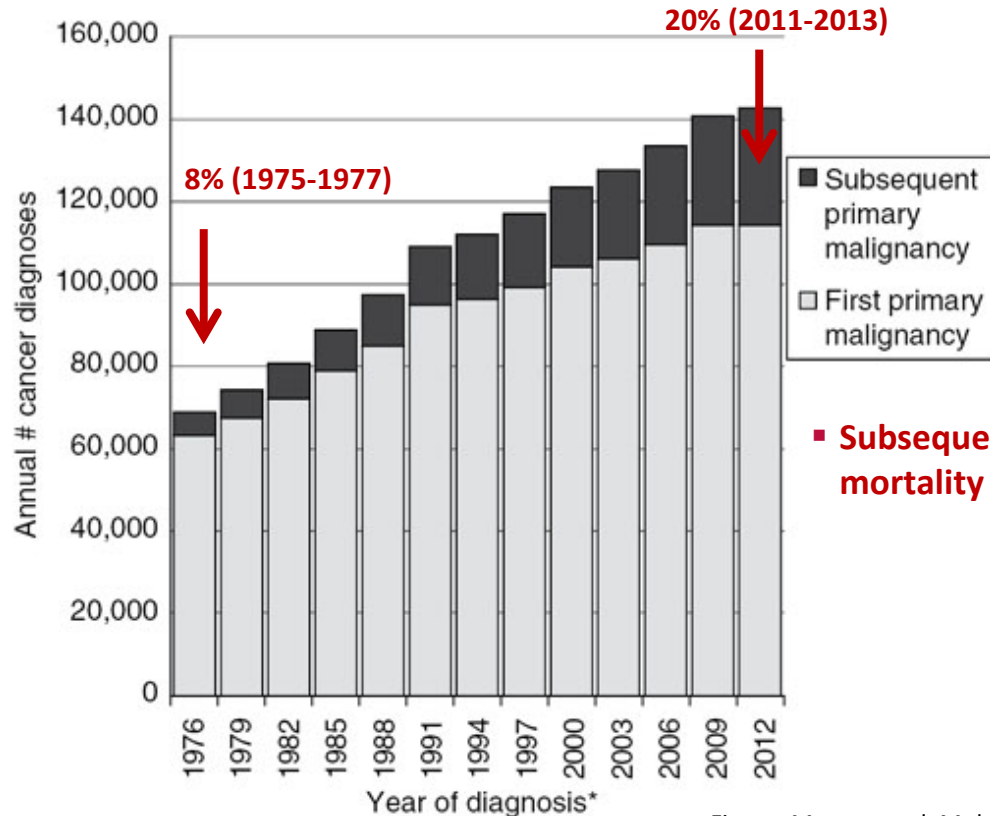
# Growing Population of Cancer Survivors



- Improved survival
- Aging population

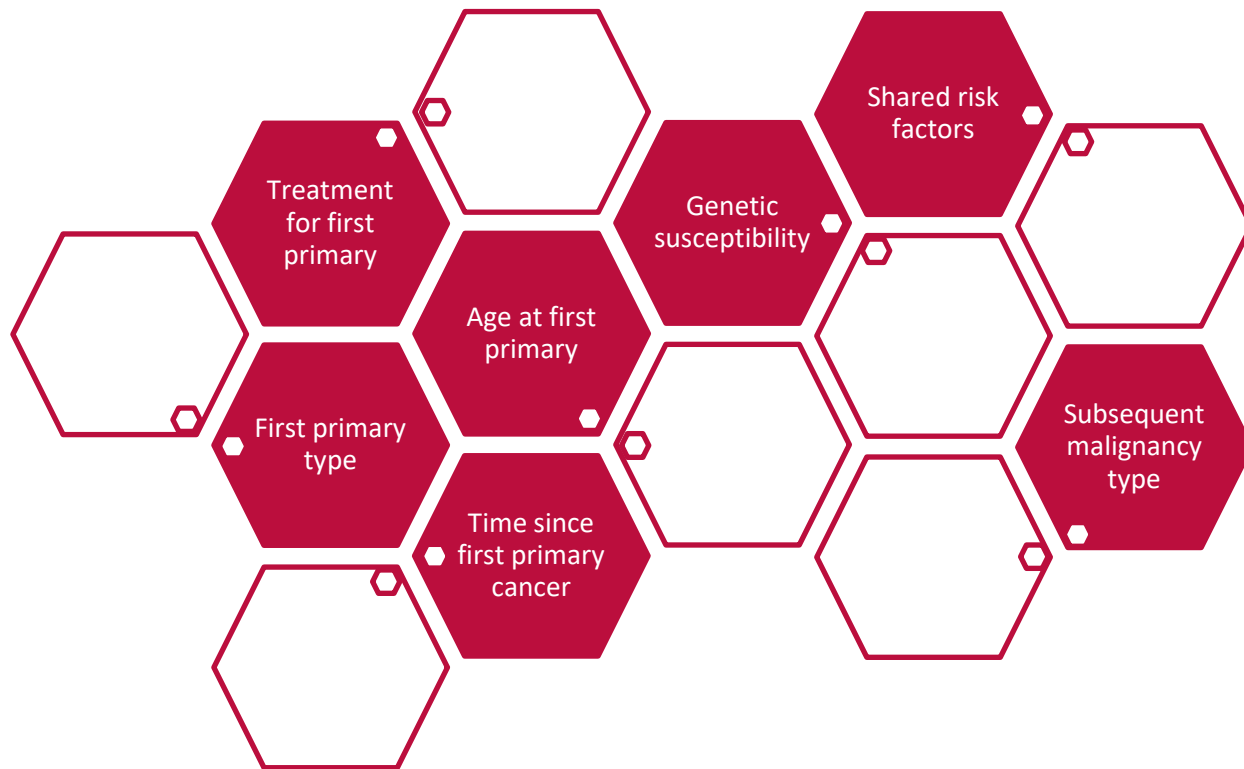
# Rising Number of Subsequent Malignancies

