Evaluation of NAACCR Survival Data

June 2012

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

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

And the NAACCR Survival Analysis Workgroup (SAWG)

NAACCR

NAACCR Survival Analysis Workgroup Members

Name	State, Province or Agency
Deb Hurley	SC (co-chair)
Chris Johnson	ID (co-chair)
Glenn Copeland	MI
Larry Ellison	Stat Cam
Monique N. Hernandez, Ph.D.	FL
Bin Huang	KY
Angela Mariotto	NCI
Zoran Miladinovic	Stat Can
Cyllene Morris	CA
Xiaoling Niu	NJ
Arti Parikh-Patel	CA
Paulo S. Pinheiro, MD PhD	NV
Trevor Thompson	CDC
Donna Turner	MB
Baozhen Qiao	NY
Zhenguo Qiu	AB
Kevin Ward	GA
Hannah Weir	CDC
Reda Wilson	CDC
Brad Wohler	FL
Kevin Zhang	MACRO

NAACCR

Overview

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

What is Population-Based Survival

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

NAACCR

Advantages and Disadvantage of Relative vs. Cause Specific Survival

	Advantages	Disadvantages
Relative	Relies on fact of death not cause of death	Life tables may not be available for all populations
Cause Specific	Not limited to populations with life tables	Death Certificates may not be reliable (e.g., may be coded to site of mets or recurrence)

NAACCR

Overview

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

NAACCR

Data

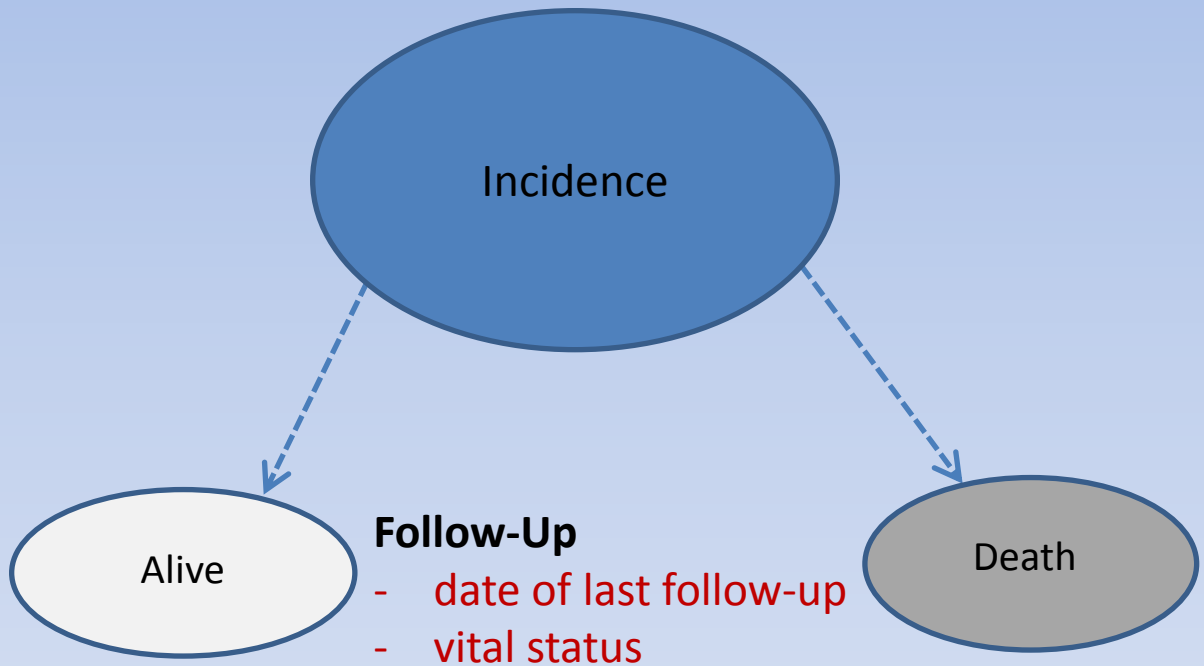
- CINA (1995-2008) 2010 data submission
- First year requested follow-up data
- Excluded Canadian data due to coding of vital status variable
- Registries
  - SEER: CA (LA, SF), Detroit, HI, IA, KY, LA, NJ, NM, UT, Seattle
  - NPCR: remaining states
    - 2 NPCR state cancer registries not included

NAACCR

Data Elements

- **Patient Demographics**
  - date of birth
  - sex
  - race/ethnicity
  - name
  - SS#
- **Tumor Record**
  - site
  - histology
  - behavior
  - stage
  - date of diagnosis
  - type of reporting source

-



Follow-Up

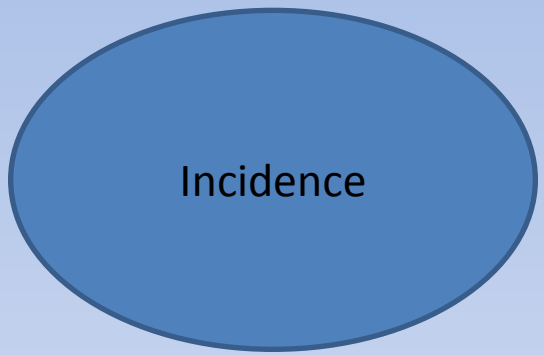
- date of last follow-up
- vital status
- cause of death
- follow-up source central

NAACCR

Evaluation Criteria

- % Sex, Age or Race Unknown
- % DCO/AO
- % Vital status Unknown
- % Edi Errors
- % MV
- % Missing Cause of Death
- % Alive with 0 Survival Time
- % Death within 1 Month of Diagnosis

