**THE FUTURE OF CANCER DATA:** UNLOCKING INSIGHTS WITH PATHOLOGY REPORTING

# **Pathology Report to Public Health and Back:** Supporting Cancer Research and Discoveries

Alison Van Dyke, MD, PhD

OCTOBER 6 | 3:00–3:45 PM CT

COLLEGE of AMERICAN PATHOLOGISTS
Laboratory Quality Solutions

CAP23 | CHICAGO

#PATHDATA

50TH ANNIVERSARY SEER

TURNING CANCER DATA INTO DISCOVERY

# Pathology Report to Public Health and Back: Supporting Cancer Research and Discoveries

Alison Van Dyke, MD, PhD, FCAP

October 6, 2023

NIH NATIONAL CANCER INSTITUTE

# Conflicts of Interest & Support

I have no conflicts of interest or other financial disclosures to declare.

The work described in this presentation and the NCI/SEER and CDC/NPCR cancer registry systems are funded and operated by the U.S. Federal Government.

I am employed by the National Cancer Institute of the National Institutes of Health.

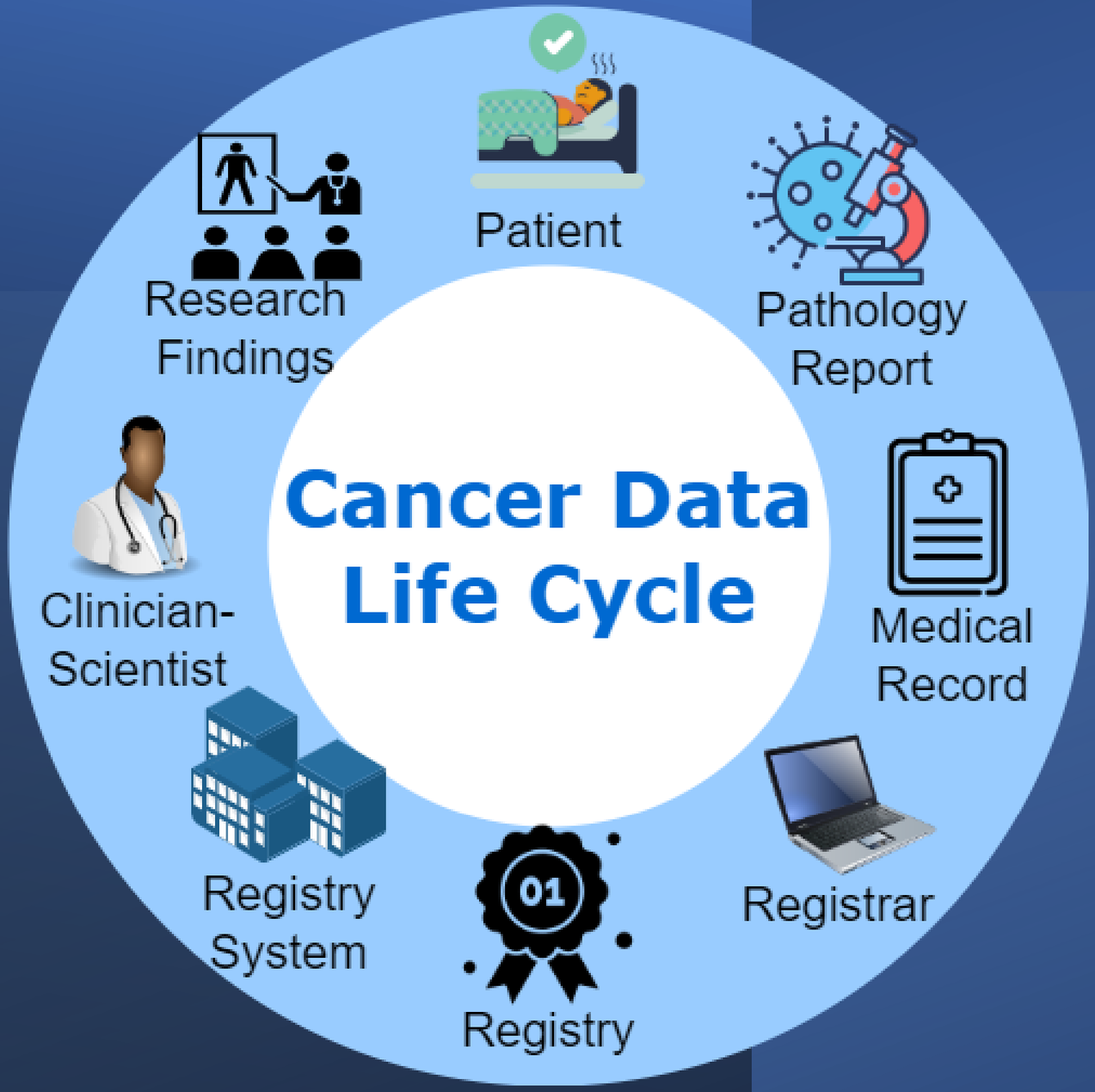
# Objectives

Understand the life cycle of pathology report data in cancer surveillance and how it can lead to advances in cancer care

