Developing and Working with Survival Data

NAACCR 2010-2011 Webinar Series

NAACCR

Questions

NAACCR

Fabulous Prizes!

Question of the Month!
- The participant that submits the best question of the session will receive a fabulous Prize!

Tip of the Month!
- The participant that sends in the best tip related to the topic will win a spectacular prize!



Agenda

- Overview
- Need for death clearance
- Survival and NDI
- Data quality
- Break
- SEER*Prep; SEER Stat
- Survival issues:
- NAACCR 2011
- Closing remarks

NAACCR

Survival statistics, and surviving statistics! An overview and update about cancer survival rates June 2, 2011

Dr. Donna Turner, Epidemiologist
Provincial Director, Population Oncology
Cancer Care Manitoba

Dr Hannah Weir, Epidemiologist
Division of Cancer Prevention and Control
Centers for Disease Prevention and Control

NAACCR

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.

Overview

- The evolution of population-based cancer survival
- EUROCARE and CONCORD studies
- Cancer Control in the UK and Canada
- Useful websites

NAACCR

Clinical vs. Population-based Survival

- Clinical trials - highest ***achievable*** survival
  - Patient focus "How long do I have, doc?"
  - Clinical focus Value of one treatment vs. another
- Population - survival ***achieved***
  - Impact of cancer control initiatives (across the spectrum of initiatives)
    - Targeting and monitoring cancer control initiatives
  - Policy-setting
    - Effectiveness of healthcare delivery - standard measure of cancer system performance

NAACCR

Population-based Cancer Survival

Why are there variations in cancer survival?*

- Timely diagnosis and good prognosis …
  - Stage of cancer at diagnosis
  - Screening (availability, access and participation)
  - Diagnostic access
  - Public's awareness of cancer symptoms
  - Types of cancer/disease diagnosed (aggressive variants
- Appropriate treatment …
  - Equitable access to treatment
  - Implementation of best practices (use of practice guidelines)
  - Organization of treatment services (timeliness, smooth transition)
- Access to healthcare (insurance) and human and financial resources

NAACCR

**Adapted from Coleman MP: Opinion: why the variation in breast cancer survival in Europe? [commentary]. http://breast-cancer-research.com/vol1no1/07oct99/editorial/*

Population-based Cancer Survival

- Crude survival:
  … how many individuals diagnosed with cancer are alive after five years?
  … endpoint is death from any cause
- Cause-specific survival:
  … how many individuals diagnosed with
  of cancer after xx years?
  … endpoint is death from cancer only

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

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

* Life tables

NAACCR

Advantages and Disadvantage of Relative vs. Cause Specific Survival

	Advantage	Disadvantages
Relative	Relies on fact of death not cause of death Enables estimation of avoidable deaths (excess mortality)	Life tables may not be available for all populations
Cause Specific	Not limited to populations with life tables	Death Certificates are not reliable (e.g., site of mets or recur)

NAACCR

Population-based Survival - Focus on Relative Survival *(Example)*

Suppose that in a jurisdiction far, far away …

Five-year survival is 60% for women aged 15-99 diagnosed with breast cancer

but

Five-year survival is only 80% for women in general

then

Relative survival is 60% / 80% or 75%.

NAACCR

Additional advantages of Relative Survival*

- Answers the question "how much is [my/my patient's] survival decreased as a result of a cancer diagnosis?"
  - speaks directly to **excess mortality** among cancer patients[1]
- Adjusts for increasing "background" mortality in a population
  - accounts for the fact that our risk of death increases as we age, whether we have cancer or not
- Adjusts for differences in "background" mortality between populations
  - allows assessment of differences in cancer survival between populations that might have large variations in mortality generally[2] (e.g.. racial/ethnic differences, international comparison, etc.)

**Adapted from Rachet B, Woods LM, Mitry E, Riga M, Cooper N, Quinn MJ, Steward J, Brenner H, Estève J, Sullivan R, Coleman MP. Cancer survival in England and Wales at the end of the 20th century. Br J Cancer 2008; 99, S2 – S10.*

[1]Estève J, Benhamou E, Croasdale M, Raymond L. Relative survival and the estimation of net survival: elements for further discussion. Stat Med. 1990;9:529-538.

[2]Micheli A, Baili P, Mugno E, Queen M, Capocaccia R, Grosclaude PC,EUROCARE Working Group. Life expectancy and cancer survival in the EUROCARE-3 cancer registry areas. Ann Oncol 2003; 14(Suppl 5): 28-40.

NAACCR

Relative survival: cohort and period approaches

- The basic cohort method[1-3]
  - Uses everyone diagnosed with cancer in the past, who has had sufficient follow up time
  - Traditional approach to survival statistics; reflect the survival expectations of patients diagnosed many years ago (i.e., everyone in the cohort must have had five years of follow up)

[1]Berkson J, Gage RP. Calculation of survival rates for cancer. Proc Staff Meet Mayo Clinic 1950;25:270-286.

[2]Cutler SJ, Ederer F. Maximum utilisation of the life table method in analyzing survival. J Chron Dis 1958;8:699-712.

[3]Ederer F, Axtell LM, Cutler SJ. The relative survival: a statistical methodology. Natl Cancer Inst Monogr 1961;6:101-121.

NAACCR

Relative survival: cohort and period approaches

- The Period approach[1]
  - Provides more 'up-to-date' estimates of long-term survival rates, incorporates the survival experience of recently diagnosed cases into the analysis.
    e.g., 5-year survival for people diagnosed 2003-2007, with follow-up to the end of 2007

    1-year estimate will include the 1-year survival experience of people diagnosed in 2003-2007
    2-year estimate will include the survival experience for people diagnosed in 2003-2006
    3-year estimate will include 2003-2005 follow-up,
    …. And so on.

[1]Brenner H, Gefeller O. An alternative approach to monitoring cancer patient survival. Cancer 1996;78: 2004–2010.

NAACCR

Relative survival estimates: still evolving

- *One primary or multiple primaries*
- SEER vs. IARC rules for multiple primaries

NAACCR

A Tale of Two Studies, Two Countries and Action Plans

EUROCARE Studies
- U.K. - NHS Cancer Plan
- International Benchmarking Study

CONCORD Study
- High Resolution (Patterns –of-care) studies
- The Canadian Strategy for Cancer Control: a cancer plan for Canada
- Canadian Partnership Against Cancer (CPAC)
  - C-SPAN

NAACCR

EUROpean CAncer REgistry-based study on survival and care of cancer patients

- Initiated in Italy (1989)
  - Istituto Nazionale Tumori (Milan)/Istituto Superiore di Sanità (Rome)
  - 12 population-based (European) cancer registries
- Versions …
  - EUROCARE-1 (1978-1984)
  - EUROCARE-2 (1978-1989)
  - EUROCARE-3 (1983-1994)
  - EUROCARE-4 (1988-2002)
  - EUROCARE-5 (2000-2007)
- Now includes 93 population-based registries in 23 European countries
- Objective of EUROCARE-5: To update the existing EUROCARE data bank by including data of patients diagnosed up to 2007. Follow up will be updated to the most recent possible dates in order to analyze both long and short term survival rates of cases diagnosed more recently.

NAACCR

EUROCARE: Findings

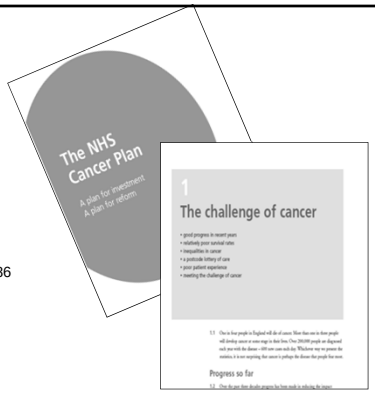
- Survival for most solid tumours (breast, colorectal, stomach, cutaneous melanoma) was:
  - highest in Finland, Sweden, Norway and Iceland
  - **lower in the UK and Denmark**
  - lowest in the Czech Republic, Poland and Slovenia
- Countries with higher expenditure on health generally had best survival (exceptions: **Denmark and UK**)
- Survival for Europe lower than for the US for nearly all cancers

Sant M, Allemani C, Santaquilani M, Knijn A, Marchesi F, Capocaccia R, and the EUROCARE Working Group. EUROCARE-4. Survival of cancer patients diagnosed in 1995-1999. Results and commentary. Eur J Cancer 2009,45:931-991.

NAACCR

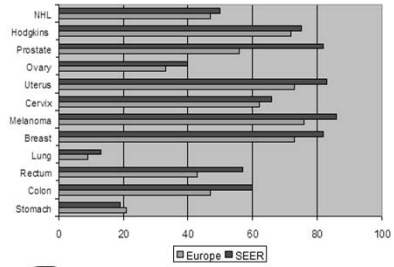
"If the survival rates among the poorest matched those among the richest in England and Wales, 12,700 untimely deaths could have been prevented amongst those diagnosed between 1986 and 1990."