SEER 9: (1975-2013)

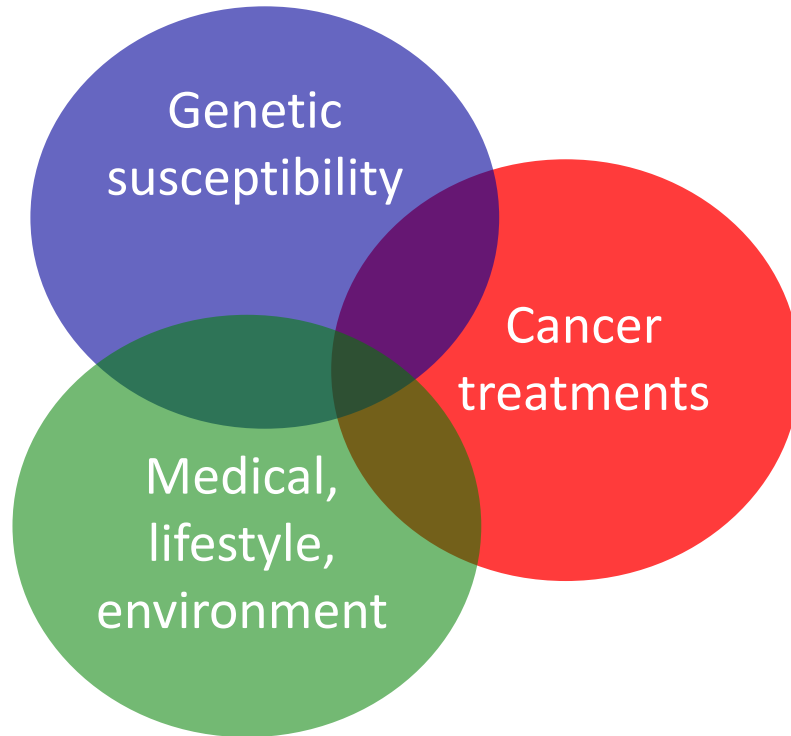


▪ Subsequent malignancies → morbidity, mortality

# Variation in subsequent cancer risks



# 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



# 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?

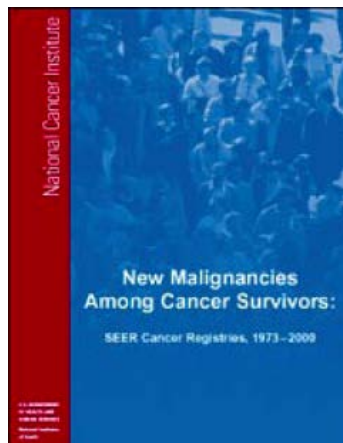
MP-SMR

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

# 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

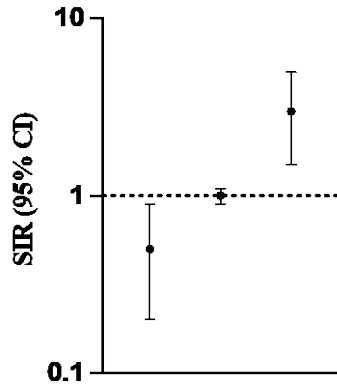
# Observed and Expected: Relative and absolute measures

## Standardized incidence/mortality ratio

- Ratio: Observed/Expected
- Measure of relative risk
- Measure the strength of association

## Excess absolute risk (EAR)

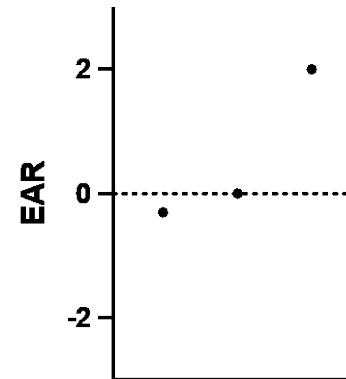
- Difference: (Observed-Expected)/PYR
- Often expressed per 10,000 PYR
- Measure of absolute risk, burden