The Foundation for Population-Based Survival



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

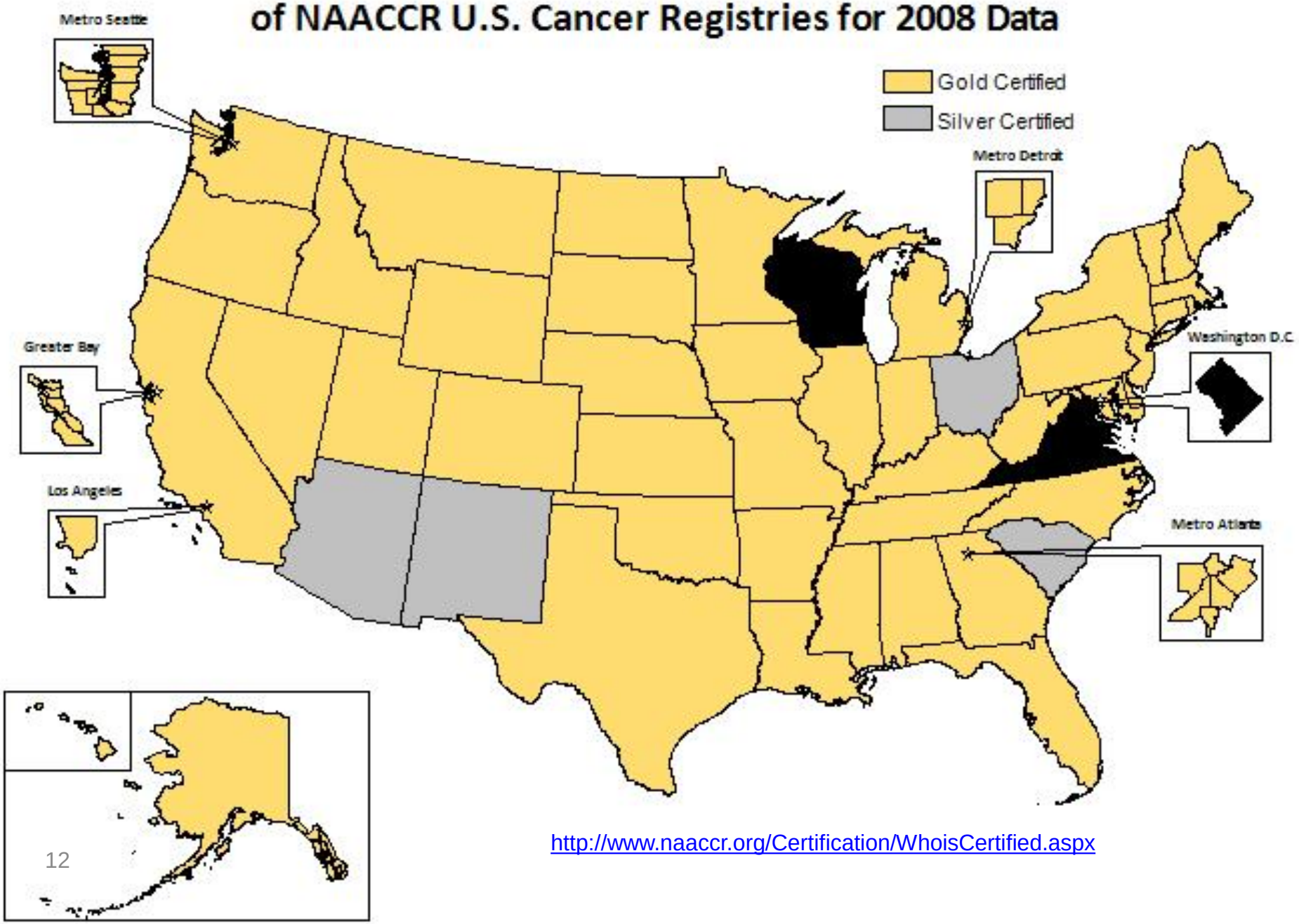
-

NAACCR

Factors that Impact Incidence

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

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



http://www.naaccr.org/Certification/WhoisCertified.aspx

Factors that Impact Incidence

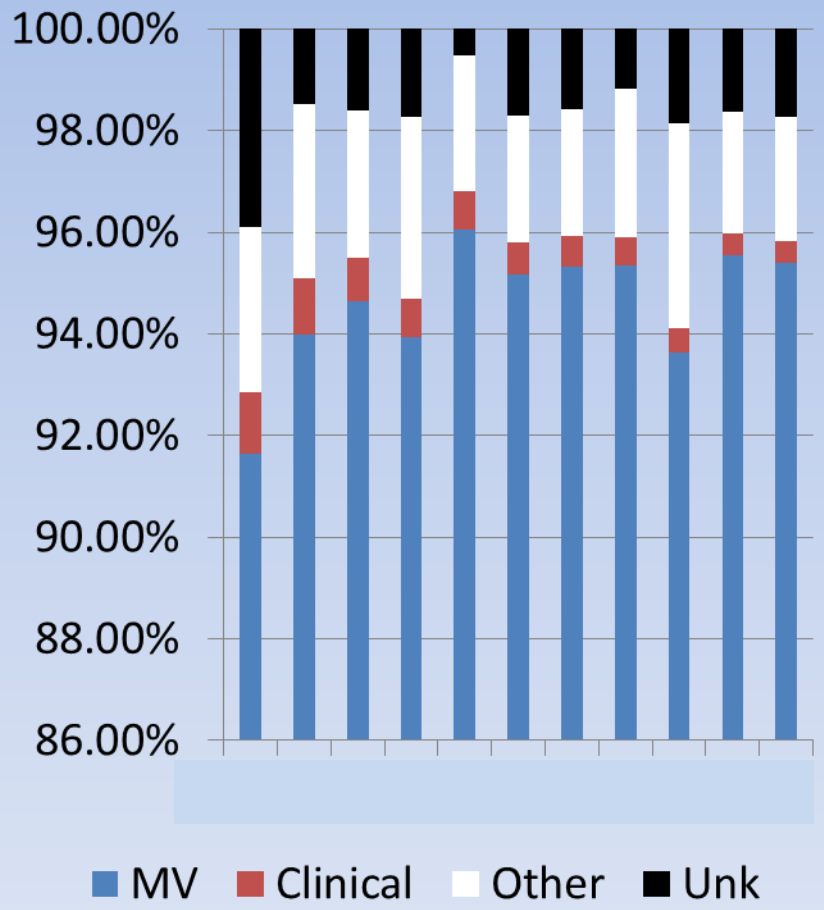
- NAACCR Certification
  - Completeness of case ascertainment
  - DCO/ autopsy
  - Missing critical information (age, sex, race)
  - Edits
  - Duplicates
- Population Coverage
  - 1995 - 19 US registries NAACCR Certified
  - 2008 - 53 US registries Certified