'Challenging cancer', Dept of Health, May 1999



NAACCR

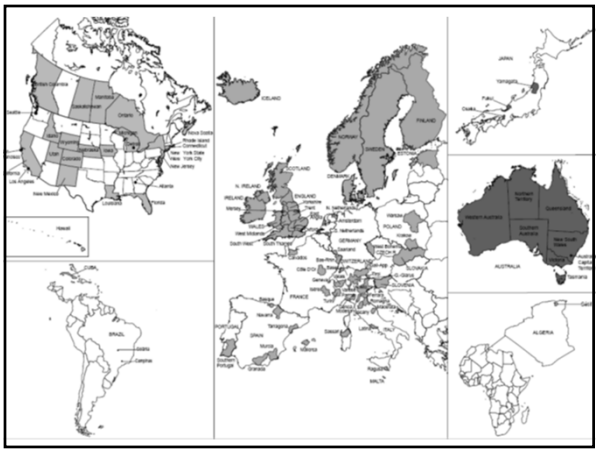
EUROCARE vs. US (SEER)
5- year Relative Survival

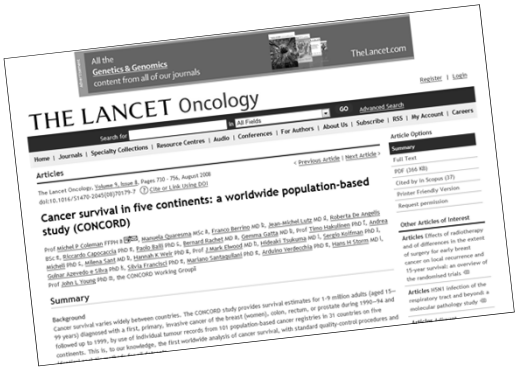


NAACCR

Gatta G, Capocaccia R, Coleman M, Reis L, et al. *Toward a Comparison of Survival in American and European Cancer Patients.* Cancer 2000 I89(4): 893-900

NAACCR

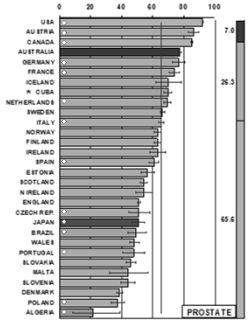




NAACCR

US Paradox – High Survival Rates Worldwide….

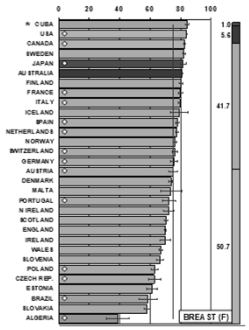
5-year relative survival (%) -
prostate cancer,
(15-99 years)



NAACCR

US Paradox – High Survival Rates Worldwide….

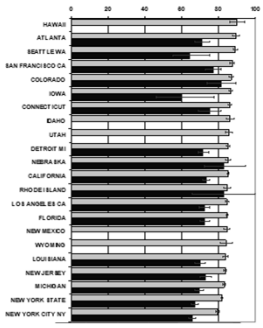
5-year relative survival (%) -
female breast cancer,
(15-99 years)



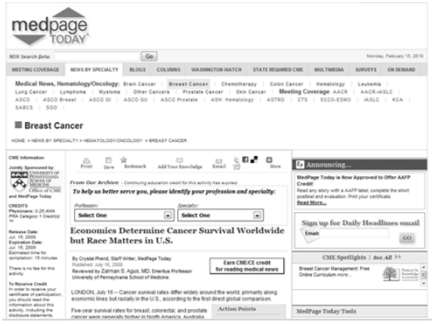
NAACCR

But Large and Consistent Racial Disparities ….

5-year relative survival (%) - female breast cancer, (15-99 years) in US by race



NAACCR



NAACCR

Canadian Strategy for Cancer Control

The Canadian Strategy for Cancer Control: A Cancer Plan for Canada

Discussion Paper

NAACCR

Canadian Strategy for Cancer Control

The Ca

Indicators - System Performance Report - 2009

Cancer Control Continuum	Indicator	Database			
		CCHS	CCR	CBCSD	P/T Cancer Agencies or Similar
PREVENTION	Smoking Prevalence	✓			
	Smoking Quit Attempts	✓			
	Overweight and Obesity	✓			
	Alcohol Consumption	✓			
SCREENING	Mammography Rates	✓		✓	
	Cervical Screening Rates	✓			
	Colorectal Screening Rates	✓			
DIAGNOSIS AND TREATMENT	Wait Times: Abnormal breast screen to resolution			✓	
	Wait Times: Radiation Treatment				✓
	Radiation Therapy				✓
SUPPORTIVE CARE & SURVIVORSHIP	Symptom Assessment				✓
LONG-TERM OUTCOMES	Cancer Incidence Rates		✓		
	Cancer Mortality Rates		✓		
	Relative Survival Rates		✓		
KNOWLEDGE MANAGEMENT/SURVEILLANCE	Capture of Stage Data				✓
RESOURCES/ CAPACITY	PET Scanner Capacity				✓
	Radiotherapy Capacity (LINACS)				✓

For a full description of indicator definitions, please refer to the Technical Appendix at the end of this report.

NAACCR

About Us Courses Research Departments News & Events

You are here: Home > IPH > NCDEU > Cancer Survival Group

CONCORD study: cancer survival world-wide

CONCORD homepage
Pilot study
Phase 1
Phase 2
Steering
Cancer

The CONCORD study was planned in three phases:

Phase 1 (low resolution) was designed to quantify international differences in population-based relative survival by age, sex, country and region for patients diagnosed during 1990-94 with a cancer of the breast, colon, rectum or prostate. The study was published online in *Lancer Oncology* on 17 July 2008.

Phase 2 (high resolution) was designed to help interpret those international differences in survival, by using a sub-set of registries that could re-abstract detailed clinical data, including stage at diagnosis and treatment, from the original medical records, for large random samples of patients diagnosed with one of the same cancers during 1996-98. Data collection in a number of US and European registries is complete.

Phase 3 was designed to [illegible] pathological diagnosis for a sub-set of patients from Phase 2, to assess the extent to which international survival differences [illegible] differences in the pathological definition of disease in participating countries. This study has [illegible]

CONCORD – 2 Study

- 150+ registries on all 5 continents
- 10 cancers of public health importance in low, medium and high income countries
- Diagnosed 1995-2007 with follow-up through 2007+
- Monitor UICC World Cancer Declaration Target 11
- Call for data: 2011-2012

NAACCR

Interesting cancer survival websites (check it out)

- EUROCARE: www.eurocare.it
- Paul Dickman (www.pauldickman.com) (Sweden)
- International Agency for Research on Cancer (IARC) http://www.iarc.fr/
- UK Cancer Survival Group: www.lshtm.ac.uk/ncdeu/cancersurvival/
- SEER: www.seer.gov/cancer
- Statistics Canada: www.statcan.gc.ca/
- Canadian Partnership Against Cancer: www.partnershipagainstcancer.ca
- Portal: CancerViewCanada: www.cancerview.ca

NAACCR

Death Clearance

Key Component to Developing Survival Statistics

NAACCR

Objectives

- Describe Death Clearance
- Function of Death Clearance
- Importance to Survival Analysis
- References

NAACCR

Death Clearance Process

- Identify Death to Cancer Patients
  - Link to Mortality Files
  - Update Vital Status
  - Identify Missed Cases
    - Unreported Patient
    - Unreported Multiple Primary
- Follow Back Unlinked Cancers
  - Confirmation of Condition
  - Residence at Diagnosis
  - Case Details

NAACCR

Death Clearance in Canada

- Local (provincial) death clearance
- National Statistics Canada

NAACCR

Example: Saskatchewan Death Clearance

NAACCR

Why Death Clearance

- Originally developed by NCI-SEER
  - Avoid unnecessary patient follow up
  - Establish vital status
  - Identify unreported cancer cases
- Critical to Survival Statistics
  - Key for Active and for Passive Follow Up
    - Reduce follow up cost for active follow up
    - Substitute for active follow up if passive
      - Must combine with NDI, SSDI or other
  - Identifies *biased* group of unreported case
    - Without death clearance will overstate survival

NAACCR

What about Cancer Types for…. *Death Certificate First* Cases?

Site	Cases	Death 1st	% Death 1st
Female Breast	86,206	2,328	2.7
Prostate	82,084	2,184	2.7
Colorectal	56,366	4,446	7.9
Lung	74,045	15,818	21.9
Pancreas	11,278	3,708	32.9
Esophagus	5,469	943	17.2

Michigan Resident Cases Diagnosed between 1994-2003
Death certificate first is a case first identified through death clearance.

NAACCR

What about ……Stage at Diagnosis?

Site	Late Stage	
	All Cases	Death 1st
Female Breast	22.1	31.6
Prostate	11.7	23.1
Colorectal	47.1	55.5
Lung	63.9	61.5
Pancreas	68.2	58.2
Esophagus	47.2	50.0

Michigan resident cases diagnosed between 1994 and 2003
with regional or distant stage at diagnosis

NAACCR

What about ……..demographics?

Percent Death Certificate 1st by age and Race

Site	Percent by Age		Percent by Race	
	65 and Younger	Over 65*	Black	White
Female Breast	1.1	4.7	2.7	2.6
Prostate	1.0	7.7	2.7	2.7
Colorectal	4.7	9.5	8.1	7.9
Lung	17.1	58.9	19.6	21.6
Pancreas	25.8	36.5	28.8	33.8
Esophagus	14.7	19.1	18.9	17.0

* Over 75 for prostate cancers

Michigan resident cases diagnosed between 1994 and 2003

NAACCR

Death Clearance is Required

- Required by NPCR
- Required by NCI/SEER
- NAACCR Standard Requirement
- Necessary for NAACCR Certification
  - Completed within 23 months

NAACCR

Death Clearance is Required

- NAACCR Standard I.B.9
  - Must
    - Be able to perform mortality linkage
    - Have adequate staff for follow back
  - Should
    - Establish formal agreement with vital records
    - Track progress and results
    - Follow back on potential multiples
    - Provide quality control feedback
      - *Identify case-finding issues*

NAACCR

What have you done ……
…… when you are done?

- Confirmed Vital Status
  - Nearly all deceased patients
  - 97% of Deaths to Cases in Michigan
  - Provide vital status updates to reporters
- Acquired Critical Missed Cases
  - Poor Prognosis
  - Tend to be Older
  - Geographically Biased
- Assured Most Accurate Surviv
  - Prompt improved reporting of clinical diagnoses

NAACCR

References

- NAACCR Death Clearance Manual
  - http://www.naaccr.org/StandardsandRegistryOperations/RegOpsGuidelines.aspx
- NAACCR Standards Vol. 3 – pp 20-21
  - http://www.naaccr.org/StandardsandRegistryOperations/VolumeIII.aspx
- SEER Data Management System – Chapter 17
  - http://www.seer.cancer.gov/seerdms/manual/

NAACCR

Best Practices for Developing and Working with Survival Data:

NDI Linkages: What they are and why they matter.