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



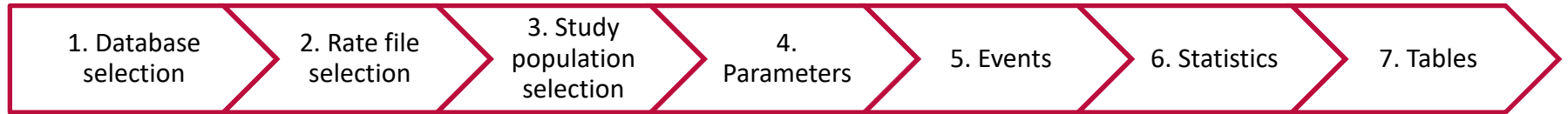
## Resources

- <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

SEER\*Stat 8.4.0  
File Edit Session Window Profile Help

Server Data: ssp://seerstat.imsweb.com:2038

MP-SIR Session-1

Data Rates Selection Parameters Events Statistic Table Output

Database Name	Linked To	Links
Incidence - SEER Research Data, 8 Registries, Nov 2021 Sub (1975-2019)		
Incidence - SEER Research Data, 12 Registries (excl AK), Nov 2021 Sub (1992-2019)		
Incidence - SEER Research Data, 17 Registries (excl AK), Nov 2021 Sub (2000-2019)		
Incidence - SEER Research Data, 8 Registries, Nov 2021 Sub (1975-2019) for SMRs		
Incidence - SEER Research Data, 12 Registries (excl AK), Nov 2021 Sub (1992-2019) for SMRs		
Incidence - SEER Research Data, 17 Registries (excl AK), Nov 2021 Sub (2000-2019) for SMRs		
Incidence - SEER Research Plus Data, 8 Registries, Nov 2021 Sub (1975-2019)	County A...	State
Incidence - SEER Research Plus Data, 12 Registries (excl AK), Nov 2021 Sub (1992-2019)	County A...	State
Incidence - SEER Research Plus Data, 17 Registries (excl AK), Nov 2021 Sub (2000-2019)	County A...	State
Incidence - SEER Research Plus Data, 8 Registries, Nov 2021 Sub (1975-2019) for SMRs	County A...	State
Incidence - SEER Research Plus Data, 12 Registries (excl AK), Nov 2021 Sub (1992-2019) for SMRs	County A...	State
Incidence - SEER Research Plus Data, 17 Registries (excl AK), Nov 2021 Sub (2000-2019) for SMRs	County A...	State
Incidence - SEER Research Data, 9 Registries, Nov 2020 Sub (1975-2018)	County A...	State
Incidence - SEER Research Data, 13 Registries (excl AK), Nov 2020 Sub (1992-2018)	County A...	State
Incidence - SEER Research Data, 18 Registries (excl AK), Nov 2020 Sub (2000-2018)	County A...	State
Incidence - SEER Research Data, 9 Registries, Nov 2020 Sub (1975-2018) for SMRs	County A...	State
Incidence - SEER Research Data, 13 Registries (excl AK), Nov 2020 Sub (1992-2018) for SMRs	County A...	State

Find Original Sort Order \* Change Linked Database... View More Database Details

Suggested citation for the selected database:  
Surveillance, Epidemiology, and End Results (SEER) Program (www.seer.cancer.gov) SEER\*Stat Database: Incidence - SEER Research Data, 8 Registries, Nov 2021 Sub (1975-2019), National Cancer Institute, DCCPS, Surveillance Research Program, released April 2022, based on the November 2021 submission.

1. Database selection

# 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  *Change in latest submission – Detroit no longer included	San Francisco-Oakland SMSA Connecticut Hawaii Iowa New Mexico Seattle (Puget Sound) Utah Atlanta (Metropolitan)	<b>SEER 8</b> plus San Jose-Monterey Los Angeles Rural Georgia Alaska Native Tumor Registry (AK)*	<b>SEER 12</b> plus California excluding SF/SJM/LA Kentucky Louisiana New Jersey Greater Georgia
Proportion of US population covered	8.3%	12.2%	26.5%
Calendar years of diagnosis	1975-2019	1992-2019	2000-2019
Total number of tumors	4,765,822	5,021,215	8,712,246