NAACCR

Factors that Impact Incidence

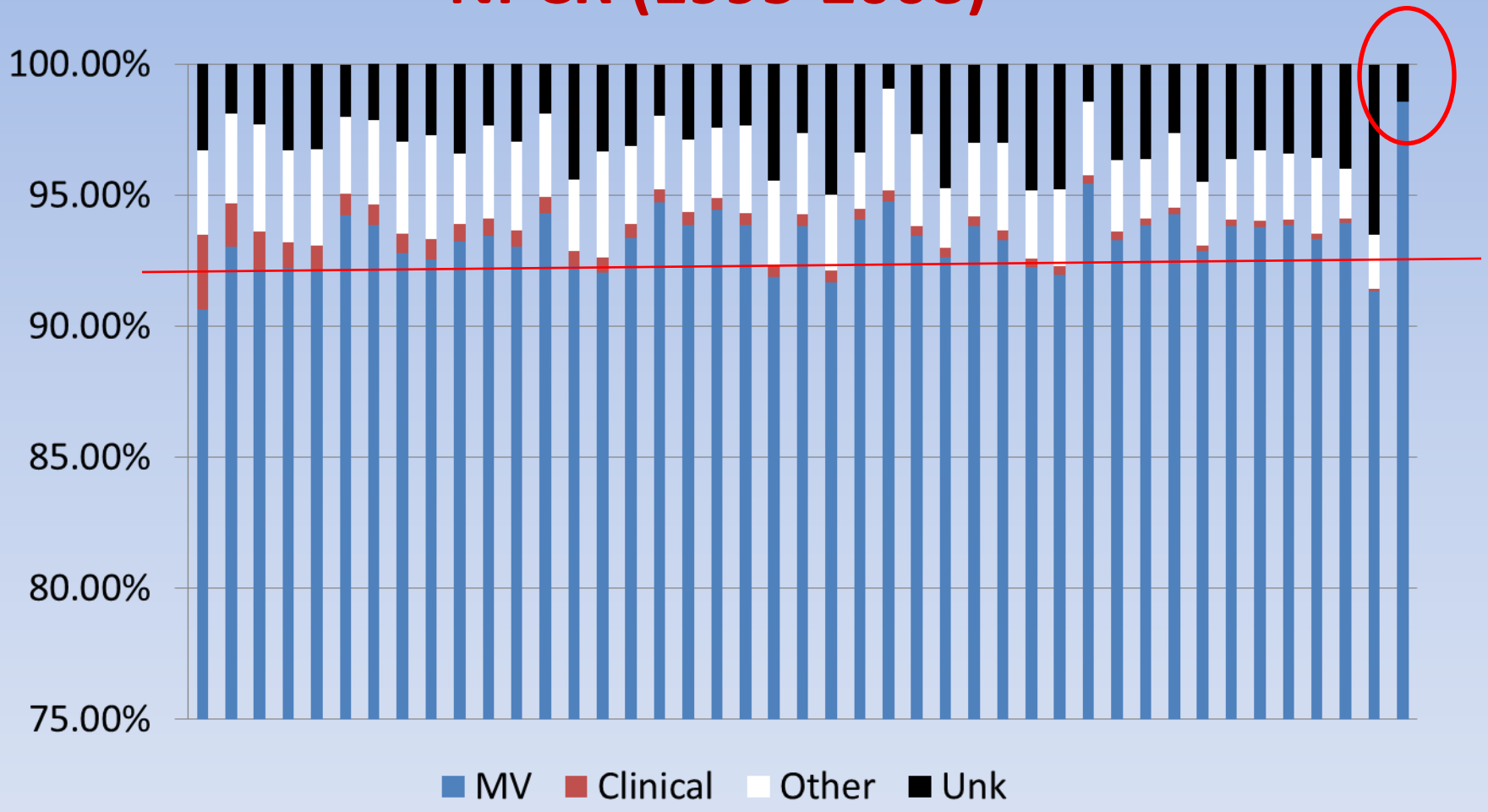
- NAACCR Certification
- Completeness of Case Ascertainment
  - Clinical vs. Microscopically Verified (%MV)

% Type Diagnostic Confirmation SEER (1992-2008)



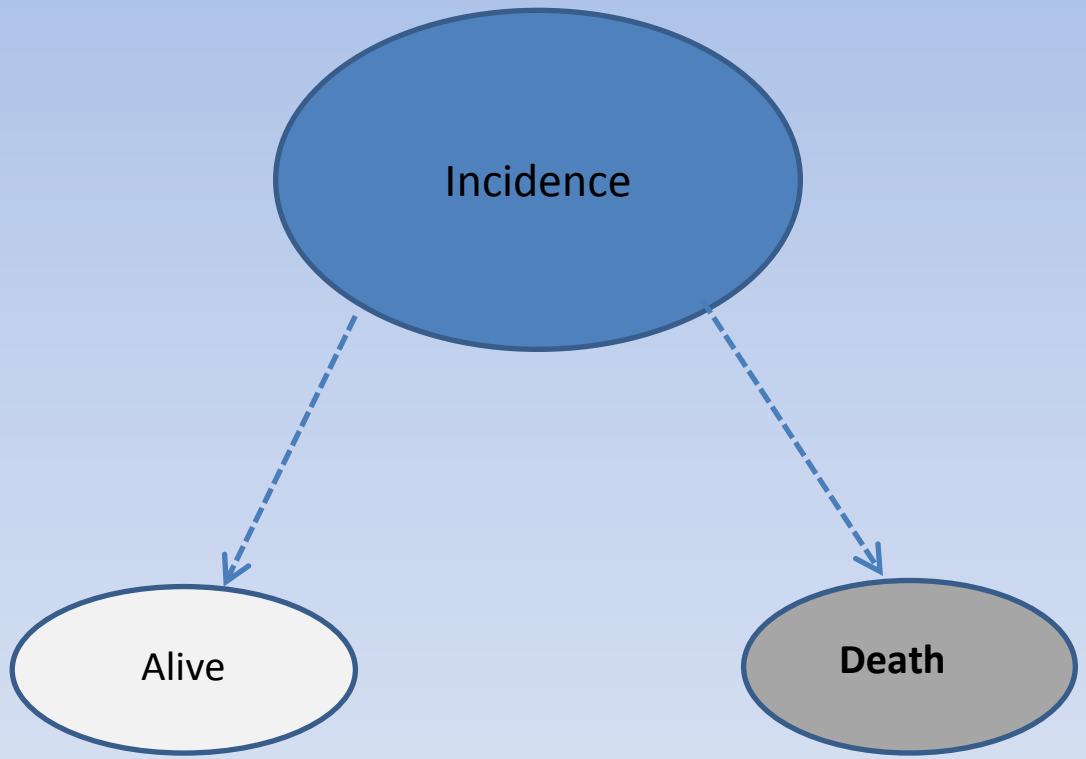
NAACCR

% Type Diagnostic Confirmation NPCR (1995-2008)



NAACCR

Follow-Up



-

NAACCR

Demographic Variables

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

% Missing - Linkage Variables

	SS#	Birth Date	Sex	Last Name	First Name
SEER - range	0.00-3.93	0.00-0.09	0.00-0.02	0.00-<0.00	0.00-<0.00
- No. states w/missing	7/10	4/10	3/10	1/10	3/10
NPCR - range	0.00-2.58	0.00-0.07	0.00-0.03	0.00-<0.02	0.00-<0.02
- No. states w/missing	30/41	21/41	22/41	9/10	15/41