Monique Hernandez, PhD
Chris Johnson, MPH
Brad Wohler, MS

NAACCR

Outline

- Brief overview of NDI linkages.
  - For more detailed information, see http://www.naaccr.org/AboutNAACCR/TownMeetings.aspx
  - http://www.cdc.gov/nchs/ndi.htm
- Examples of impact of NDI linkages on population-based survival measures.
  - CONCORD
  - Florida Cancer Data System NDI Linkage and Survival Project
  - Accuracy of Cancer Mortality Study
    - California, Colorado, Idaho

NAACCR

49

Thanks!

Lyn Almon, Georgia Comprehensive Cancer Registry
Chris Johnson, Cancer Data Registry of Idaho
Robert Bilgrad, National Death Index
Glenn Copeland, Michigan Cancer Surveillance Program
Monique Hernandez, Florida Cancer Data System
Colleen McLaughlin, New York State Cancer Registry
Hannah Weir, Centers for Disease Control and Prevention
Brad Wohler, Florida Cancer Data System

NAACCR

50

National Death Index

- The National Death Index (NDI) is a centralized registry maintained by the National Center for Health Statistics of all deaths that have occurred in the United States, Puerto Rico, and the Virgin Islands since 1979.

NAACCR

51

National Death Index - Purpose

- Identifies deceased study subjects
- Provides the following:
  - dates of death
  - states of death
  - death certificate numbers

52 NAACCR

National Death Index - Coverage

- All 50 states, District of Columbia, NYC, Puerto Rico, & Virgin Islands
- 65 million NDI records
- All deaths from 1979-2008
- 2009 deaths expected July 2011

53 NAACCR

NDI PLUS

- Implemented in 1997
- Provides researchers with
  - Underlying cause of death codes
  - Multiple cause codes
  - ICD-9 and ICD-10 codes

54 NAACCR

Death Clearance Safety Net

- Late cases
  - Cancer registry
  - Vital Statistics
- Missed cases
  - Out of state

NAACCR

55

National Death Index - Process

- Select candidate records for submission to NDI – unknown vital status
- Run EDITS, Inter-Record Edits
- Cut file using NPCR Extract utility
- Complete forms and submit them with data
- <NDI processes file>
- Receive results from NDI
- Process results using SAS algorithm available from NPCR docserver
- Manual review component
- Update central registry database with NDI results
- Data sharing with other states

NAACCR

56

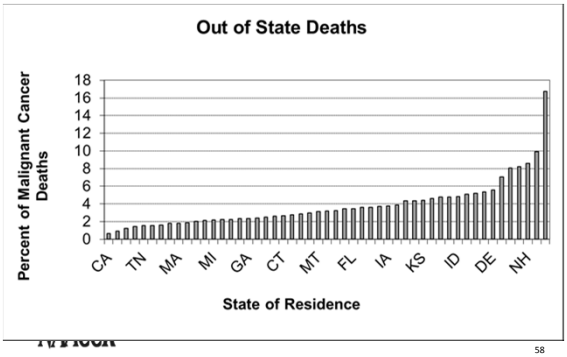
Evidence

- Indirect
- CONCORD
- Florida
- ACM

NAACCR

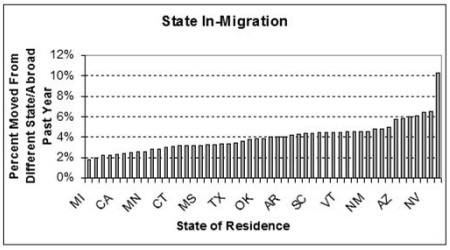
57

Indirect Evidence



58

Indirect Evidence



- From 2005-2007, 12 states had at least 5% of their population migrate in from another state or abroad.

59

Indirect Evidence

- Americans very mobile
- May change residence state after dx
- Don't always die in their state of residence.
- Death missed during death ascertainment
  - ↓ event count for survival calculations (numerator)
  - ↑ increase in follow-up time (denominator).

60

CONCORD

- Cancer survival in five continents: a world-wide population-based study
  - British Columbia, Manitoba, Nova Scotia, Ontario, Saskatchewan
  - California, Colorado, Connecticut, Florida, Georgia - Atlanta SEER, Hawaii, Idaho, Iowa, Louisiana, Michigan, Nebraska, New Jersey, New Mexico, New York, Rhode Island, Utah, Washington - Seattle SEER, Wyoming
- In the U.S., NDI linkages were required.
  - NDI Plus not conducted – no cause of death information, so data not useful for cause-specific survival.

NAACCR

61

CONCORD NDI Results (Partial)

Results of CONCORD NDI Match Resolution Algorithm

State:	Michigan	Idaho	NYS	Nebraska	CA	Colorado	Florida
Unknown status	31,471	6,263	91,497	9,027	42,377	16,198	105,575
Matches returned	56,042	9,774	209,615	11,446	88,396	25,701	282,106
True matches	1,131	444	5,907	166	3,364	191	7,690
True Match %	3.6%	7.1%	6.5%	1.8%	7.9%	1.2%	7.3%

NAACCR

62

CONCORD

- States submitted some cases above and beyond those sites required for the CONCORD study.
- In both Idaho and Florida, about 25% of the total NDI matches were among in-state deaths.
  - In Florida, 12% of total NDI matches were NY deaths.
  - In Idaho, 28% of total NDI matches were WA deaths.

NAACCR

63

Evidence From CONCORD

ncrease in Percent of CONCORD Cases Deceased

State	Prostate + NDI	Breast (Female) + NDI	Colorectal (Male) + NDI	Colorectal (Female) + NDI	CONCORD Total + NDI
California	1.6%	1.1%	1.7%	1.6%	1.4%
Colorado	0.7%	0.4%	1.2%	0.8%	0.7%
Florida	4.5%	3.4%	5.3%	5.0%	4.3%
Idaho	0.8%	0.6%	1.0%	1.4%	0.8%
Michigan	2.4%	1.7%	2.0%	2.0%	2.1%
New York State	2.6%	2.1%	2.7%	2.8%	2.5%
NPCR CONCORD Avg	1.5%	0.8%	1.6%	1.9%	1.2%

NAACCR

64

Florida Cancer Data System NDI Linkage and Survival Analysis Project

- 1981 - 2005
- 1,115,558 records submitted to NDI

NAACCR

65

FCDS NDI Linkage and Survival Analysis Project: Review NDI Results

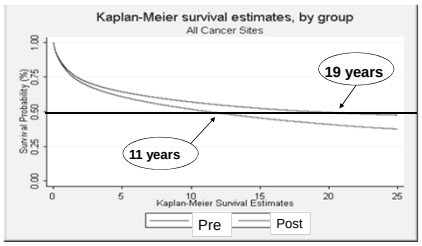
- Bottom Line
  - Vital status changed to deceased = 125,648 patients
    - Affecting 147,211 tumors
- Death clearance safety net = 9,854
- Remaining 115,794 died out of state
  - FL VS does not re-release info on FL residents who died out of state

NAACCR

66

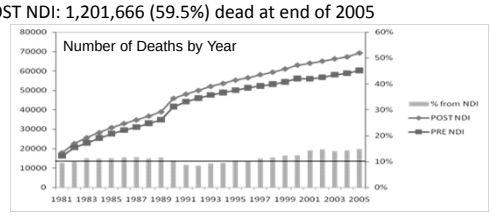
FCDS NDI Linkage and Survival Analysis Project:
Kaplan-Meier Survival Curve

- Median survival time: The time at which exactly half of the population has survived



FCDS NDI Linkage and Survival Analysis Project

- 1981-2005 there were 2,020,387 people DX with >= 1 tumor in FL
- PRE NDI: 1,076,018 (53.5%) dead at end of 2005
- POST NDI: 1,201,666 (59.5%) dead at end of 2005



FCDS NDI Linkage and Survival Analysis Project:
Percent Patient Survival Pre and Post at Time T in Years

Survival Estimates										
	All		Lung		CRC		Prostate		Breast	
years	pre	post	pre	post	pre	post	pre	post	pre	post
5	64	61	28	21	70	66	90	89	88	86
10	57	52	24	16	65	60	83	79	81	78
15	52	46	23	14	64	58	77	70	77	72
20	49	41	22	13	63	57	73	62	74	68

Site with most impact: lung and bronchus cancer
Site with least impact: Breast cancer

Greater than 5% difference

Accuracy of Cancer Mortality Study

- What it is?
- California, Colorado, Idaho
- The Impact of National Death Index Linkages on Population-Based Cancer Survival Rates.
  - A separate data collection effort merged cancer registry data to the National Death Index (NDI) to find deaths that occurred out of state and to obtain cause-of-death information for these deaths.
  - State vital statistics linked deaths were thus augmented with linkages to the National Death Index (NDI).

NAACCR

70

Accuracy of Cancer Mortality Study

- The Impact of National Death Index Linkages on Population-Based Cancer Survival Rates
- We investigated the impact on 5-year cancer survival rates of performing the NDI linkage component of the ACM study.
  - 1993-1995 cases with linkages to state vital statistics and NDI through 2004
  - Measured the impact of NDI linkages on cause-specific and relative cancer survival statistics.

NAACCR

71

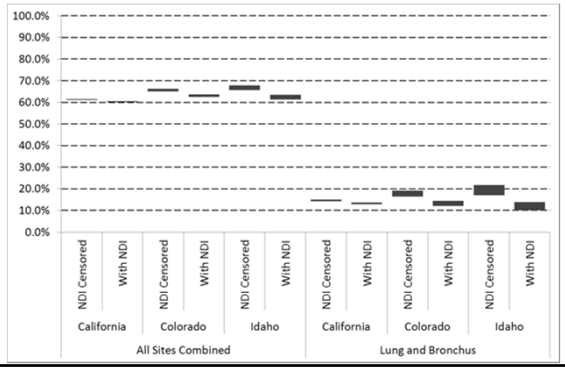
Accuracy of Cancer Mortality Study Impact of NDI Linkages on Survival Statistics

- Two datasets created:
- One dataset included deaths ascertained through state vital records linkages augmented with deaths ascertained through NDI linkages.
- The second dataset included only deaths ascertained through state vital records linkages;
  - all NDI deaths were censored at the end of the study period (vital status alive as of Dec 31, 2004), as if the NDI linkages had not been performed.