\*Excluded from MP-SIR/SMR

## 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

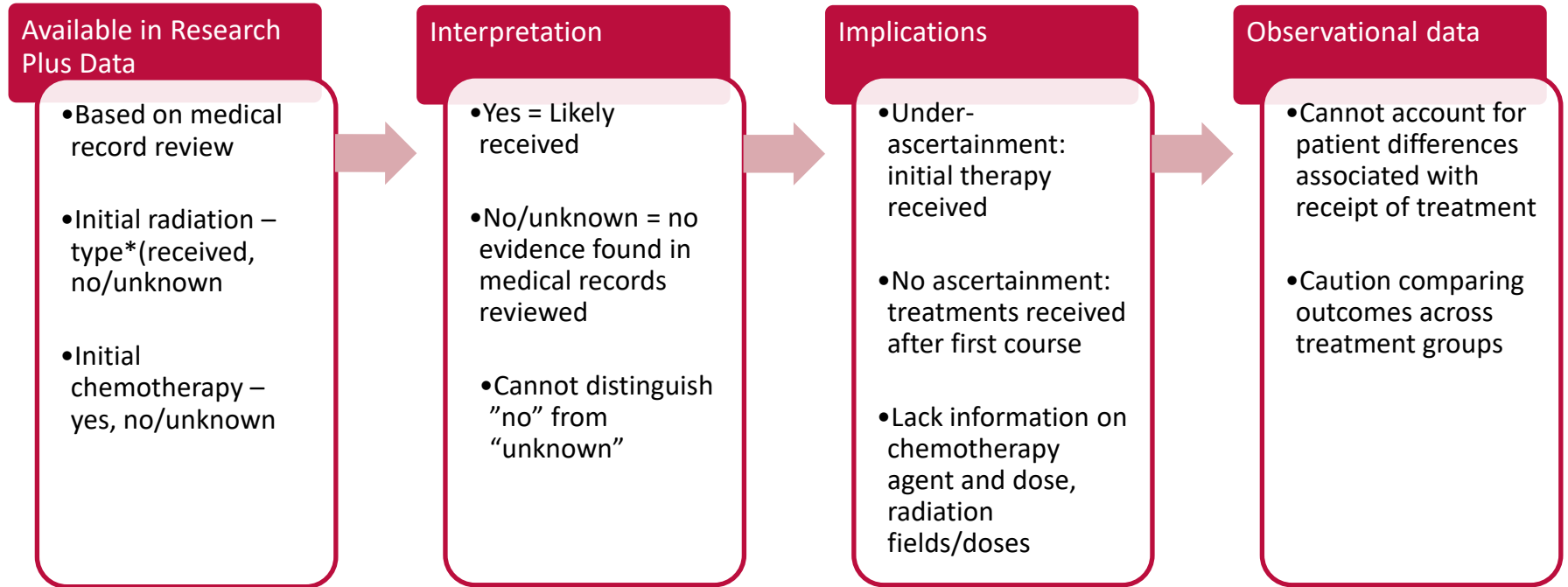
## Additional notes

- 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





# SEER\*Rx

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

Understand changes (since 2005)

Access at:  
<https://seer.cancer.gov/tools/seerrx/>

NIH Surveillance, Epidemiology, and End Results Program

Search SEER

Home Cancer Statistics SEER Data & Software Registry Operations News & Events About

Home / Registry Operations / Tools & Software / SEER\*Rx - Interactive Drug Database

## SEER\*Rx - Interactive Antineoplastic Drugs Database

Updated September 30, 2014 (view [Revision History](#))

**Tools & Software**

- Glossary for Registrars
- ICD Conversion Programs
- SEER\*Rx - Interactive Drug Database
  - Summary of Changes
- Data Documentation & Variable Recodes
- SEER Abstracting Tool (SEER\*Abs)
- SEER Application Programming Interface (API)
- File\*Pro
- SEER Data Management System (SEER\*DMS)

**Important Update:** SEER\*Rx has a new look! SEER\*Rx now has a new and improved search engine that does faster and more intelligent full text searching of all fields, with a sortable results table and a new relevance column so you can tell how relevant each search result is to your entered search string. Additionally, each drug and regimen is now displayed in its own page so that you can bookmark specific entries.

SEER\*Rx was developed as a one-step lookup for coding oncology drug and regimen treatment categories in cancer registries. The information in this database is effective for cancer diagnoses made on January 1, 2005 and after. Review and recoding of drugs from previous years is not required or recommended.

### How to Access SEER\*Rx

The **SEER\*Rx - Interactive Antineoplastic Drugs Database** is provided in a web-based format:

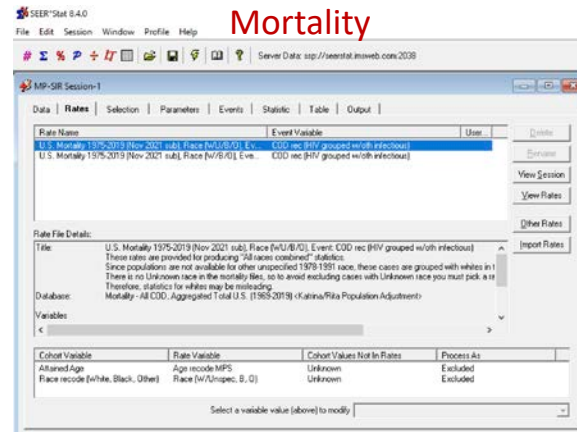
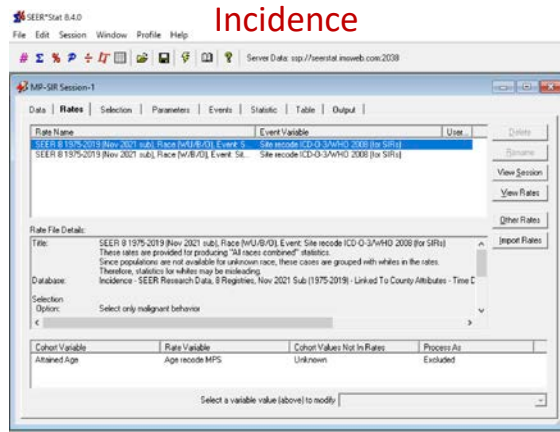
- Updates are automatic: users do not have to install anything to access the latest revisions.
- Allows access from any computer or device with an Internet connection.
- Eliminates problems for users who do not have permission to install software on their work computers.

Please note: The stand-alone version of SEER\*Rx is no longer provided. The web-based tool provides the most up-to-date information.

**Support Resources**

- Questions? [Ask a SEER Registrar.](#)
- [Join the SEER Registrar News listserv](#) to receive announcements of upcoming changes.

## 2. Rate file Selection



### Standard rate files

#### Incidence (SIR)

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 (SMR)

Mortality: Cause of death recode

[https://seer.cancer.gov/codrecode/1969\\_d03012018/index.html](https://seer.cancer.gov/codrecode/1969_d03012018/index.html)

Race (WU/B/O): Unknown race included with white  
Race (W/B/O): Use for analyses where you plan to stratify by race

## 2. Rate file Selection

# Take time to know the rate file

SEER\*Stat 8.4.0  
File Edit Session Window Profile Help  
Server Data: ssp://seerstat.inovweb.com:2038

MP-SIR Session-1

Data | Rates | Selection | Parameters | Events | Statistic | Table | Output

Rate Name	Event Variable	User...	Delete
SEER 8 1975-2019 (Nov 2021 sub), Race (W/U/B/D), Event: S...	Site recode ICD-O-3/A/H/O 2008 (for SIRs)		
SEER 8 1975-2019 (Nov 2021 sub), Race (W/B/D), Event: S...	Site recode ICD-O-3/A/H/O 2008 (for SIRs)		

Rate File Details:

Title: SEER 8 1975-2019 (Nov 2021 sub), Race (W/U/B/D), Event: Site recode ICD-O-3/A/H/O 2008 (for SIRs)  
These rates are provided for producing "All races combined" statistics. Since populations are not available for unknown race, these cases are grouped with whites in the rates. Therefore, statistics for whites may be misleading.