Understand how variability of terminology used in pathology reports and outdated standards lead to inconsistencies and inaccuracies in data captured

Know about a cancer surveillance initiative with CAP to facilitate accurate cancer registry data collection and how to gain access to population-based cancer data for research

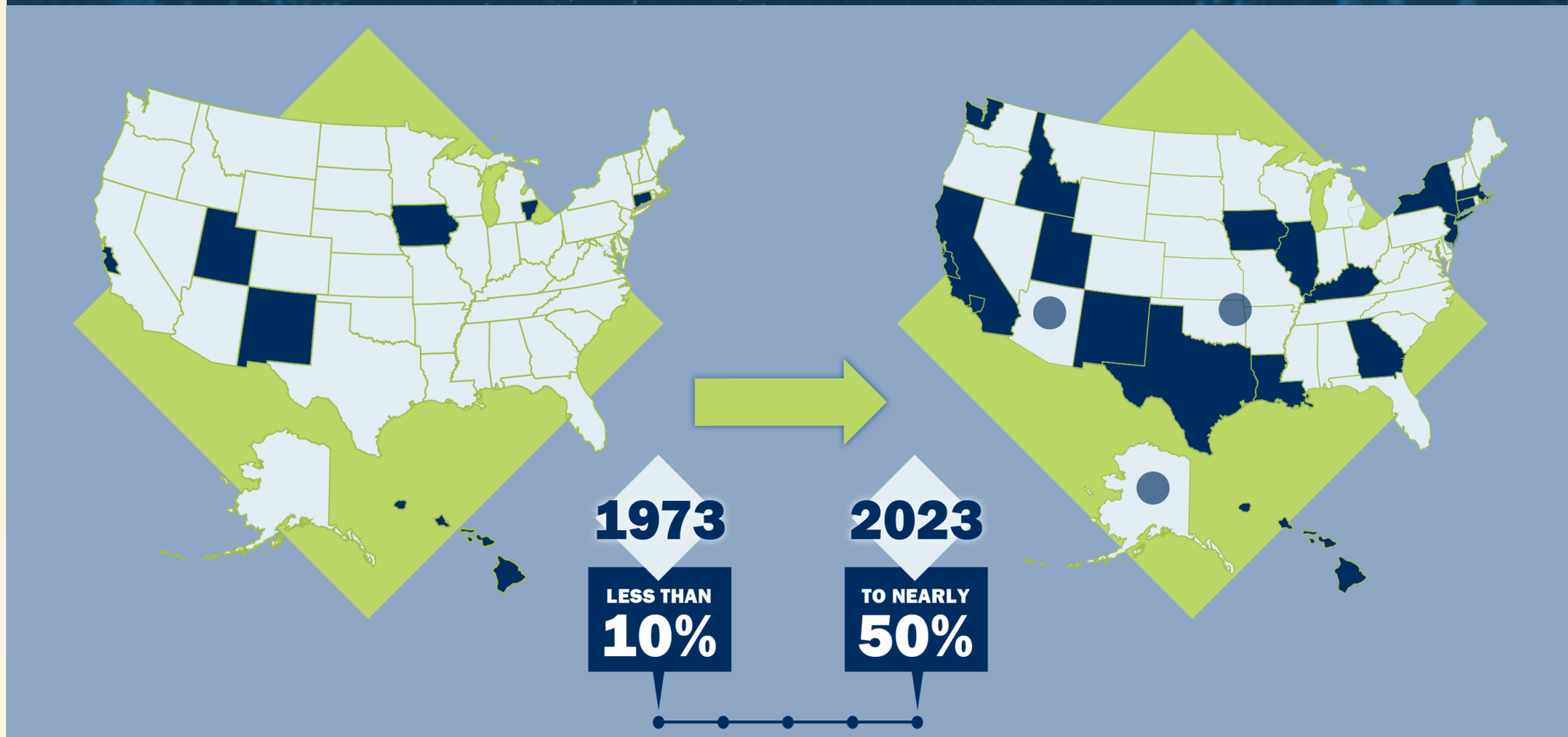
# Cancer Surveillance & Pathologists: Data Quality Partners