Source: M Jim, IHS linkage data, variable years of diagnosis

NAACCR

Follow-Up Variables: Inter-Field and Intra-Record Edits

Data Variables and Edits

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

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

Edits associated with vital status variables needed for survival analysis
Age, Histologic Type, COD, ICDO3 (SEER IF43)
Cause of Death (SEER COD)
Date of Last Contact (NAACCR DATEEDIT)
Date of Last Contact Flag (NAACCR)
Date of Last Contact, Date Flag(NAACCR)
Date of Last Contact, Date of Diag. (NAACCR IF19)
Follow-Up Source (COC)
Follow-up Source Central (NAACCR)
Follow-Up Source Central, Vital Status (NPCR)
Follow-Up Source, Vital Status (COC)
ICD Revision Number (NPCR)
ICD Revision Number, Cause of Death (SEER IF37)
ICD Revision, Vital Stat, Date Last Contact (NPCR)
Type of Rep Srce(DC),Seq Num--Cent,ICDO3(SEER IF04
Type of Report Srce (AO), Date of Dx (SEER IF02)
Type of Report Srce(DC/AO), COD (SEER IF09)
Type of Report Srce(DC/AO), Diag Conf (SEER IF05)
Type of Report Srce(DC/AO), Vital Stat (SEER IF08)
Type of Reporting Source (SEER RPRTSRC)
Vital Status (Subm)
Vital Status, Cause of Death (Subm)
Verify cause of death same on all records for a patient (SEER IR11)
Verify date of follow-up same on all records for a patient (SEER IR08)
Verify vital status same on all records for a patient (SEER IR10)

NAACCR

Follow-Up Requirements

Alive Status

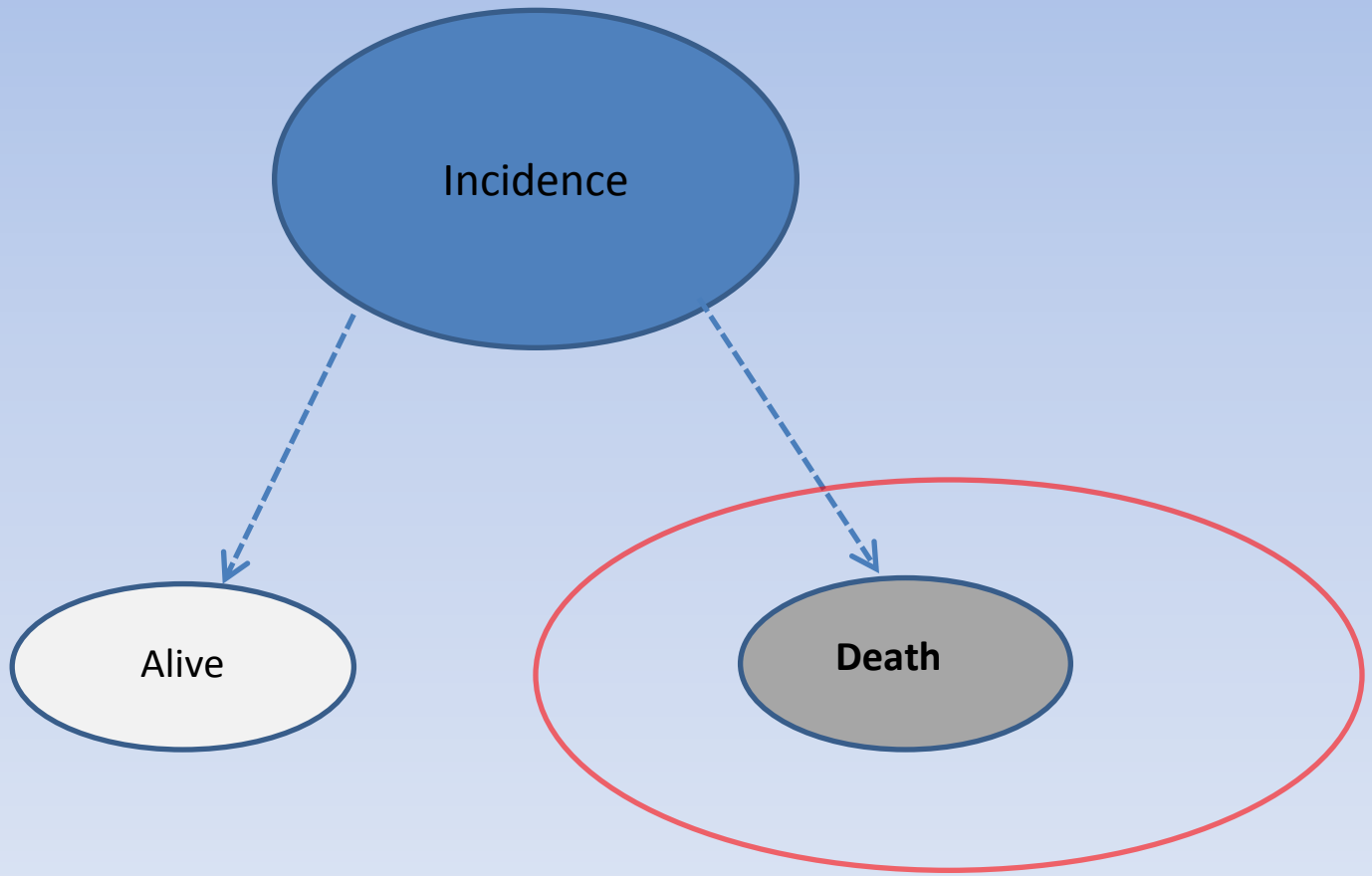
- SEER Program requires all SEER registries to follow alive patients
  - 95% patients have last contact date within 18 months of the annual date of submission
- NPCR registries are not required to follow patients

Death Staus

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

NAACCR

Events in Follow-Up



-

NAACCR

The Importance of Death Ascertainment

Johnson CJ, Weir HK, Yin D, Niu X. ***The impact of patient follow-up on population-based survival rates.*** J Registry Manag. 2010 Fall;37(3):86-103.

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

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

CONCLUSIONS:

- Complete death ascertainment important for producing accurate cancer survival statistics, and
- Ascertainment of deaths only should generally be sufficient for survival analysis.