Database: Incidence - SEER Research Data, 8 Registries, Nov 2021 Sub (1975-2019) - Linked To County Attributes - Time C

Selection Option: Select only malignant behavior

Cohort Variable	Rate Variable	Cohort Values Not In Rates	Process As
Allured Age	Age recode MPS	Unknown	Excluded

Select a variable value (above) to modify

**View Session** (indicated by red arrow)

SEER\*Stat 8.4.0  
File Edit Session Window Profile Help  
Server Data: ssp://seerstat.inovweb.com:2038

MP-SIR Session-1

Rate Session-1

Data | Statistic | Selection | Table | Output

Display Variables

- Page
  - Site recode ICD-O-3/A/H/O 2008 (for SIRs)
  - Sex M/F
  - Race recode (W/U/B, D)
- Row
  - Age recode MPS
- Column
  - Year of diagnosis (thru 2019)

Available Variables

- Age at Diagnosis
- Race, Sex, Year Dx
- Site and Morphology
- Stage - Summary/Historic
- Stage - 8th edition
- Stage - 7th edition
- Stage - 6th edition
- Stage - Other
- Therapy
- Site-Specific Data Items
- Extent of Disease

Move Up, Move Down, Remove, Page, Row, Column, End...

## 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

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

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

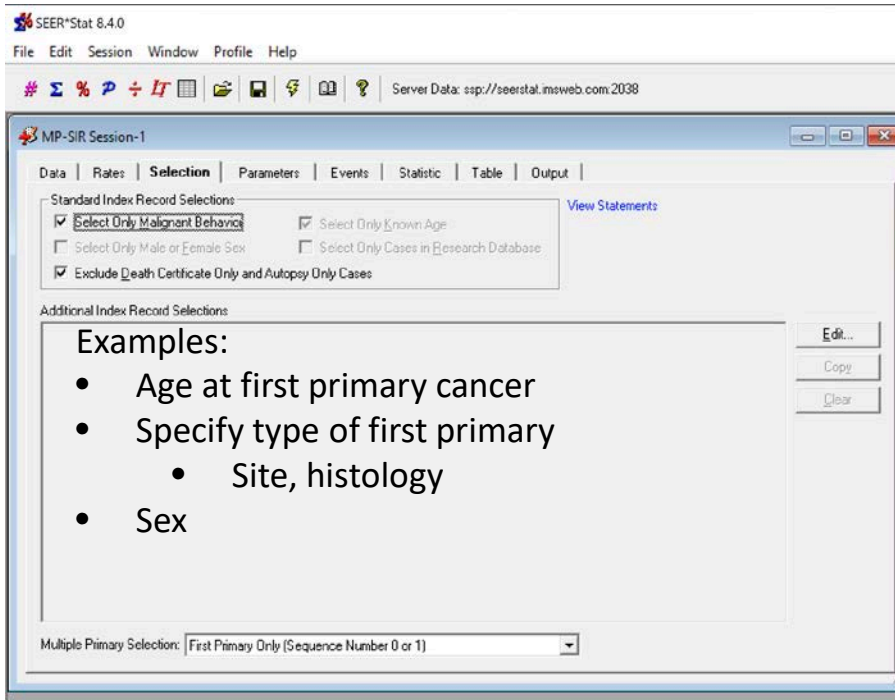
Events based on other SEER recodes such as:

- AYA Site Recode
- International Classification of Childhood Cancer
- Lymphoma Subtype Recode

Alt. grouping population characteristics

### 3. Study population selection

## 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

## 4. Parameters

SEER\*Stat 8.4.0  
File Edit Session Window Profile Help  
# Σ % P ÷  $\int$  Server Data: ssp://seerstat.imswb.com:2038

MP-SIR Session-1

Data Rates Selection **Parameters** Events Statistic Table Output

Study Time  
Exposure Date:   
Latency Exclusion Period (Months):   
Start Date:   
Cutoff Dates  
Month Year  
Start:    
End:

Cut-Points for Time-Dependent Variables:  
Latency (months and/or years):   
(e.g. 6m,1y,5y,10y)  
Attained Age:   
Attained Calendar Years:

- 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



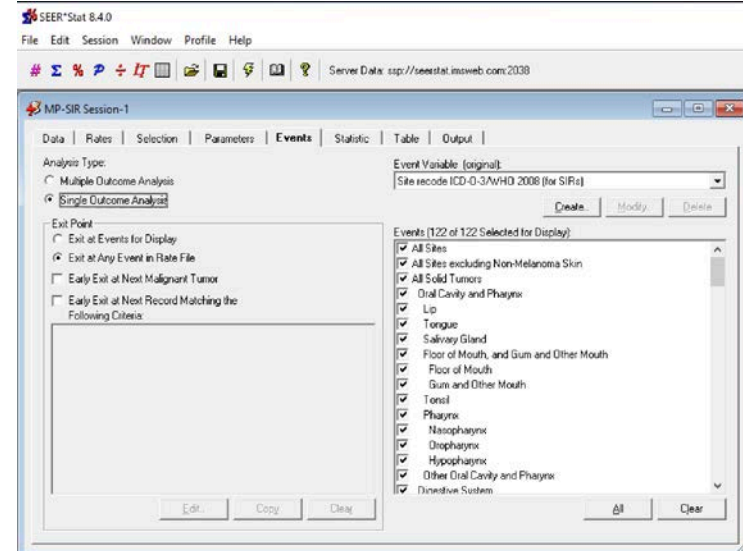
# Multiple vs Single Outcome Analysis (SIR only)

## 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



5. Events

# Example: Subsequent cancer incidence after first primary thyroid cancer

1<sup>st</sup>: thyroid

2<sup>nd</sup>: colon

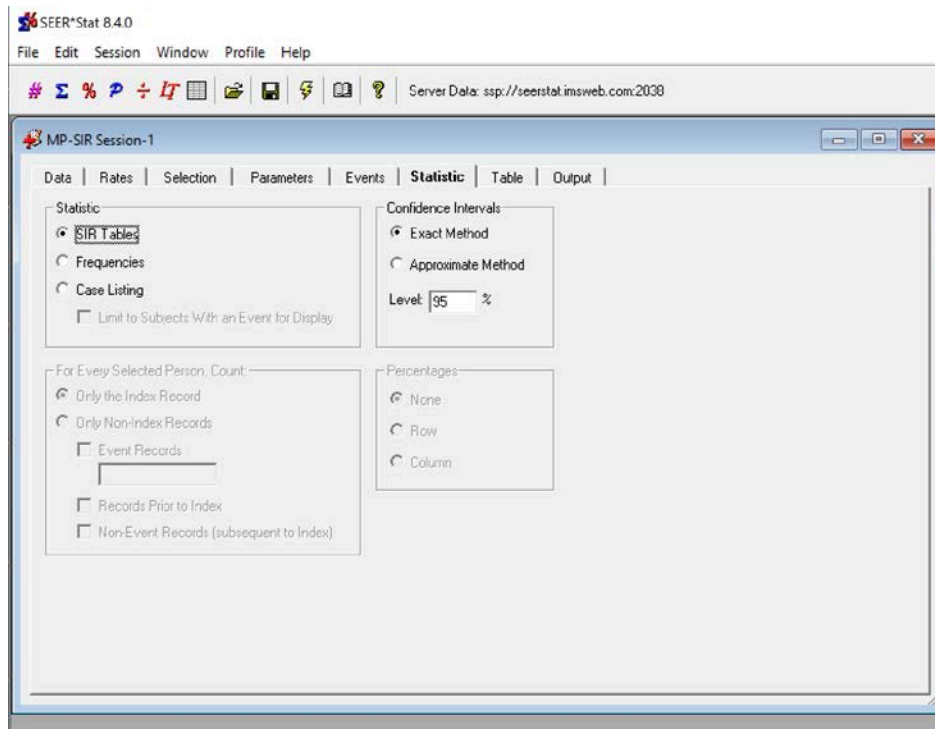
3<sup>rd</sup>: breast

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



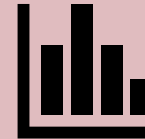
## Objective

- 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

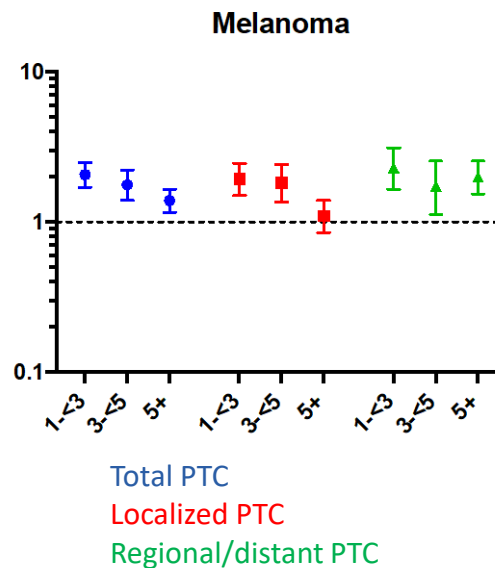
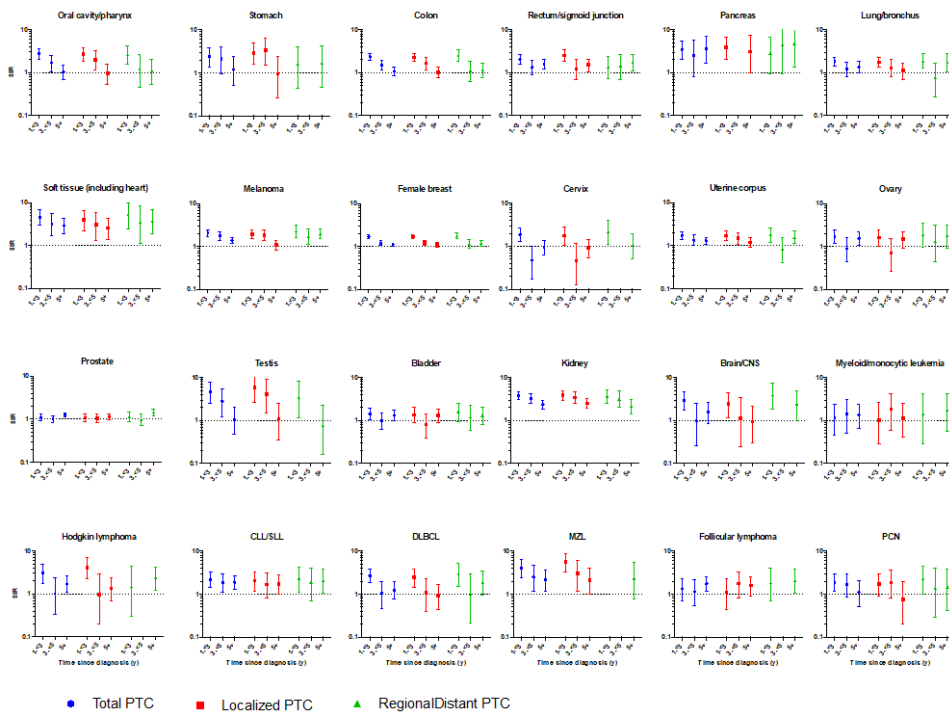
- 3,175,216  $\geq$  1-year adult survivors of non-thyroid malignancies
- First primary diagnosis: 2000-2014, followed through 2015 (SEER 18)
- Ages 20-84 at first primary diagnosis