50TH ANNIVERSARY SEER

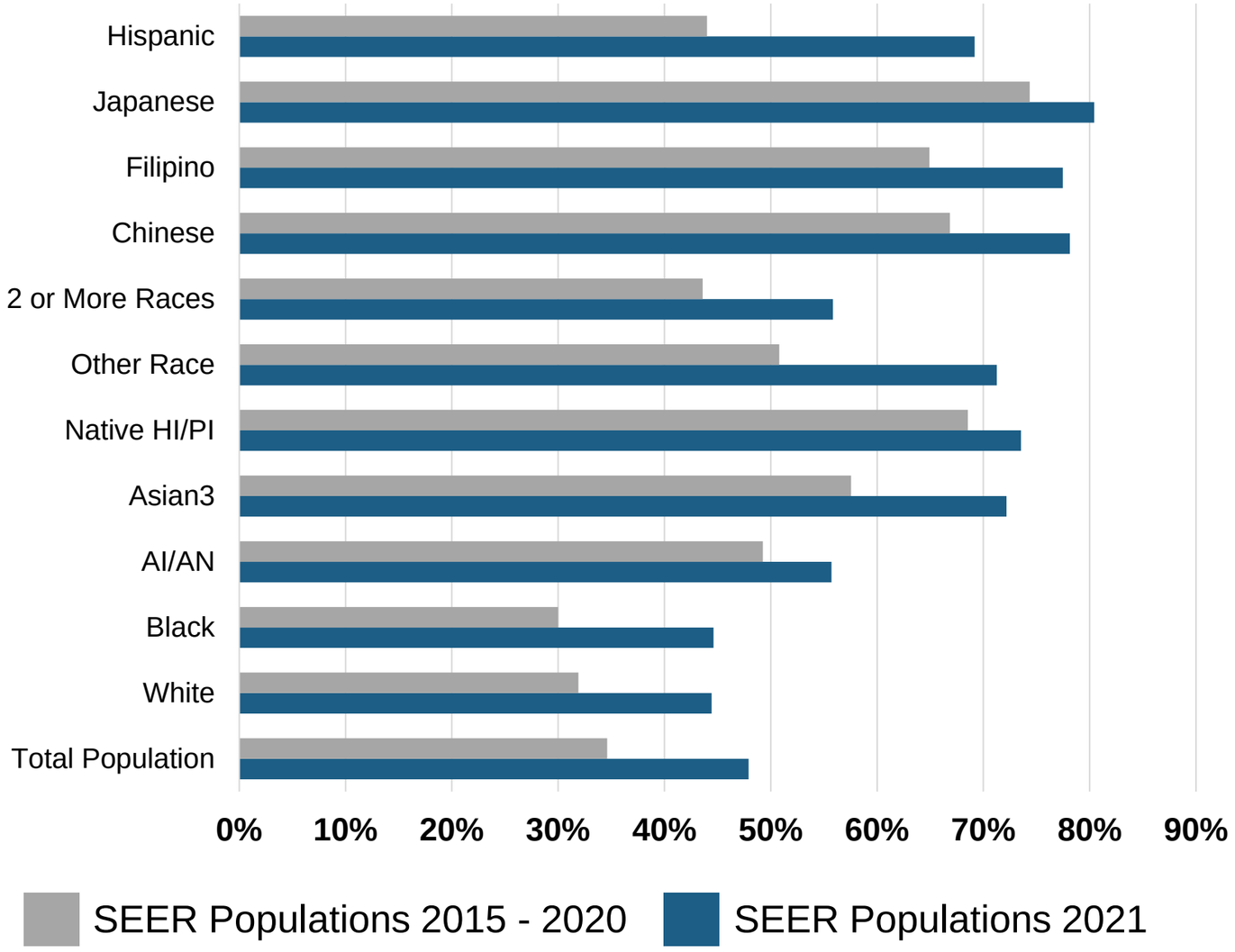
TURNING CANCER DATA INTO DISCOVERY

- >850,000 cases/year
- Rare cancers
- Cases with rare outcomes
- Additional cancers
- Understudied populations