Full Dates vs. Partial Dates

- Date of Birth
- Date of diagnosis

Age at diagnosis needed for Life Tables

- Date of last contact

Survival interval

SEER Program uses month and year

Example: Patient diagnosed April 2000 and dies May 2000. Survival interval could be 1 – 60 days

NAACCR / NPCR uses month, day and year

NAACCR

Survival Interval
Full Dates vs. Partial Dates

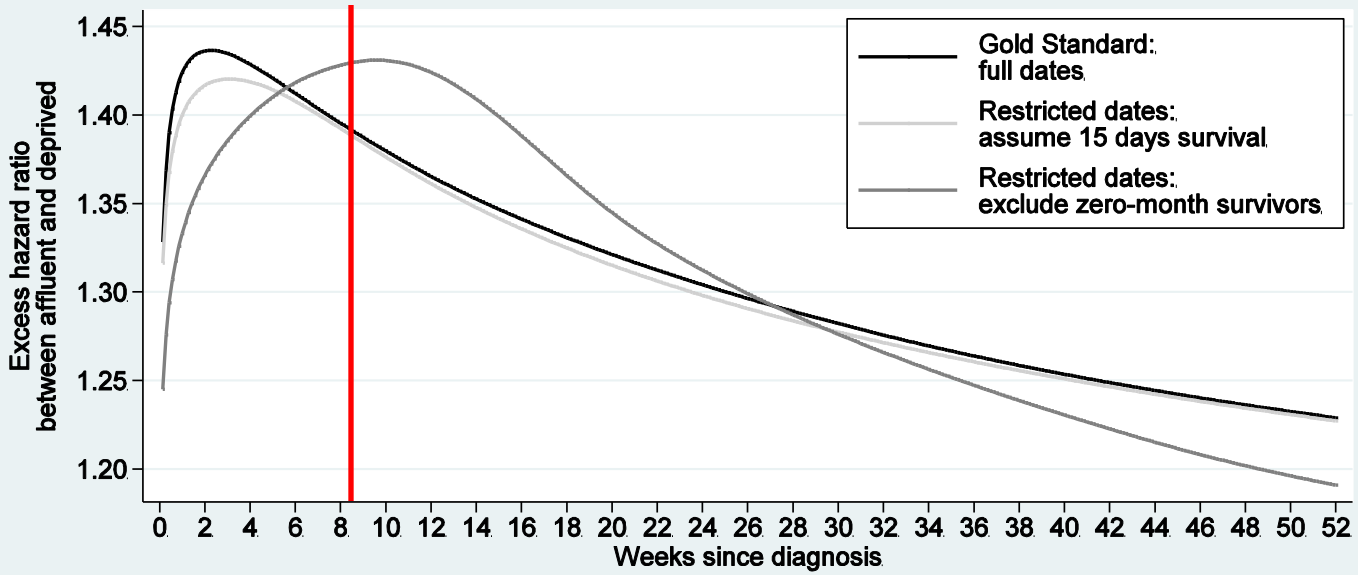
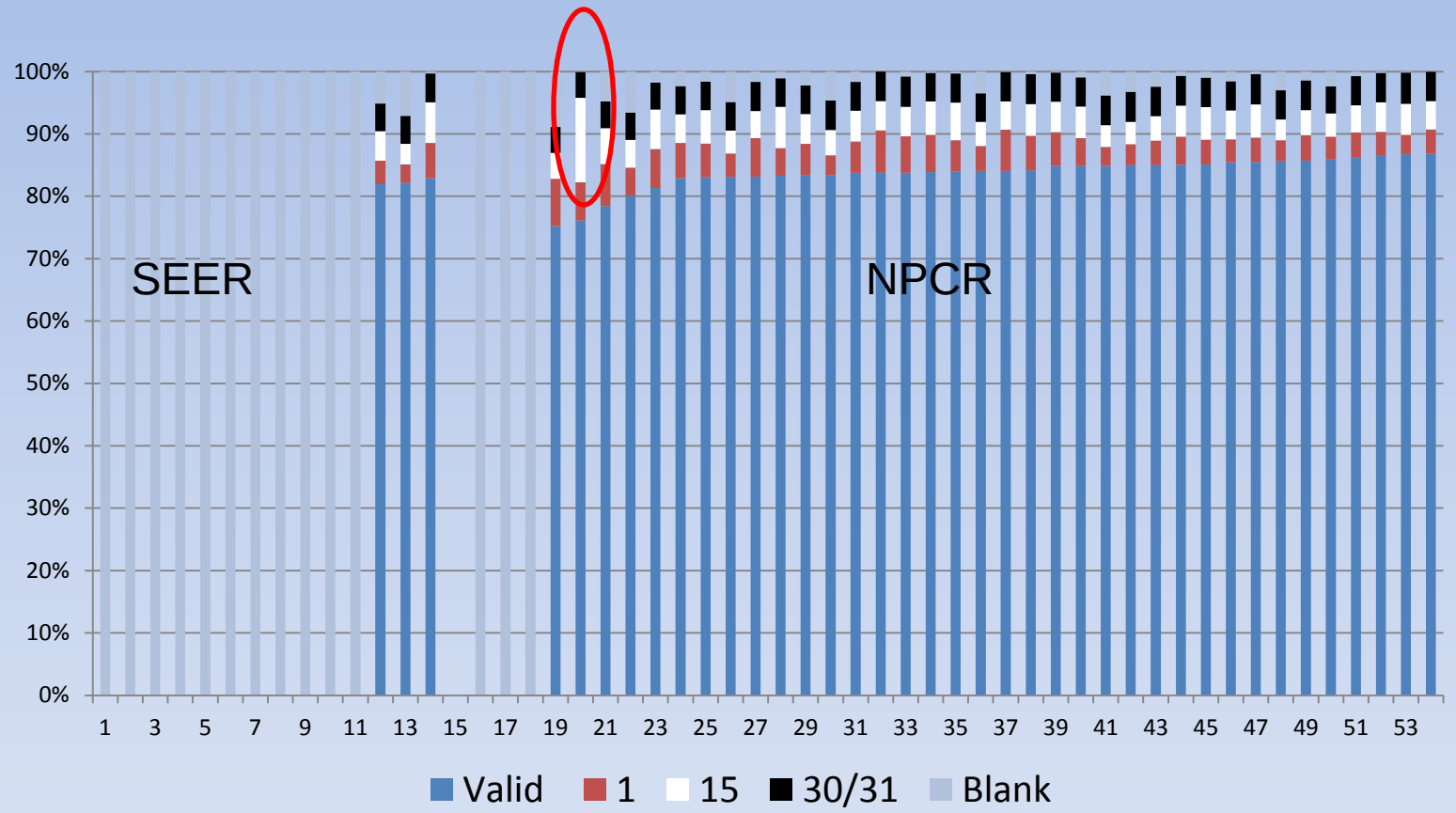