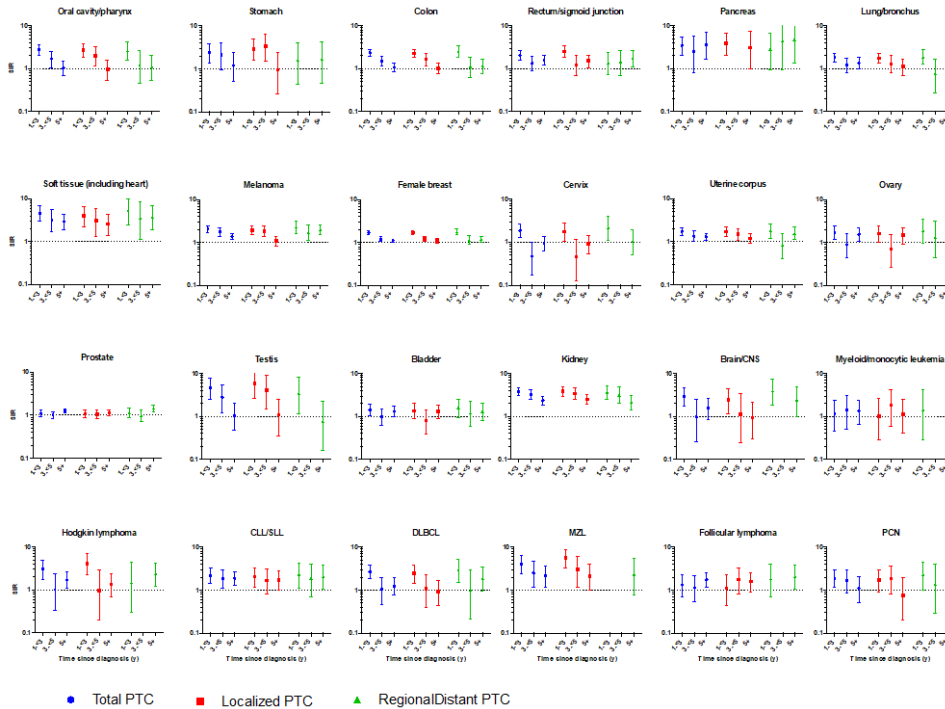
## Statistical analyses

- SIRs (second cancer)
- Poisson regression – tests of trends, heterogeneity (not in SEER\*stat)