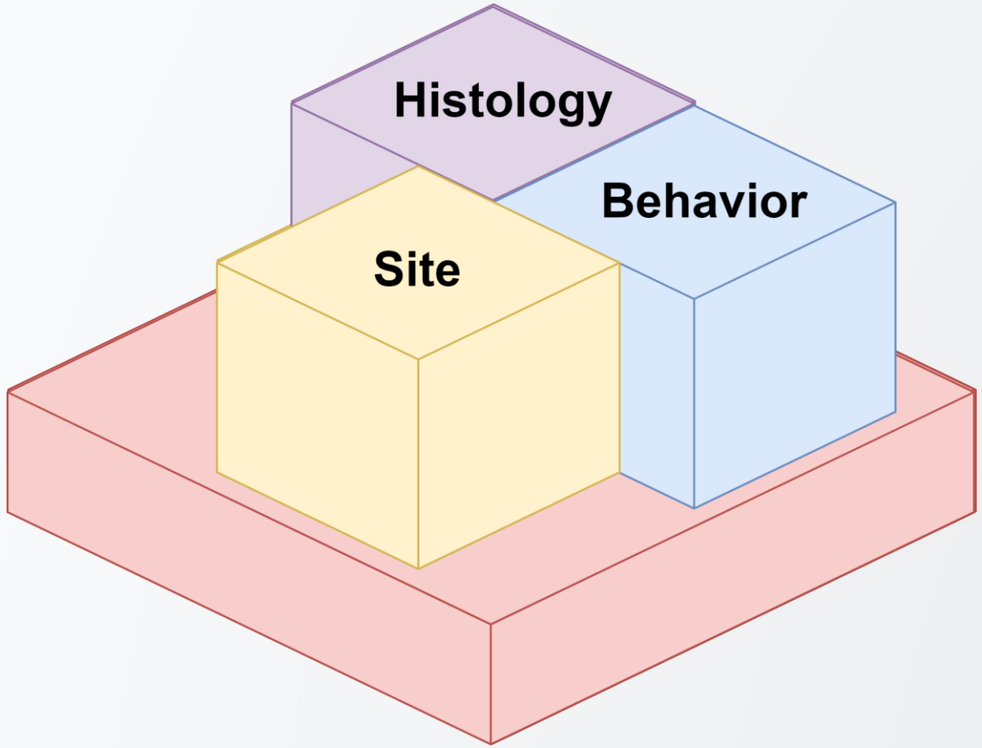

NIH NATIONAL CANCER INSTITUTE

# Increase in Representation of Population Subgroups with SEER Expansion




| Percent Increase for US Population Subgroups with SEER Expansion 2021 | |
|---|---|
| **Total Population** | **13.3** |
| White | 12.5 |
| Black | 14.6 |
| American Indian/Alaska Native | 6.5 |
| Asian | 14.6 |
| Native Hawaiian/Pacific Islander | 5.0 |
| 2 or more races | 12.3 |
| Chinese | 11.3 |
| Filipino | 12.5 |
| Japanese | 6.1 |
| Hispanic | 25.2 |

# Foundation of Cancer Surveillance Data Capture




## Site Specific Data Items (SSDIs) – ex: Breast

- ER, PR, HER2
- Axillary nodal involvement
- Oncotype DX
- Multigene Signature

## Statistics Reported for

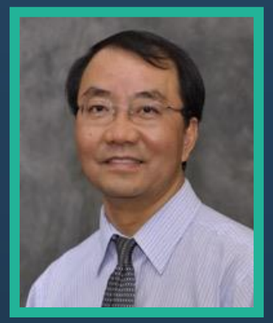
- Incidence
- Outcomes
- Trends
- Additional Cancers

# SEER*ClinCORE Pathologists

**Aaron Auerbach**
Hematopathology

**James Connolly**
Breast Pathology

**Brent Harris**
Neuropathology



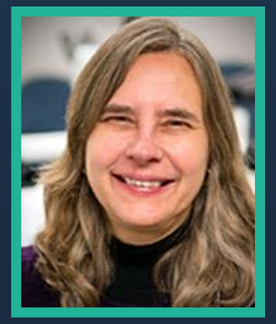
**Pei Hui**
GYN Pathology

**Peter Humphrey**
Male Genital/Urinary Pathology

**Jim Lewis Jr.**
Head/Neck Pathology & HPV

**Ricardo Lloyd**
Endocrine Pathology

**Jessica Davis**
Bone/Soft Tissue & Pediatric Pathology



**Kay Washington**
GI Pathology

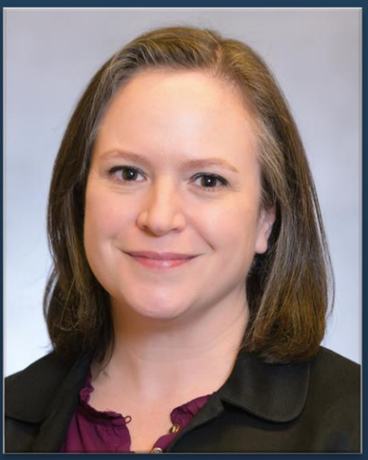
**Priya Nagarajan**
Dermatopathology

# CAP-NCI Problem Solving – Site-Morphology Combinations, Terminology & Coding

Richard Moldwin
CAP

Keren Hulkower
CAP

Serban Negoita
NCI



Alison Van Dyke
NCI

- Different terminology used across standard-setters & stakeholders
- Variation in terminology & coding over time
- ~24-month timeline for implementing new histology standards in cancer surveillance
- Cancer surveillance standards overdue for major overhaul