Fig 2 - colorectum (Females)

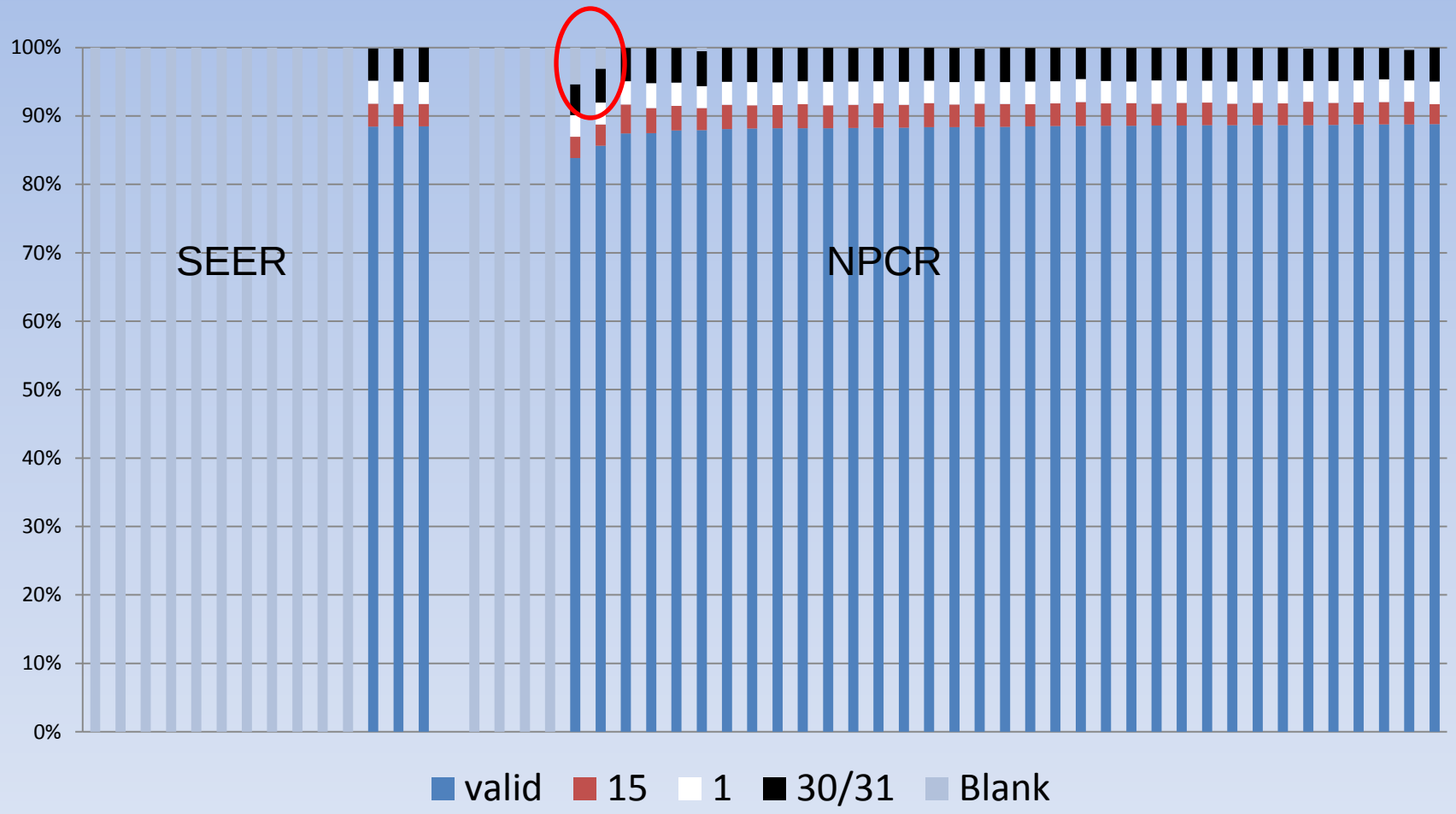
Woods LM, Rachet B, Ellis L, Coleman MP
Full dates (day, month, year) should be used in population-based cancer survival studies

NAACCR

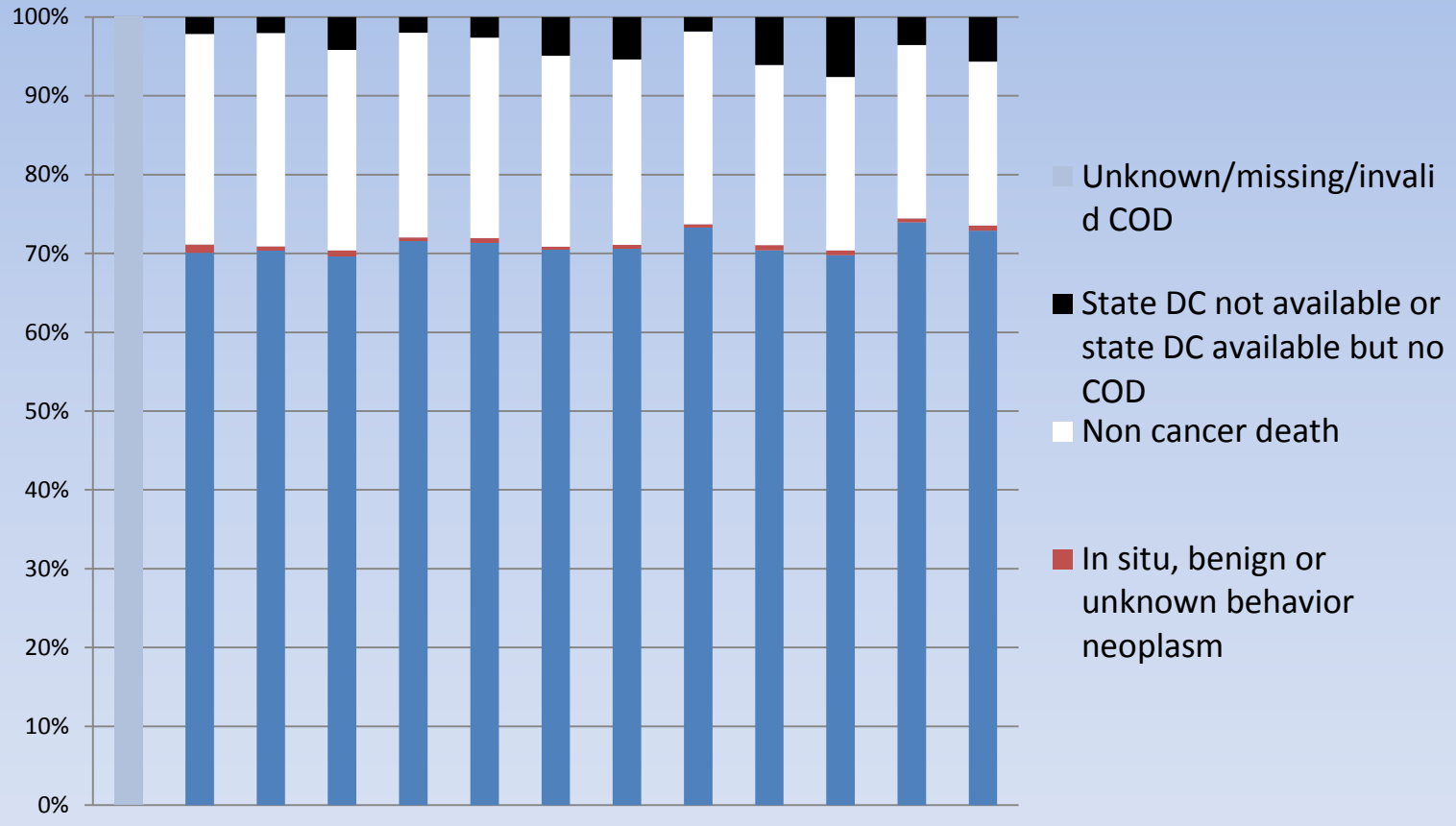
Day of Diagnosis (2004-2008)