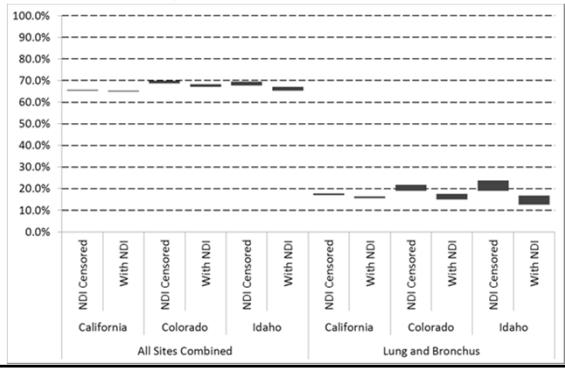
NAACCR

72

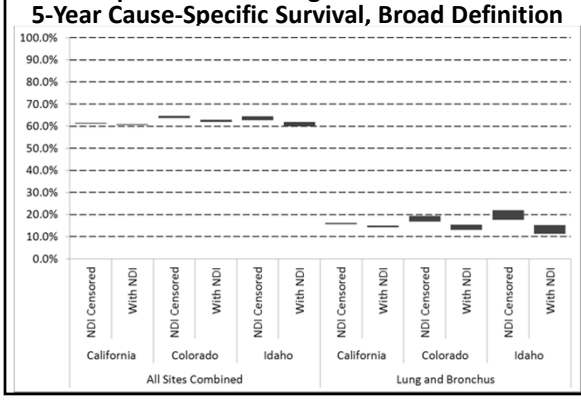
ACM - Impact of NDI Linkages on Survival Statistics
5-Year Relative Survival



ACM - Impact of NDI Linkages on Survival Statistics
5-Year Cause-Specific Survival, Narrow Definition



ACM - Impact of NDI Linkages on Survival Statistics
5-Year Cause-Specific Survival, Broad Definition



Conclusions

- Annual linkage of central cancer registry data with NDI data is highly recommended.
- Death clearance safety net
- Access to info on state residents who die out of state
- Access to info on cases who move out of state after dx
- Fee support via CDC or NCI
- The benefits of NDI linkage include improved follow-up for more accurate survival statistics

NAACCR

76

Thanks!

Lyn Almon, Georgia Comprehensive Cancer Registry
Chris Johnson, Cancer Data Registry of Idaho
Robert Bilgrad, National Death Index
Glenn Copeland, Michigan Cancer Surveillance Program
Monique Hernandez, Florida Cancer Data System
Colleen McLaughlin, New York State Cancer Registry
Hannah Weir, Centers for Disease Control and Prevention
Brad Wohler, Florida Cancer Data System

NAACCR

77

Best Practices for Developing and Working with Survival Data:

Data Quality for Survival Analysis

Contributors/Presenters: Katherine Fradette, Deborah Hurley, Hannah Weir, Donna Turner

NAACCR

Data Quality for Survival: Two Main Considerations

1. The quality of information about the cases
   - Missing, incomplete or poor quality reporting of cases can lead to a biased picture of survival
2. The quality of the death data
   - Missing, incomplete or poor quality reporting of death information can also lead to a biased picture of survival (usually over-estimation)

NAACCR

Input Quality Affects Output Quality

For each regional registry included in analysis, the quality and comprehensiveness of information about cases and deaths is of primary importance

- Type of follow-back
- Routine data quality checks and clean-up
- Coding Rules
- Death related information

NAACCR

Type of Follow-Back: Sources

Active follow-back (medical records)

- Cancer Registry Initiated
  - Contact physician or reporting hospital
  - National, State or Province data exchange agreements
- Hospital/Physician Office Initiated
  - Data sharing agreement with CCR or VR

Passive follow-back (data linkages)

- Regional Vital Records
- National Death Index (US only)
- Social Security Death Index (US only)
- Canadian National Death Clearance (Canada only)

NAACCR

Data Quality Checks and Clean-Up

- Data linkage quality control
  - Manual review
  - NDI SAS utility program
  - Other data linkages
    - Voter registration
    - Health insurance data
    - Hospital discharge data
    - Government offices (motor vehicle, public safety, taxes, etc.)
- Edits
  - NAACR/SEER/NPCR edit set
  - Survival-specific edit set

NAACCR

Coding Rules

- ICD coded diagnoses and COD are preferable
- Different jurisdictions sometime use slightly different rules for coding multiple primary cancers
- Prior to analysis and quality assessment, registry data can be transformed to a common rule structure for consistency (e.g., the International Agency for Research in Cancer (IARC) rules)

NAACCR

Death Related Information

- Updated vital status
- Date of death (or date last seen)
  - Complete dates are preferable (MDY)
- Accurate and complete COD information
  - Non-missing COD preferable
  - ICD coded COD preferable
  - Primary & underlying COD information preferable

NAACCR

Incomplete Date Information

- If complete dates are not available, imputation solutions can be used to produce an estimated survival time
- Example: C-SPAN mean imputation method
  - Used in the case of missing month or day of death (or diagnosis)
  - A SAS algorithm written by Larry Ellison at Statistics Canada returns an imputed a mean survival time

NAACCR

Incomplete Dates: Mean Survival Imputation

- An exact interval SAS macro with the imputation algorithm is available at: http://www.cancerview.ca/idc/groups/public/documents/webcontent/cspan_intervalmacro.sas
- The imputed value is a function of all potential values and the likelihood of their occurrence
- If either the diagnosis year or the death year is unknown then the survival is undefined
- If the month is missing from a date value then the day is also assumed to be missing

NAACCR

Incomplete Dates: Mean Survival Imputation

- Example 1: If only the **day of death is missing**:
  - If diagnosis and death month and year are the same
    - Imputed survival time is equal to half of the time between the date of diagnosis and the last day of the month of death
  - If diagnosis and death month and/or year are different
    - Imputed survival time is equal to the middle of the month of death (the 15th or 16th, depending on the month) minus the date of diagnosis

NAACCR

Incomplete Dates: Mean Survival Imputation

- Example 2: If the **month and day of death are missing**:
  - If diagnosis and death year are the same
    - Imputed survival time is equal to half of the time between the date of diagnosis and the last day of the year of death (December 31st)
  - If diagnosis and death year are different
    - Imputed survival time is equal to the middle of the year of death (July 2nd) minus the date of diagnosis

NAACCR

Final Data Quality for Survival Analysis

- Final data quality must be specially appraised before survival is calculated using protocols designed to highlight potential areas of error or bias
- To provide a picture of data quality in the survival context, make an inventory of ineligible, eligible and excluded records

NAACCR

Ineligible Records

- Following international protocols, criteria for ineligibility may include:
  - Basal and squamous cell skin cancers
  - Adolescent bone cancers
  - In situ cancers (with the exception of in situ bladder)
  - Tumours of benign or uncertain behaviour

NAACCR

Excluded Records

Following international protocols, criteria for exclusion may include:

- Age (<15 and >99 years at diagnosis)
- Unknown vital status
- Unknown sex
- Sex-site incompatibility
- Unknown year of birth, diagnosis or death
- Invalid sequences of dates
- Records where the diagnosis method was autopsy and the survival time was zero
- Records where the diagnosis method was death certificate only (DCO)
- In the case of first primary tumour analyses, second or subsequent tumours

NAACCR

Quality Assessment of Included Records

- To ensure completeness of the included records, a data quality assessment might involve enumerating:
  - Microscopically confirmed records
  - Records with missing month or day of birth, diagnosis, or death
  - Records where the diagnosis method is autopsy but survival time is greater than zero
  - Records where survival time is zero but diagnosis method is not DCO or autopsy (considered a "true zero survival time")

NAACCR

Reporting Quality Information

- To provide a comprehensive picture of data quality for survival analysis, the following might be provided by jurisdiction, site, diagnosis period and sex, where applicable:
  - Percentage of ineligible and excluded records
  - A description of completeness of the records retained in survival analyses after exclusions
  - Percentage of all primary records included in survival analyses
  - Percentage of patients where the attained age of the patient was ≥ 100 at the end of the study period

NAACCR

The Cancer Survival and Prevalence Analytic Network (C-SPAN) Experience

- Primary data source: The *Canadian Cancer Registry* (CCR), housed at Statistics Canada
  - A collaboration among Canadian provincial and territorial cancer registries and Statistics Canada
  - Regular data quality edits, de-duplication and death clearance at a national level augment local level efforts
- Funding provided by: The *Canadian Partnership Against Cancer*

NAACCR

Data Quality Results

- Overall, potential quality threats were minimal as measured by the quality protocol for survival analysis just presented
- C-SPAN's rates of DCOs, missing demographic or date information, and microscopic confirmation are remarkably similar to those arising from international studies that have set high quality data standards

NAACCR

Inter-Provincial Differences

- Quality considerations highlighted that higher-level system (inter-provincial) differences must also be considered
- Consistent with previous analyses, Quebec's data were excluded from analyses due to differences in cancer registration practices and issues in determining vital status for Quebec cases in the CCR
- Until recently, the Newfoundland and Labrador (NL) Cancer Registry did not receive information on all death certificates that mentioned cancer
  - Since the situation was recently resolved, NL data were included in analyses and interpreted with caution (consistent with national protocol)

NAACCR

Data Quality Results

- There were 1,600,722 cancer records registered between 1992 and 2006, representing 1,565,425 cancer patients
- 6.1% of the registered records were ineligible, mostly in situ neoplasms, reflecting variations in registry practices - some provinces do not register non-invasive tumours
- Only 2.6% of all eligible patients were excluded and inclusion rates by site were high:

Lung	Colorectal	Breast	Prostate
96.6%	98.2%	99.3%	99.1%

NAACCR

Data Quality Results

- 88.7% of included records were microscopically confirmed. Rates varied by site:

Lung	Colorectal	Breast	Prostate
96.6%	98.2%	99.3%	99.1%

- Most other indicators of potential quality issues showed low rates of occurrence, particularly for missing or questionable death date-related information

NAACCR

A Need to Protect Confidentiality

- To reduce disclosure risk and maintain confidentiality:
  - Percentages were categorized in most cases
  - Any cell count less than 6 and greater than 0 was suppressed
  - One area of particular disclosure risk involved the Territories
    - Required heavy suppression due to small numbers
    - These regional data are only presented in select cases

NAACCR

Best Practices for Developing and Working with Survival Data:

*Using SEER*Prep and SEER*Stat to calculate survival statistics.*

Chris Johnson, MPH
Epidemiologist
Cancer Data Registry of Idaho

NAACCR

Outline

- The presentation will follow from a NAACCR V12 layout through using SEER*Prep to create a SEER*Stat database, then the calculation of survival statistics in SEER*Stat.
- Brief overview of what needs to be done to prepare data for use in SEER*Prep and SEER*Stat.
- Examples of calculations of more commonly used survival statistics, i.e., observed, relative.