# Cancer PathCHART Acronym

**C**ancer

**Path**ology

**C**oding

**H**istology

**A**nd

**R**egistration **T**erminology

# Collaborating Organizations

International Agency for Research on Cancer

World Health Organization

National Cancer Registrars Association

NAACCR

SEER

NPCR

NATIONAL PROGRAM of CANCER REGISTRIES

Commission on Cancer®

Statistics Canada

Statistique Canada

IACR

International Association of Cancer Registries

ICCR

International Collaboration on Cancer Reporting

COLLEGE of AMERICAN PATHOLOGISTS

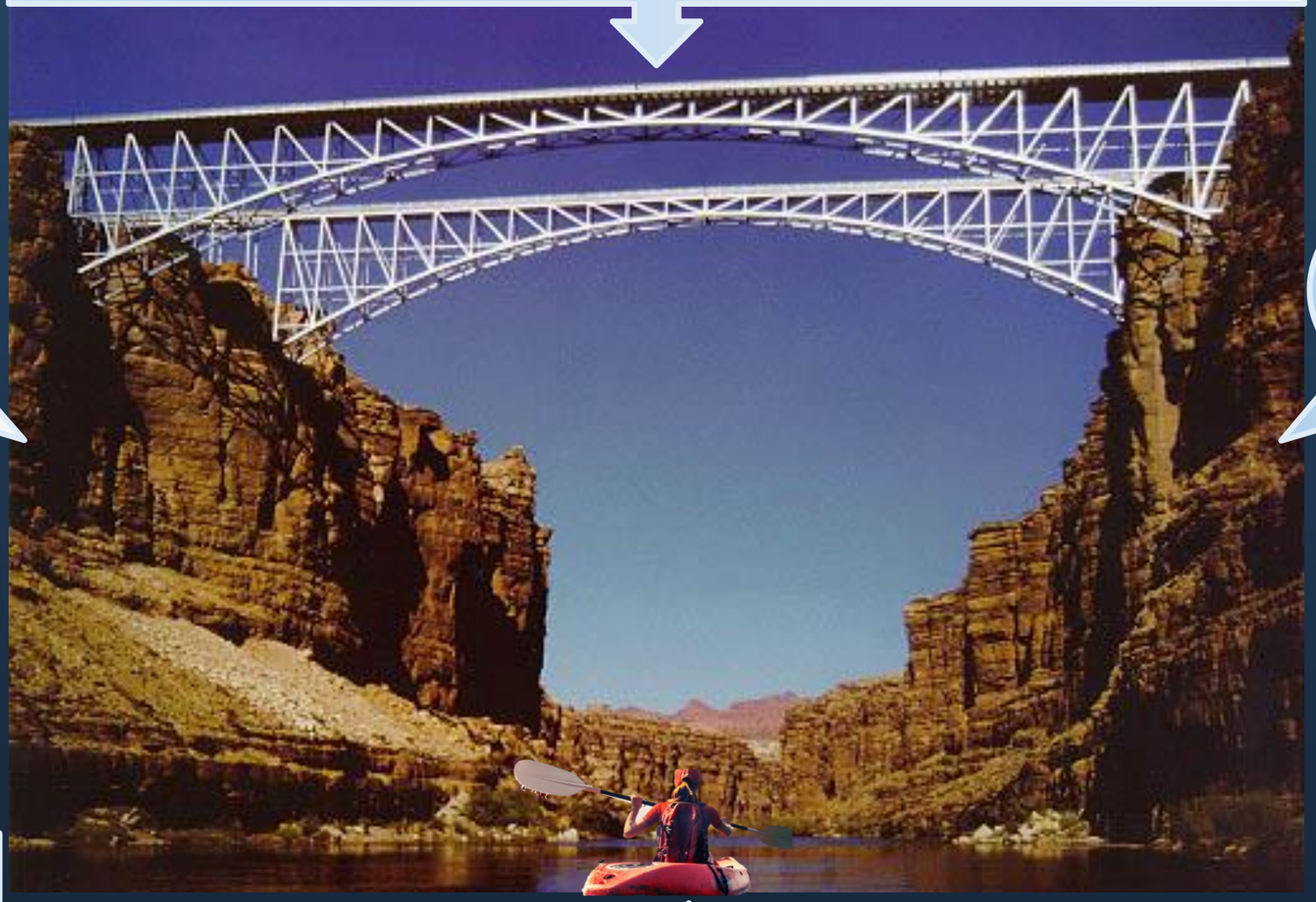
AJCC

American Joint Committee on Cancer

100+years