# SIRs by latency and stage of PTC



# 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



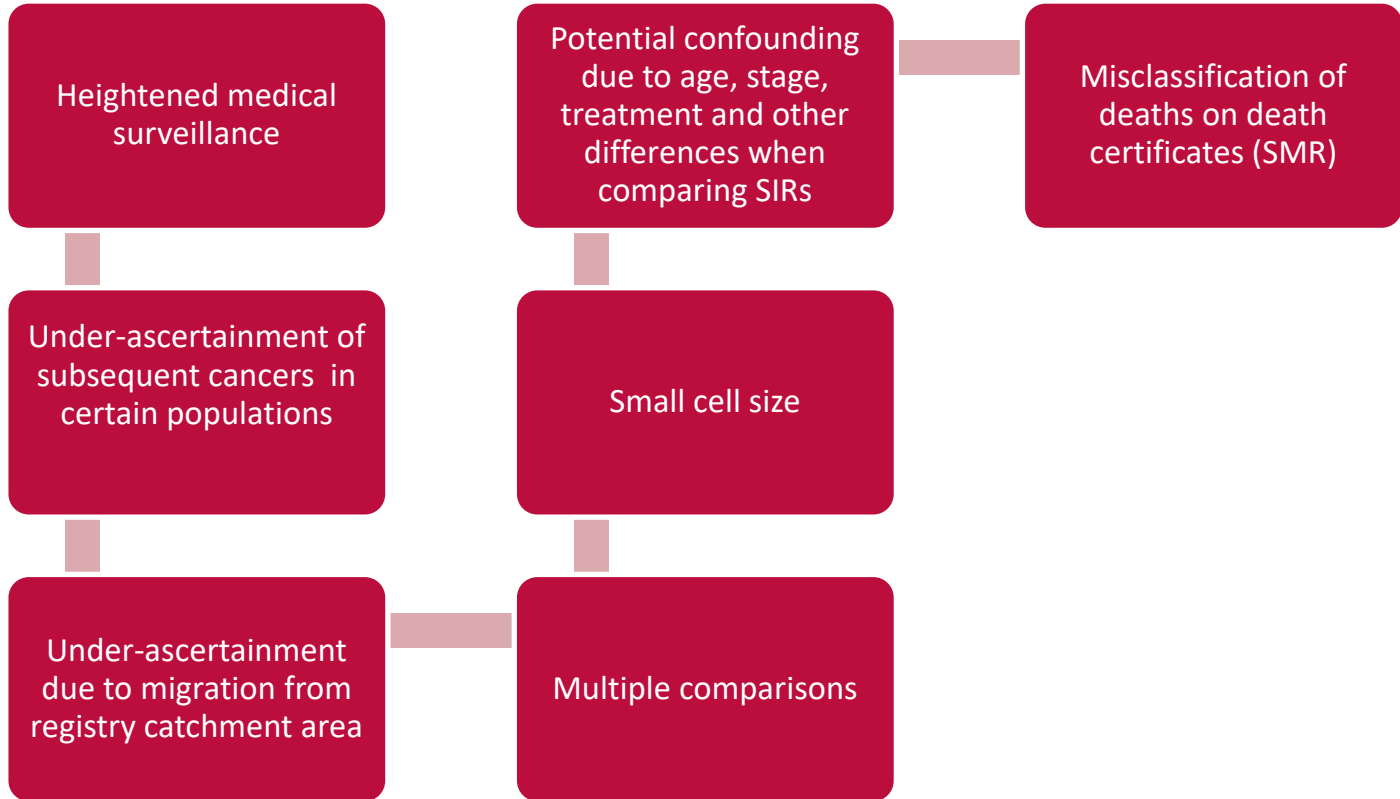
# Leading causes of excess deaths after cHL

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

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

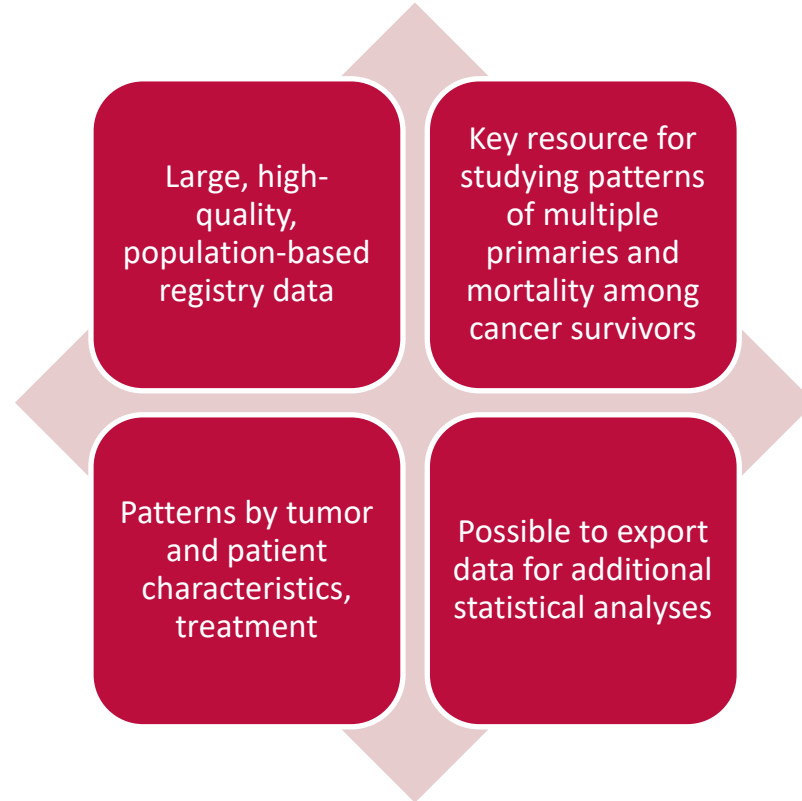
# Summary of limitations and strengths

# Be Mindful...

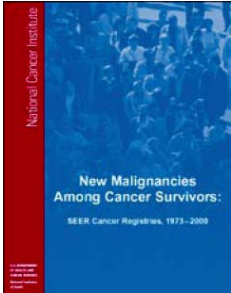


Curtis et al. 2006

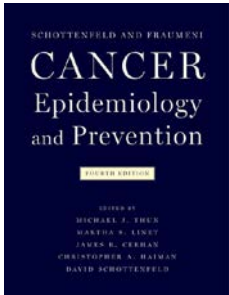
# SEER MP-SIR has great strengths!



# Resources



Curtis et al. 2006 <https://seer.cancer.gov/archive/publications/mpmono/index.html>



Chapter 60: Multiple Primary Cancers

SEER\*stat Tutorials (<https://seer.cancer.gov/seerstat/tutorials/advanced.html>)

# 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*

May 5, 2022

NCI Analytic Tools SEERies



**NATIONAL  
CANCER  
INSTITUTE**

[www.cancer.gov](http://www.cancer.gov)

[www.cancer.gov/espanol](http://www.cancer.gov/espanol)