NAACCR

101

What is SEER*Stat?

- SEER*Stat is a statistical package created for the analysis of SEER and other cancer databases.
- It was developed by Information Management Services, Inc. in consultation with the SEER Program of the National Cancer Institute (NCI).
- The SEER*Stat statistical software provides a convenient, intuitive mechanism for the analysis of SEER and other cancer-related databases.
- It is a powerful PC tool to view individual cancer records and to produce statistics for studying the impact of cancer on a population.

102 NAACCR

What is SEER*Prep?

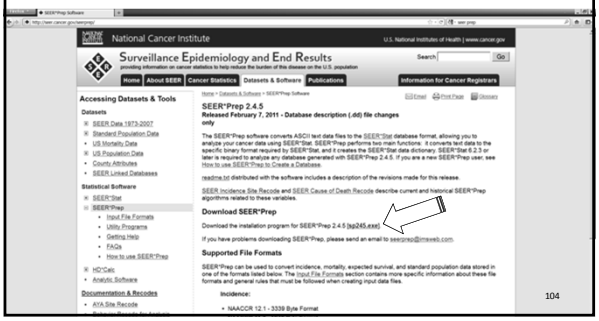
- SEER*Prep software converts ASCII text data files to the SEER*Stat database format, allowing you to analyze your cancer data using SEER*Stat.
- SEER*Prep performs two main functions:
  - it converts text data to the specific binary format required by SEER*Stat,
  - and it creates the SEER*Stat data dictionary.

NAACCR

103

How to obtain SEER*Prep software

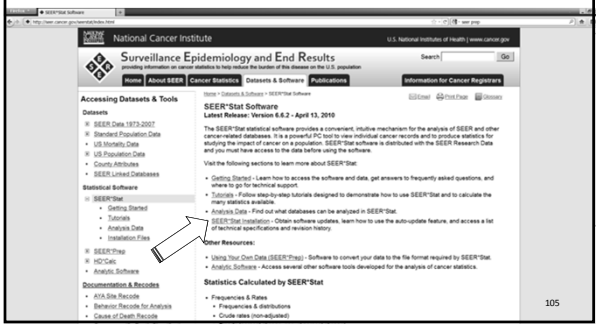
- http://seer.cancer.gov/seerprep/



104

How to obtain SEER*Stat software

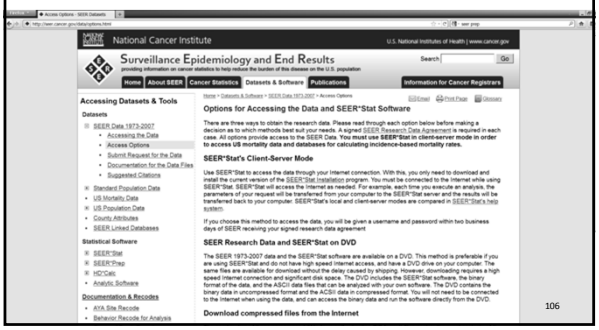
- http://seer.cancer.gov/seerstat/



105

How to access the SEER Research Data.

- http://seer.cancer.gov/seerstat/



Create a NAACCR V12 Incidence file

1. Query your database for state/provincial residents diagnosed over the range of years you have completed death clearance/follow-up activities/(NDI linkages for U.S.).

- e.g.

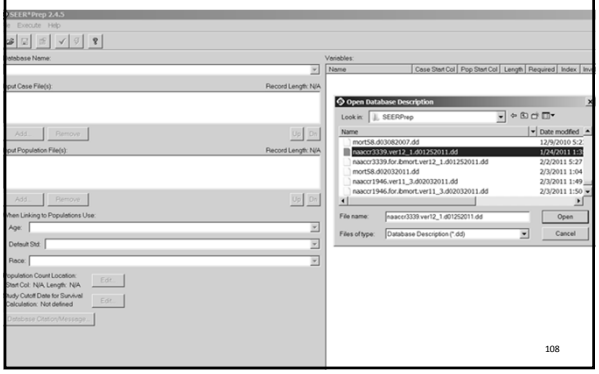
NAACCR Item	Min	Max
390	bbbb1970	99992008
80	ID	ID
others?		

2. Sort the query result by NAACCR Item 20 (Patient ID Number) and NAACCR Item 380 (Sequence Number—Central).
3. Export a NAACCR V12 Incidence (3339 column width) file for the queried cases with a .txd file extension.

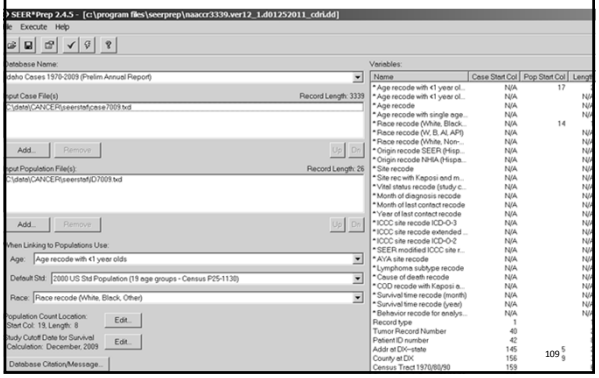
NAACCR

107

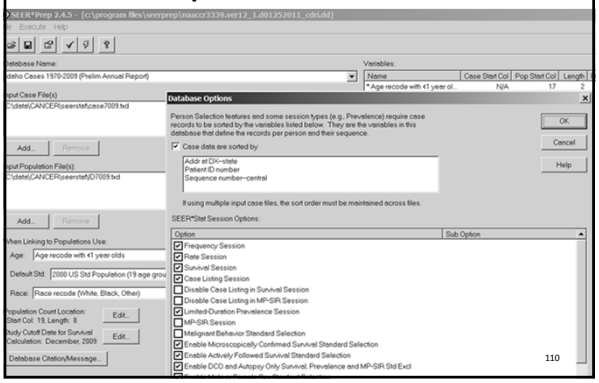
Use SEER*Prep to create a SEER*Stat dataset



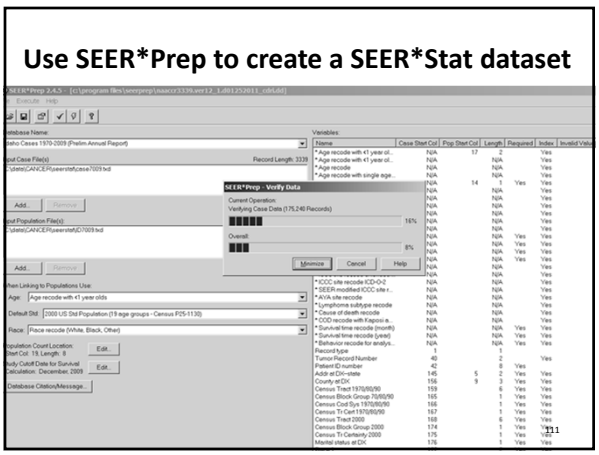
Use SEER*Prep to create a SEER*Stat dataset



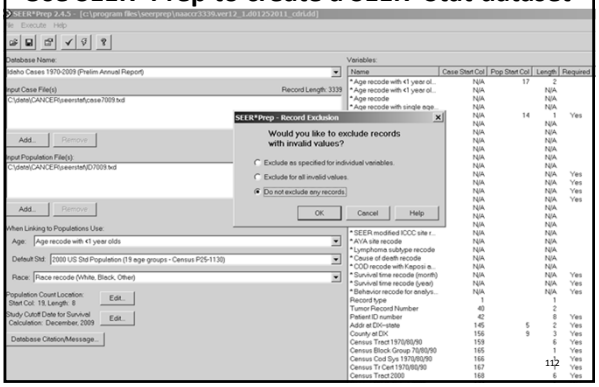
Use SEER*Prep to create a SEER*Stat dataset



Use SEER*Prep to create a SEER*Stat dataset

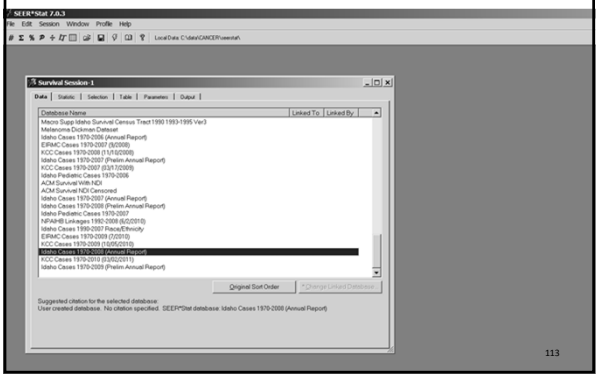


Use SEER*Prep to create a SEER*Stat dataset



112

Using SEER*Stat for survival analysis



113

Population-based Cancer Survival Statistics Overview

- Cancer survival is the proportion of patients alive at some point subsequent to the diagnosis of their cancer, or from some point post-diagnosis (conditional survival).
- It is represented as the probability of a group of patients "surviving" a specified amount of time (e.g. 3 years, 5 years, 20 years).
- (Source: NCI http://surveillance.cancer.gov/survival/)

NAACCR

114

Types of survival statistics available in SEER*Stat

- Observed Survival
  - Estimate of the probability of surviving all causes of death.
- Net Survival
  - (policy-based statistic) - The probability of surviving cancer in the absence of other causes of death. It is a measure that is not influenced by changes in mortality from other causes and, therefore, provides a useful measure for tracking survival across time, and comparisons between racial/ethnic groups or between registries.
- Conditional Survival
  - Given survival to some number of years, what is the probability of surviving some additional number of years.
- Crude Probability of Death
  - (patient prognosis measure) - The probability of dying of cancer in the presence of other causes of death.
- Survival Case Listing

NAACCR

115

Approaches to estimation of cancer-specific survival

- There are two ways to estimate Net Cancer-Specific Survival:
  - using cause of death information
  - or using expected survival tables.