Day of Death among Decedents (2004-2008)

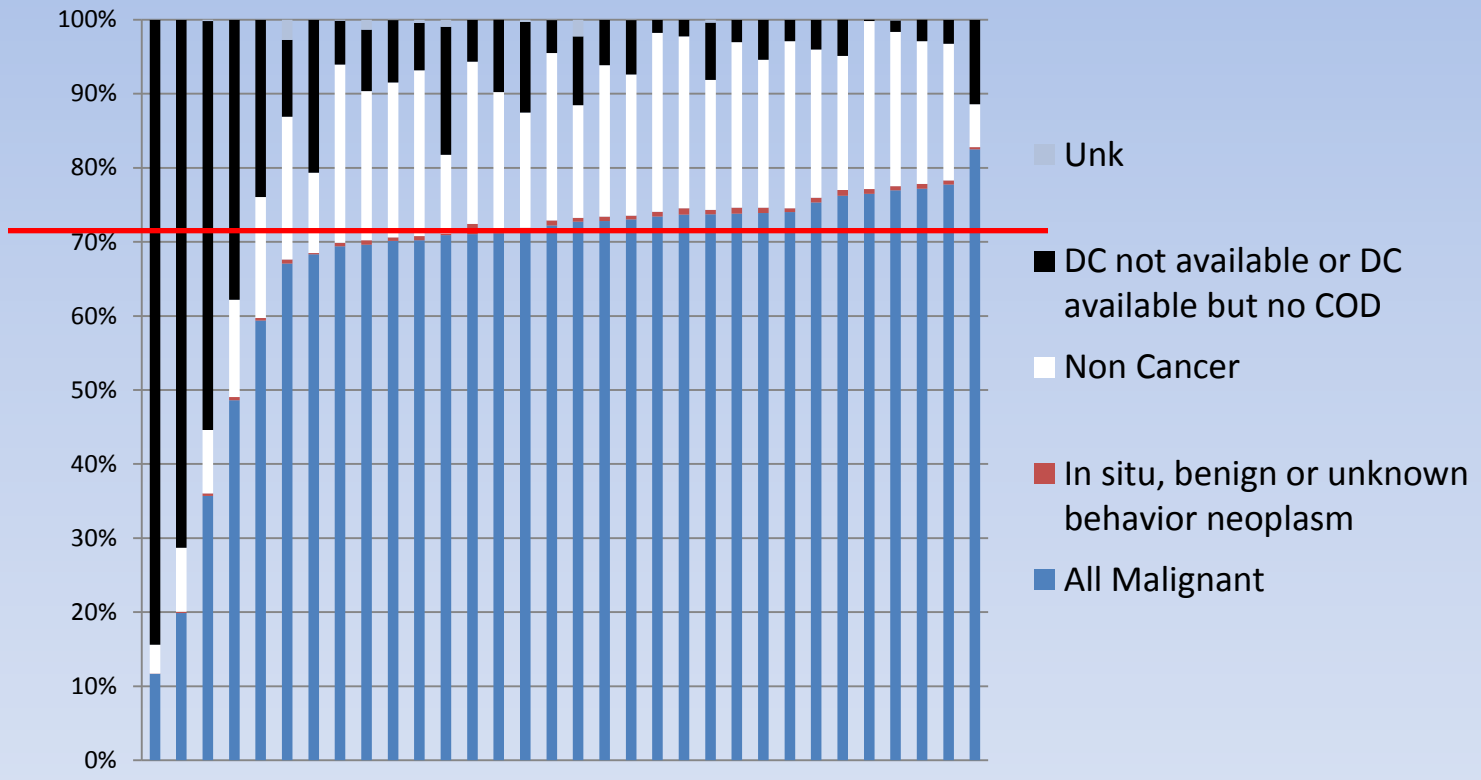


Cause of Death among Decedents SEER 1995-2008

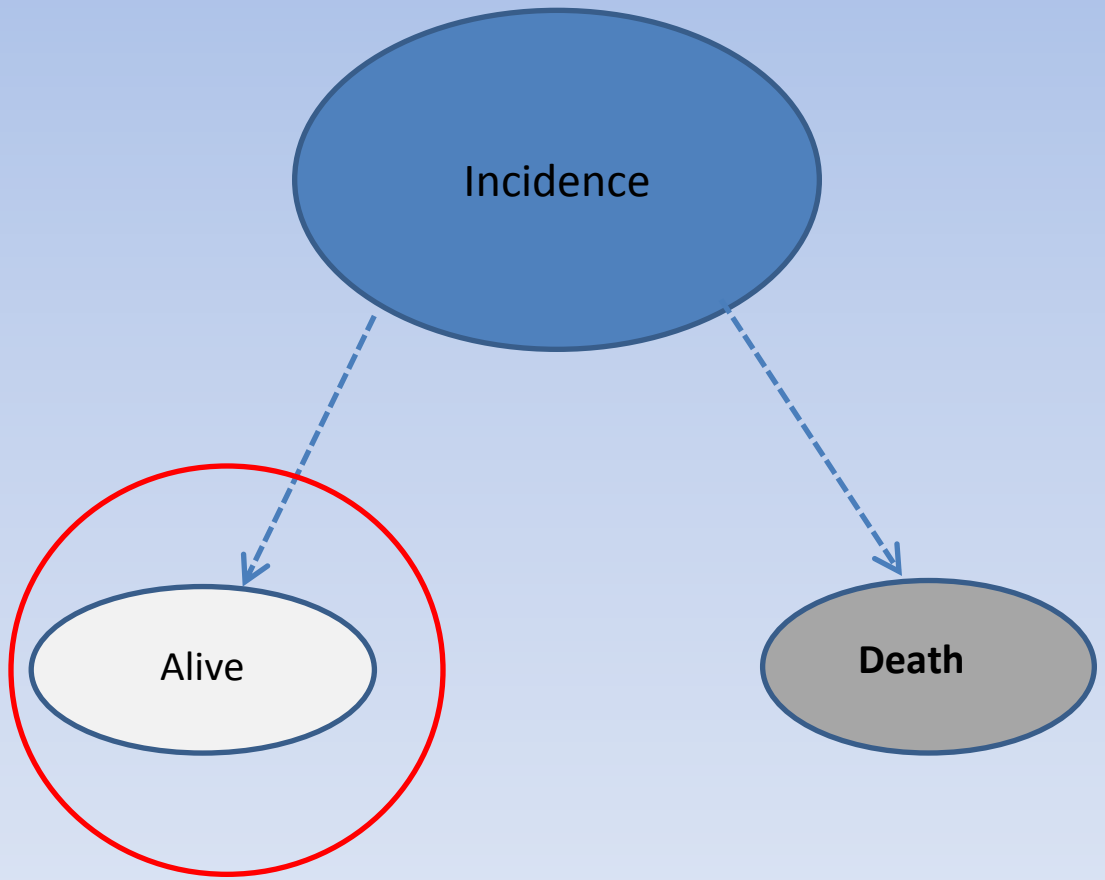


NAACCR

Cause of Death among Decedents NPCR 1995-2008



Events in Follow-Up



-

Follow-Up Requirements

Alive Status

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

Death Staus

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

NAACCR

Follow-Up Requirements

Alive Status

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

Death Staus

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

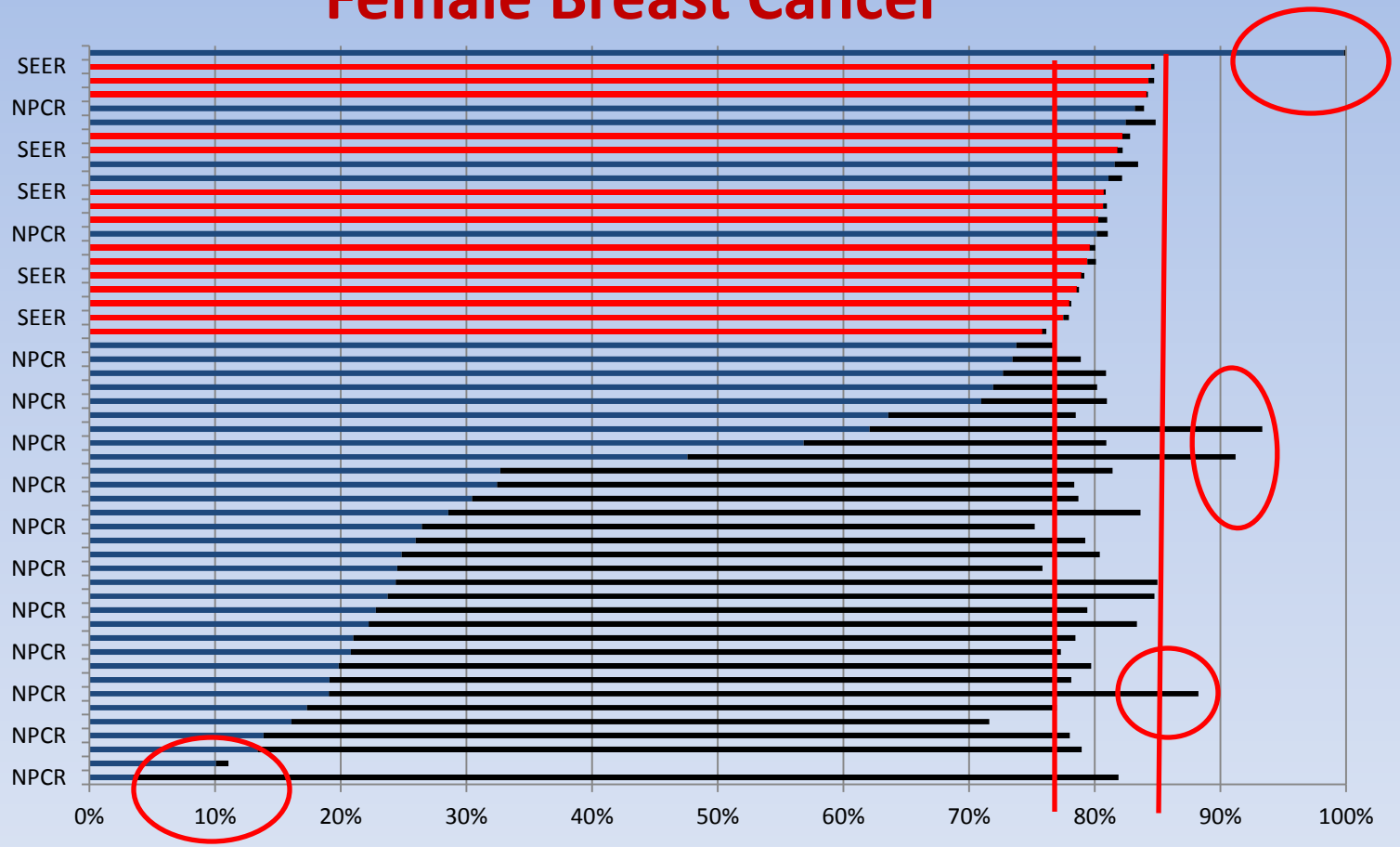
NAACCR

Overview

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

60-Month Observed Survival 2003-2007 Cases Followed Through 2008

Female Breast Cancer



SEER NPCR

NAACCR

Overview

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

SAWG - Next Steps

- Deceased with 0 survival time (and not a DCO/AO case)
  - E.g., Physician only reporting source, follow up source central (State or NDI). These events are included in analysis whereas DCO/AO cases are excluded
- Immortal cases
- Survival using full dates - SEER*Stat enhancement
- State specific life tables – available in 2012
- NAACCR Webinar June 14, 2012
- CONCORD Study

NAACCR

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

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

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

NAACCR