# Cancer PathCHART

Cancer Surveillance Standards



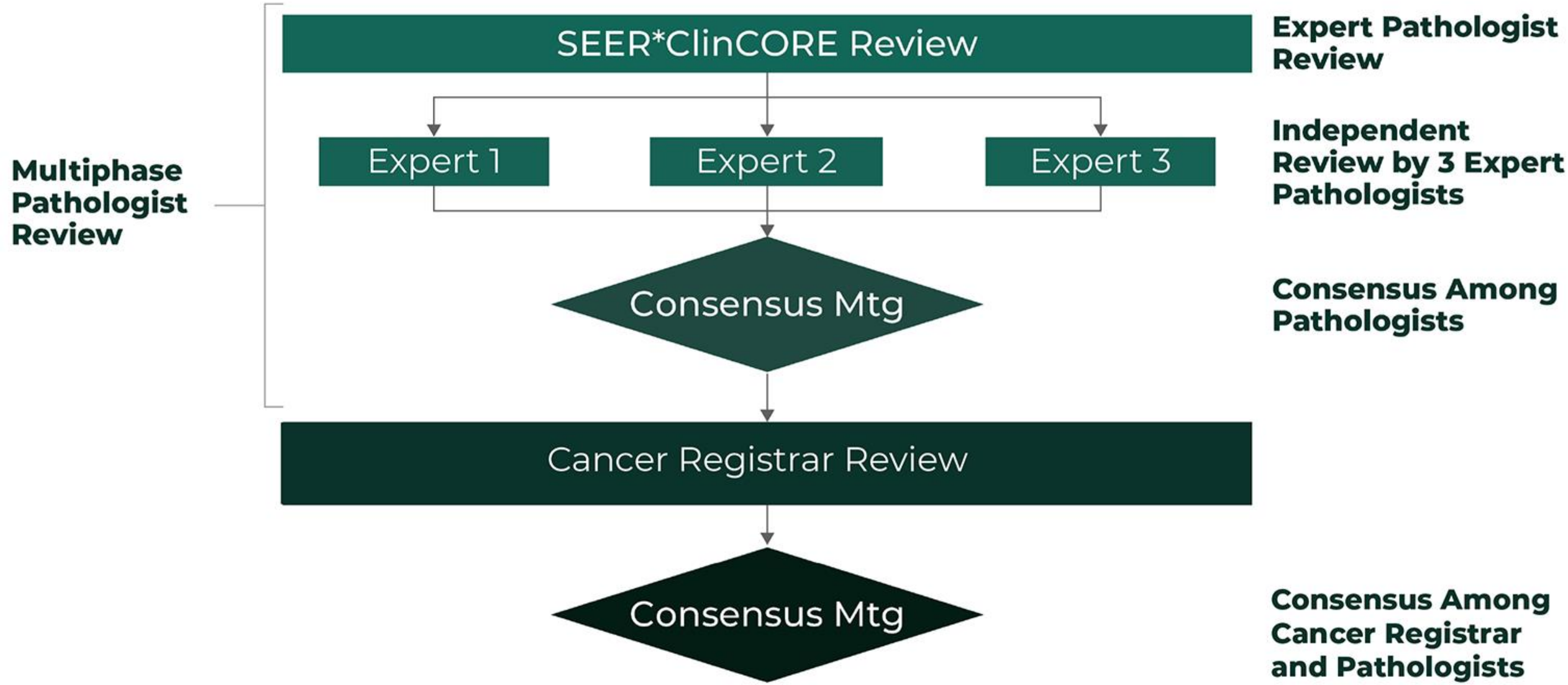
Language in Pathology Reports

Cancer Registrar

## Picture Link

Modified from: U.S. Department of Transportation, Federal Highway Administration. Accessed April 28, 2023

# Interdisciplinary Review Process



# Pathologist Reviewer Decisions

## Biologically Valid

**No further review needed**

*Example*
Adenocarcinoma of the colon & rectum

## Biologically Unlikely

**Histology is unlikely in this site/organ system and may be an error**

*Example*
Squamous cell carcinoma in situ of the rectum (more likely of the anal canal)