NAACCR

116

Net cancer-specific survival

- Cause-specific survival
  - Estimates are calculated by specifying the cause of death. Individuals who die of causes other than those specified are considered to be censored.
- Relative survival
  - Uses population life tables to estimate expected survival. Relative survival is defined as the **ratio** of the proportion of observed survivors (all causes of death) in a cohort of cancer patients to the proportion of expected survivors in a comparable cohort of cancer-free individuals.
    - Assumes independent competing causes of death. Since a cohort of cancer-free individuals is difficult to obtain, we use expected life tables and assume that the cancer deaths are a negligible proportion of all deaths.

NAACCR

117

Overview of SEER*Stat

- SEER*Stat allows you a great deal of freedom to request the cancer statistics/values/methods you want for your analysis.
- Part 1: Session
- Part 2: Execute
- Part 3: Matrix

NAACCR

118

Overview of SEER*Stat

- Part 1: Session
- The analysis is set up in the session window. Each session consists of tabs on which you select the database subset, statistics, and appearance of your output matrix.
- You should work through each tab in order from left to right and from top to bottom to ensure that all options have been considered.
  - However, changes can be made in any order.
  - It is possible to work on multiple sessions simultaneously.

NAACCR

119

Overview of SEER*Stat

- Part 2: Execute
- Once the session is set up, you are ready to execute it as a job.
- While the job is executing, you can change the session or begin a new one without affecting the original job.
- It is possible to execute more than one job at a time.

NAACCR

120

Overview of SEER*Stat

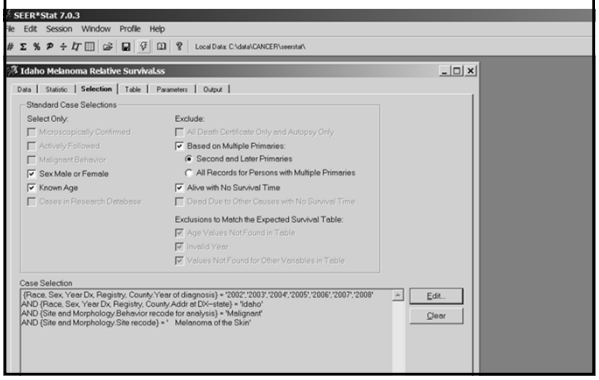
- Part 3: Matrix

- When the job has finished executing, the output matrix you requested is displayed.
- You can change the appearance of the output matrix, print it, copy it to the Windows clipboard, and/or export the statistics/values so they may be used in another application.

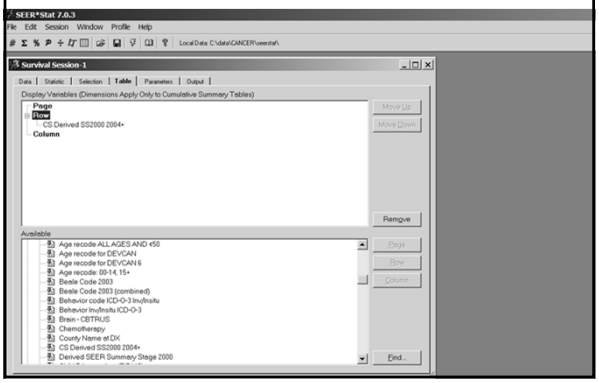
NAACCR

121

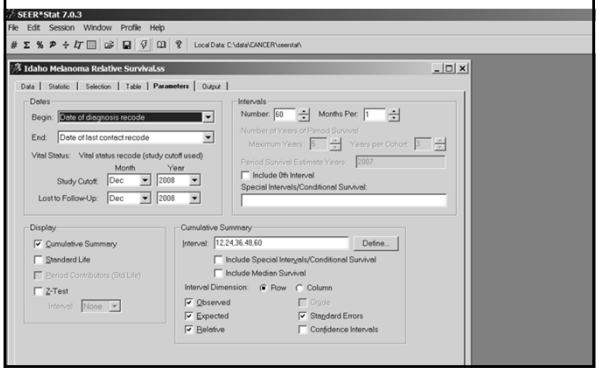
Dataset used for SEER*Stat examples



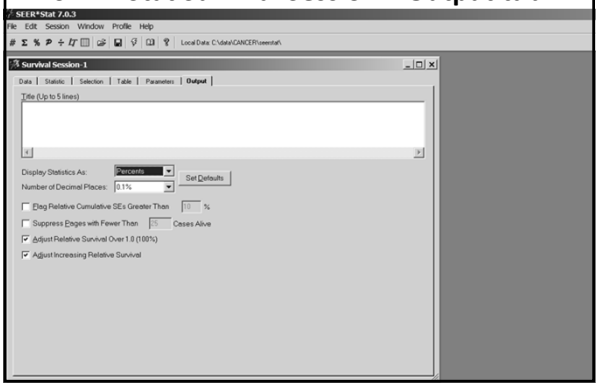
SEER*Stat Survival Session – Table tab



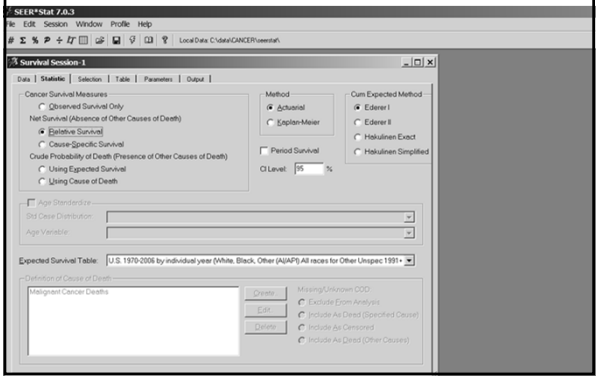
SEER*Stat Survival Session – Parameters tab



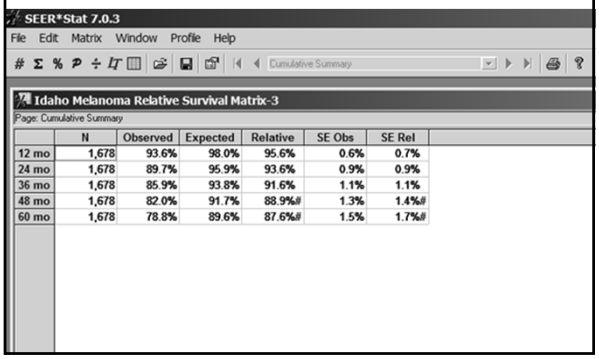
SEER*Stat Survival Session – Output tab



SEER*Stat Survival Session – Statistics tab

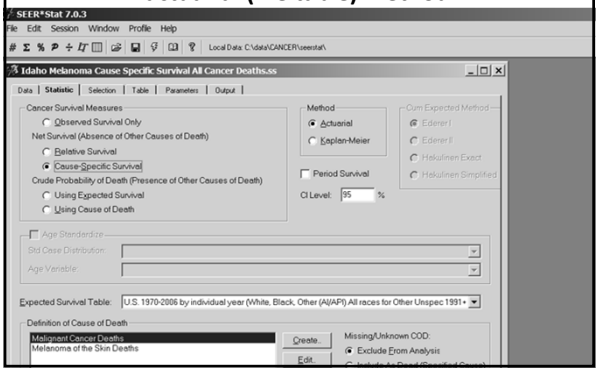


Example 1: Observed and Relative Survival using the actuarial (life table) method

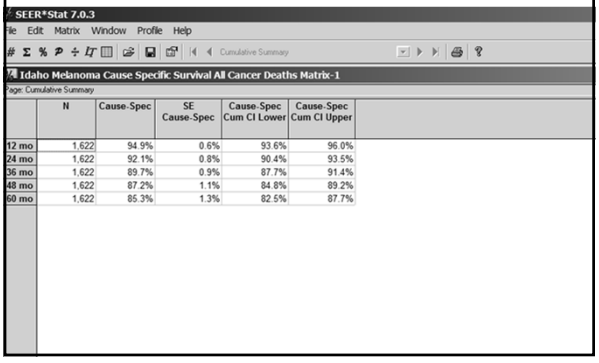


	N	Observed	Expected	Relative	SE Obs	SE Rel
12 mo	1,678	93.6%	98.0%	95.6%	0.6%	0.7%
24 mo	1,678	89.7%	95.9%	93.6%	0.9%	0.9%
36 mo	1,678	85.9%	93.8%	91.6%	1.1%	1.1%
48 mo	1,678	82.0%	91.7%	88.9%#	1.3%	1.4%#
60 mo	1,678	78.8%	89.6%	87.6%#	1.5%	1.7%#

Example 2: Cause-specific survival using the actuarial (life table) method

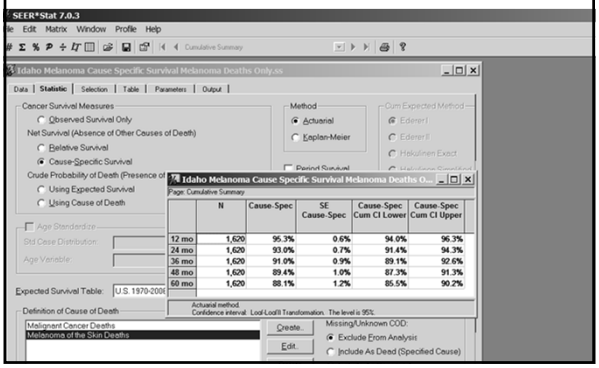


Example 2: Cause-specific survival using the actuarial (life table) method



	N	Cause-Spec	SE Cause-Spec	Cause-Spec Cum CI Lower	Cause-Spec Cum CI Upper
12 mo	1,622	94.9%	0.6%	93.6%	96.0%
24 mo	1,622	92.1%	0.8%	90.4%	93.5%
36 mo	1,622	89.7%	0.9%	87.7%	91.4%
48 mo	1,622	87.2%	1.1%	84.8%	89.2%
60 mo	1,622	85.3%	1.3%	82.5%	87.7%

Example 3: Cause-specific survival using the actuarial method, melanoma of the skin deaths



Survival Proportion Calculations

- Five-year survival is calculated as the product of the conditional probabilities for surviving each single year interval.

$S_{5yr} = S_{1yr} * S_{2yr|1yr} * S_{3yr|2yr} * S_{4yr|3yr} * S_{5yr|4yr}$

NAACCR

131

Cohort versus Period survival

- Survival estimates from cancer registry data are usually dated measures of current-year survival, because of the time needed to observe survival and lag between available data and the current year.
- There are different approaches of grouping survival experience with respect to year of diagnosis and follow-up to obtain more up-to-date estimates of patients recently diagnosed.

Observed Survival by Year of Follow-up and Year of Diagnosis
Regional Female Breast Cancer, 1997-2001, SEER 9 Registries

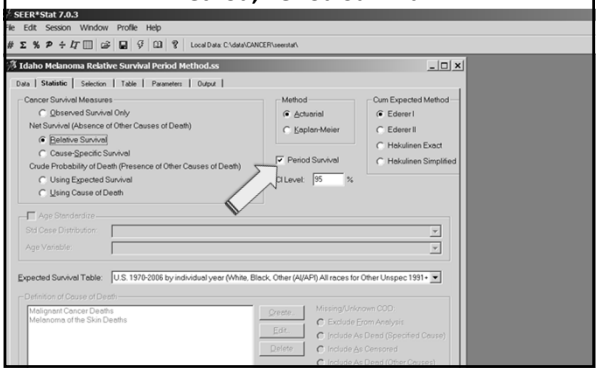
	1997	1998	1999	2000	2001
1	96.4%	96.5%	96.3%	97.0%	96.8%
2	93.8%	93.8%	94.5%	94.4%	
3	93.8%	94.3%	94.7%		
4	93.6%	94.0%			
5	94.1%				

Cohort 74.7 %
Mutiple Year 75.2 %
Period 76.6 %

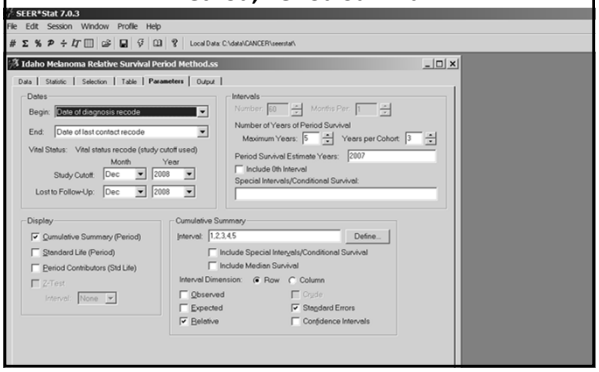
NAACCR

132

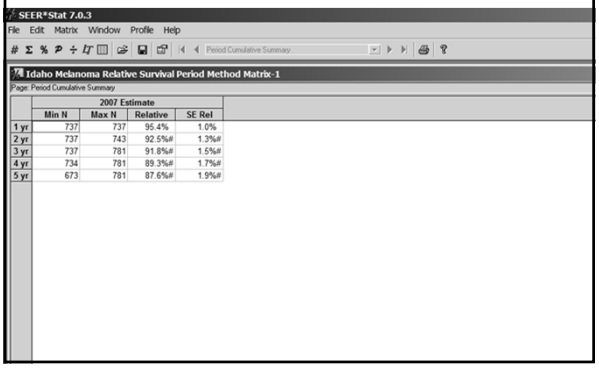
Example 4: Relative survival using the actuarial method; Period Survival



Example 4: Relative survival using the actuarial method; Period Survival

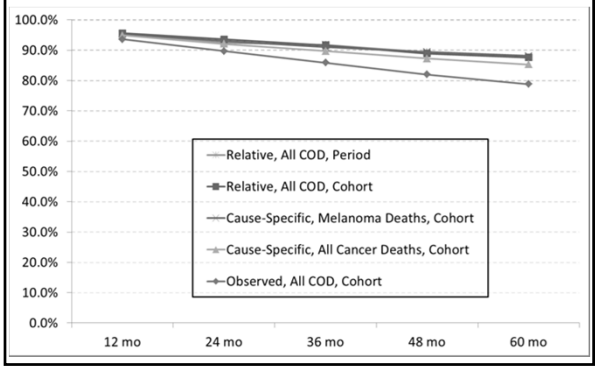


Example 4: Relative survival using the actuarial method; Period Survival



	2007 Estimate			
	Min N	Max N	Relative	SE Rel
1 yr	737	737	95.4%	1.0%
2 yr	737	743	92.5%#	1.3%#
3 yr	737	781	91.8%#	1.5%#
4 yr	734	781	89.3%#	1.7%#
5 yr	673	781	87.6%#	1.9%#

Comparison of common population-based survival methods



Conclusions: SEER*Stat

- Advantages of SEER*Stat over other statistical tools:
- Simple to use GUI
- Facilitates comparisons with SEER data
- Can paste results into other Windows programs
- SEER/NCI is responsible for keeping it updated and standardized
- Well supported by IMS

NAACCR

137

Summary
Using SEER*Prep and SEER*Stat to calculate survival statistics.

- The presentation followed a NAACCR V12 layout through SEER*Prep to create a SEER*Stat database, then demonstrated the calculation of survival statistics in SEER*Stat.
- Brief overview of what needs to be done to prepare data for use in SEER*Prep and SEER*Stat.
- Examples of calculations of more commonly used survival statistics, i.e., observed, relative.

NAACCR

138

Some issues related to survival

....

Hannah Weir, PhD
Trevor Thompson, BS
Division of Cancer Prevention and Control
Centers for Disease Prevention and Control

The findings and conclusions in this presentation are those of the presenters and do not necessarily represent the official position of the Centers for Disease Control and Prevention.

NAACCR

LIFE TABLES

NAACCR

140

General mortality varies by area

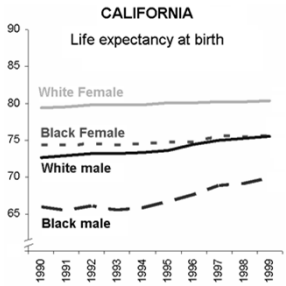
Life expectancy at birth - all races, 1990-1999

Cancer Registry areas	e_0 Male	e_0 Female
Hawaii	76.11	82.19
Utah	75.78	80.85
Iowa	74.75	80.84
Connecticut	74.56	80.46
California	73.94	79.98
Wyoming	73.83	79.62
New Mexico	73.31	79.98
USA	**72.76**	**79.09**

NAACCR

141

General mortality varies by calendar year (principally in male population)



NAACCR

142

CONCORD Study - Relative Survival using two LTs

Characteristic	Patient example	NCHS LT*	CONCORD LT
SEX	Male	Male	Male
RACE	Black	Black	Black
YEAR	1996	1990	1996
AREA	Utah	US	Utah

NAACCR

* US Census 1990

143

Life expectancy at birth in 1990 – all races

	Male		Female	
	CONCORD	NCHS	CONCORD	NCHS
California	72.7	72.5	79.4	79.2
Connecticut	74.0	73.6	80.2	80.0
Hawaii	75.6	75.4	81.7	81.3
Iowa	74.2	73.9	80.9	80.5
New Mexico	72.6	72.2	79.6	79.3
Utah	75.2	75.0	81.0	80.4
Wyoming	73.3	73.2	79.4	79.3

NAACCR

144

What do we expect in relative survival?

WE KNOW THAT:
- General mortality varies in the period (1990-1999) principally in male population
- General mortality varies by geographical area
- Hawaii is the area with major differences in comparison with the USA (also in female population)

Using CONCORD life tables versus US Census (NCHS) life tables in relative survival estimates WE EXPECT THAT:
- Major differences will be present in male cancer sites
- Hawaii cancer relative survival estimates will have major differences

NAACCR

145

5-yr crude relative survival
Male colorectal cancer - all races

Areas	# Cases	NCHS LT (1)	CONCORD LT (2)	Difference (2) – (1)
California	30,379	63.0%	61.1%	- 1.9%
Connecticut	4,559	63.6%	61.5%	- 2.1%
Hawaii	1,493	69.6%	65.7%	- 3.9%
Iowa	4,043	61.6%	60.1%	- 1.5%
New Mexico	1,335	60.9%	59.0%	- 1.9%
Utah	1,258	64.5%	61.4%	- 3.1%
Wyoming	357	57.7%	56.7%	- 1.0%

NAACCR

146

5-yr crude relative survival
Female colorectal cancer- all races

Areas	# Cases	NCHS LT (1)	CONCORD LT (2)	Difference (2) – (1)
California	29,204	61.7%	61.0%	- 0.7%
Connecticut	4,406	63.0%	61.6%	- 1.4%
Hawaii	1,089	68.5%	66.2%	- 2.3%
Iowa	4,519	66.1%	64.4%	- 1.7%
New Mexico	1,214	61.9%	60.7%	- 1.2%
Utah	1,096	60.5%	59.6%	- 0.9%
Wyoming	391	59.1%	58.6%	- 0.5%