## Biologically Impossible

**Cancer registrars cannot record this combination in the cancer registry database**

*Example*
Hepatocellular carcinoma of the prostate

## Send for Consensus

**Determination to be made via consensus among multiple pathologists and CTRs**

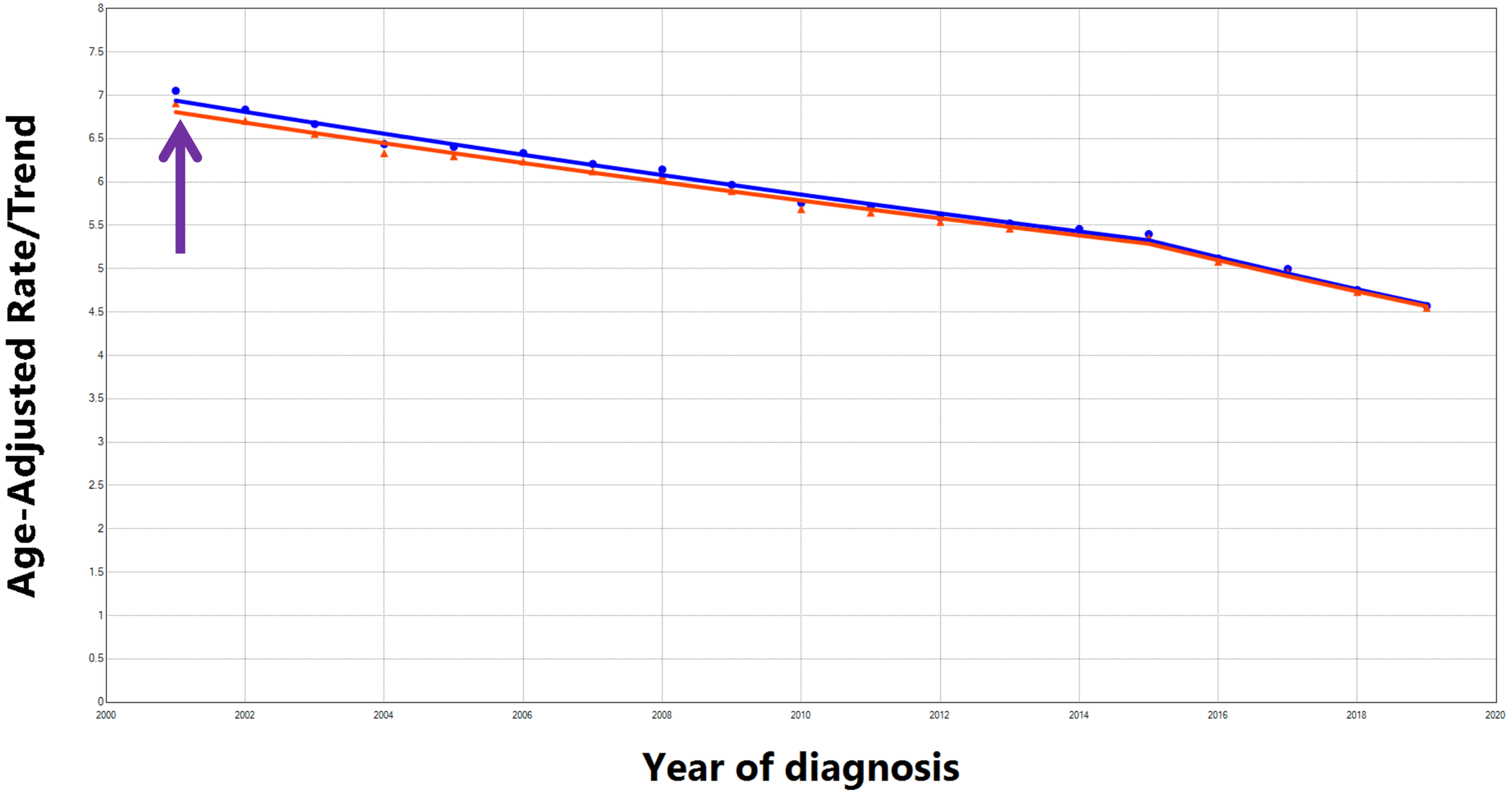
# Previously Valid Ovarian Histologies Deemed Impossible - examples

| Morphology | ICD-O-3.2 Term | Count |
|---|---|---|
| 8051/3 | Verrucous carcinoma, NOS | 0 |
| 8052/2 | Papillary squamous cell carcinoma, non-invasive | 0 |
| 8070/2 | Squamous cell carcinoma in situ, NOS | 41 |
| 8230/2 | Ductal carcinoma in situ, solid type | 0 |
| 8261/2 | Adenocarcinoma in situ in villous adenoma | 0 |
| 8261/3 | Adenocarcinoma in villous adenoma | 0 |
| 8262/3 | Villous adenocarcinoma | 4 |
| 8263/2 | Adenocarcinoma in situ in tubulovillous adenoma | 0 |
| 8263/3 | Adenocarcinoma in tubulovillous adenoma | 21 |
| 8510/3 | Medullary carcinoma, NOS | 2 |

# Ovarian Cases: Impact estimate of CPC changes

| Before Review | # of Histologies | 2019 Case Count | % of Total Cases | Expert Review Designation | # of Histologies | 2019 Case Count | % of Total Cases |
|---|---|---|---|---|---|---|---|
| SEER Site/Type Validation List | 107 | 17,932 | 99.0 | Valid | 64 | 17,806 | 98.7 |
| | | | | Unlikely | 3 | 9 | <0.1 |
| | | | | Impossible | 40 | 117 | 0.6 |
| Manual Review/Override | 228 | 110 | 1.0 | Valid | 5 | 4 | <0.1 |
| | | | | Unlikely | 40 | 33 | 0.2 |
| | | | | Impossible | 183 | 73 | 0.4 |
| Impossible | 3 | 0 | 0 | Valid | 0 | 0 | 0 |
| | | | | Unlikely | 0 | 0 | 0 |
| | | | | Impossible | 3 | 0 | 0 |
| New WHO Code/Term | 7 | 0 | 0 | Valid | 0 | 0 | 0 |
| | | | | Unlikely | 0 | 0 | 0 |
| | | | | Impossible | 7 | 0 | 0 |
| **Total** | **345** | **18,042** | **100** | **Total** | **345** | **18,042** | **100** |