NAACCR

147

5-yr crude relative survival
Female breast cancer- all races

Areas	# Cases	NCHS LT (1)	CONCORD LT (2)	Difference (2) – (1)
California	82,868	86.3%	85.8%	- 0.5%
Connecticut	11,288	86.2%	85.1%	- 1.1%
Hawaii	2,854	91.2%	89.5%	- 1.7%
Iowa	9,131	87.8%	86.3%	- 1.5%
New Mexico	3,793	85.5%	84.6%	- 0.9%
Utah	3,505	86.3%	85.3%	- 1.0%
Wyoming	1,073	84.3%	83.9%	- 0.4%

NAACCR

148

5-yr crude relative survival
Male prostate cancer - all races

Areas	# Cases	NCHS LT (1)	CONCORD LT (2)	Difference (2) – (1)
California	91,613	96.9%	93.5%	- 3.4%
Connecticut	11,306	96.4%	92.8%	- 3.6%
Hawaii	3,480	99.9%	94.1%	- 5.8%
Iowa	10,742	95.2%	92.8%	- 2.4%
New Mexico	5,389	96.6%	93.2%	- 3.4%
Utah	5,777	99.2%	94.2%	- 5.0%
Wyoming	1,551	95.9%	93.9%	- 2.0%

NAACCR

149

Recent updates to SEER*Stat (V 7.0.4)

US LT available for individual years 1970-2006
by gender and race (All, W, B and O)

NAACCR

Cancer cause specific survival - an alternatives to relative survival when life tables not availalbe

LT matched to cancer patients according to risk factors (age, calendar period, geographic area and race/ethnicity)
- SES, smoking status, etc.

RS can underestimate or overestimate the actual survival experience when there is a mismatch between the LT and cancer patient cohort (e.g., tobacco related cancers)

Howlader et al, 2010 published broader definition of caused related death variable.

NAACCR

Age Standardized Survival Estimates

NAACCR

152

Age-Standardized Survival Estimates

- Survival generally depends on age
- Age distribution among cancer patients may vary across comparison groups
- Standardization is needed to remove the confounding effect of age when comparing survival estimates
- Which standard population should be used?

NAACCR

Commonly Used Standards

- Internal site-specific age distribution of a study
  - Derived from observed age distribution of a specific cancer patient population
- International Cancer Survival Standards (ICSS) standard populations
  - Set of general standard cancer patient populations developed from the EUROCARE-2 study

NAACCR

ICSS Standard Populations

- Consists of three standard populations describing the main age patterns of cancer incidence
  1. Increasing with age (91.1% of EUROCARE-2 patients)
  2. Generally constant with age (7.4%)
     - Nasopharynx, soft tissues, melanoma, cervix uteri, brain, thyroid, bone
  3. Primarily affecting young adults (1.5%)
     - Testis, Hodgkin's disease, acute lymphatic leukemia

NAACCR

Which Standard to Use?

- Site-specific
  - Has the desirable property that age-standardized survival estimates are generally close to the crude survival estimates
  - Does not allow comparisons across sites
  - Does not allow comparisons across studies if internal standards are used

NAACCR

Which Standard to Use?

- ICSS
  - Standardized survival estimates can differ from crude results
  - Allows for comparisons with other sites that use the same standard
  - Allows for comparisons with other studies using ICSS weights

NAACCR

Example – Comparison of Standards

Table 1. 5-Year Relative Survival Estimates by Cancer Site, SEER 1998-2002

Site	N	Crude Relative Survival (95% CI)	Age-Adjusted* **Site Specific** (95% CI)	Age-Adjusted† **ICCC** (95% CI)
Colorectal	118,451	64.9 (64.6-65.3)	64.3 (64.0-64.7)	64.9 (64.6-65.3)
Female Breast	172,662	89.3 (89.1-89.5)	89.4 (89.2-89.7)	89.7 (89.5-90.0)
Prostate	190,464	99.6 (99.4-99.8)	99.1 (98.9-99.3)	98.7 (98.5-98.9)
Thyroid	20,415	96.9 (96.5-97.3)	95.8 (95.4-96.2)	93.9 (93.3-94.6)
Melanoma	42,132	91.5 (91.1-91.9)	90.6 (90.0-91.1)	91.3 (90.8-91.7)
Cervix	12,923	71.8 (71.0-72.7)	69.8 (68.9-70.7)	65.5 (64.4-66.5)
Testis	7,851	95.5 (95.0-96.0)	95.2 (94.6-95.8)	90.4 (87.6-93.3)
Hodgkin Lymphoma	7,174	84.1 (83.1-85.1)	80.2 (79.2-81.3)	80.0 (79.0-81.1)

* Age-standardized to the site-specific age distribution of the 2004-2006 USCS.
† Age-standardized to the appropriate ICSS standard.

NAACCR

Software Considerations

- SEER*STAT currently does not perform age-standardization of relative survival estimates
  - This may be included in future versions
  - Age-specific relative survival estimates can be calculated in SEER*STAT and exported to other packages for standardization
- R and Stata code are available for calculating age-standardized survival estimates and confidence intervals

NAACCR

The "multiple primaries issue"

- One person can have many cancers.
  - Becoming more common scenario as survivorship from cancer increases.
- Multiple primary rules differ (IARC vs. SEER/Canadian)
- Survival statistics have traditionally focused on "first primary"… but this doesn't include all the information available.
- EUROCARE now including ALL cancers diagnosed.

NAACCR

Reference

- Baili P, Micheli A, De Angelis R, *Weir HK*, Francisci S, Santaquilani M, Hakulinen T, Quaresma M, Coleman MP, CONCORD Working Group. **Life tables for world-wide comparison of relative survival for cancer (CONCORD study).** Tumori 2008 94; 658-668.
- Howlader N, Ries LA, Mariotto AB, Reichman ME, Ruhl J, Cronin KA. **Improved estimates of cancer-specific survival rates from population-based data.** J Natl Cancer Inst. 2010 Oct 20;102(20):1584-98.
- Corazziari I, Quinn MJ, Capocaccia R**. Standard cancer patient population for age standardizing survival ratios.** *Eur J Cancer 2004;* 40: 2307–16.
- Brenner H, Hakulinen T. **Patients with previous cancer should not be excluded in international comparative cancer survival studies.** Int J Cancer 2007;121:2274–8
- Rosso S, De Angelis R, Ciccolallo L, Carrani E, Soerjomataram I, Grande E, Zigon G, Brenner H, and the EUROCARE Working Group. **Multiple tumours in survival estimates**. Eur J Cancer 2009,45:1080-1094
- Ellison LF. **Measuring the effect of including multiple cancers in survival analyses using data from the Canadian Cancer Registry.** Cancer Epidemiol. 2010 Oct;34(5):550-5.
- Micheli A, Baili P, Mugno E, Queen M, Capocaccia R, Grosclaude PC, EUROCARE Working Group. **Life expectancy and cancer survival in the EUROCARE-3 cancer registry areas**. Ann Oncol 2003;14(Suppl 5): 28–40.

NAACCR

Conclusion

NAACCR

Overview and Use of Population-based Survival Data

- Population-based survival monitors the effectiveness of health care delivery - cancer control and health policy
- Adding survival data enhances the registry
- Canada and NCI/SEER routinely produce these data
- NPCR is now expanding capacity:
  - Currently 26 NCPR/SEER registries link to NDI (62% population coverage)
  - 83% coverage with additional 10 "interested" registries
  - 14 registries ???
- NAACCR Survival Workgroup is addressing issues related to the collection, analysis and interpretation of survival data
- 2011 NAACCR conference – plenary address by Prof Michel Coleman, PI CONCORD (-2) Study

NAACCR

Need for Death Clearance

- Routine DC helps with case ascertainment and provides information on vital status (~97% of deaths MI)

NAACCR

NDI and impact on survival data

- Supplements DC to provide info on ~99% deaths (L Alom)
  - Deaths out of state /residence out of state
  - Identifies duplicate cases (NY-FL dual residences issue, etc.)
- Most deaths found through DC but NDI still critical
- NPCR-NDI umbrella application
- NDI linkage at no additional cost to NPCR/SEER registries
- Tools available to help with NDI output

NAACCR

Data Quality Issues

- Complete case ascertainment and death ascertainment very important*
- Data quality indicators
  - Confidentiality issues related to complete date variables
    - Age is needed for LT
    - Survival interval immediately following diagnosis - impact on long term survival and measures of excess mortality related to treatment
  - Errors vs. non errors
    - Patients with "zero" survival time ???

* Johnson CJ, Weir HK, Yin D, Niu X. Assessment of the impact of variation in patient follow-up on survival statistics using synthetic datasets based on SEER data. JRM 2010: 37(3):96-103.

NAACCR

Tools for calculating survival statistics

- SEER*Prep and SEER*Stat are powerful tools, freely available from NCI/IMS for use in calculating survival data.
- Other stat programs are available for more complex analyses (websites listed)
- Tools are there but some of the supporting data may be lacking
  - In US, availability of State and race/ethnic specific LT is limited
  - Cause of death for cause specific survival is not consistent

NAACCR

Work Remains!!!!

- Dual residence issue
- Multiple primary rules
- Data quality – incomplete vs. suppressed or tweaked data
- Quality of cause of death information on DC
- Availability of supporting information (LT)
- Choice of standard

NAACCR

Eventually we will get to here…..

Nationwide coverage of high quality and complete population-based cancer survival data available for cancer control, health policy and research use.

NAACCR

Questions?

Complete Case Identification and Ascertainment 7/7/11

Joyce Jones

CoC trained Independent Cancer Program Consultant

NAACCR

Coming up…

- July 7, 2011
  - Complete Case Identification and Ascertainment
  - Presented by Joyce Jones
    - *CoC trained Independent Cancer Program Consultant*
- August 4, 2011
  - NAACCR Interoperability Activities and the Electronic Health Record
  - Presented by NAACCR Path Data Workgroup

NAACCR

2011-2012 NAACCR Webinar Series

- Registration is open!
  - http://www.naaccr.org/EducationandTraining/WebinarSeries.aspx

NAACCR