# Valid or Override Histologies of the Ovary



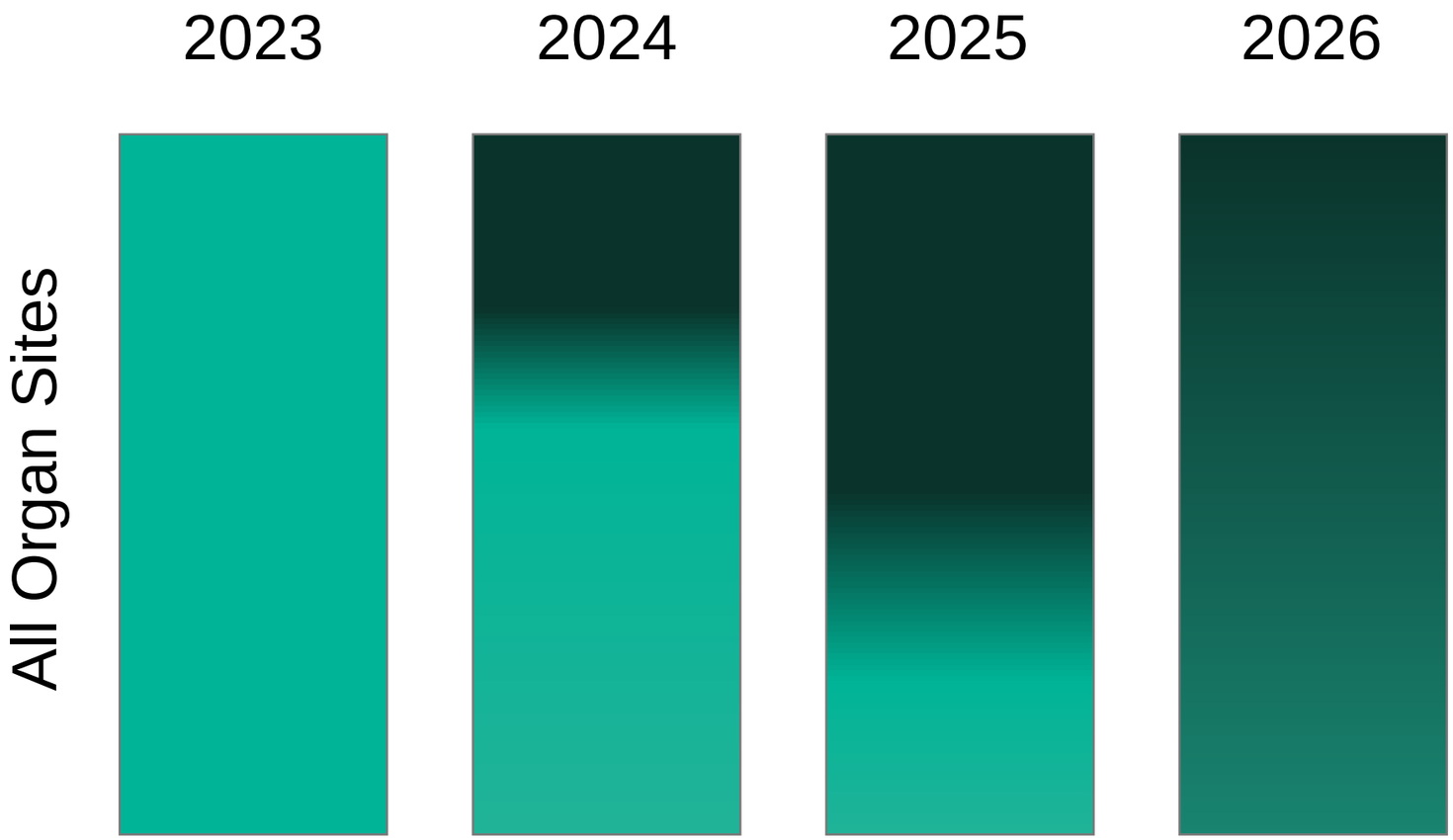

Prior to CPC

Post CPC

# Implementation Timeline

Cancer PathCHART Updated Standards

Previous Standards



# Pathologist Reviewers-completed

Bone & Soft Tissue

John SA Chrisinger, MD
Jessica Davis, MD
Karen Fritchie, MD
Paari Murugan, MD

Breast

Veerle Bossuyt, MD
James Leo Connolly, MD
Mary Elizabeth Edgerton, MD, PhD
Patrick L. Fitzgibbons, MD

Central Nervous System

Brent Harris, MD, PhD
David Louis, MD
Arie Perry, MD

Digestive System

Volkan Adsay, MD
Olca Basturk, MD
Norman Carr, MB, BS, FRCPath
Jessica Davis, MD
Dhanpat Jain, MD
Sanjay Kakar, MD
Gregory Lauwers, MD
Robert Odze, MD
Asif Rashid, MBBS, PhD
Romil Saxena, MD
Chan Juan Shi, MD, PhD
Aatur Singhi, MD, PhD
Mike Torbenson, MD
Kay Washington, MD, PhD
Tsung-The Wu, MD, PhD

# Pathologist Reviewers-completed

**Female Genital System**

Elizabeth Euscher, MD

Ian Hagemann, MD, PhD

Pei Hui, MD, PhD

Martin Kobel, MD

Uma Krishnamurti, MD, MBBS, PhD

Mohammad Ruhul Quddus, MD

Brian Rous, MD

Jian-Jun Wei, MD

**Male Genital System**

Michael  Eden, MBBS, FRCPath

Jonathan Epstein, MD

Peter Humphrey, MD, PhD

Gladell P. Paner, MD

Joseph Sirintrapun, MD

John Robert Srigley, MD, FRCPath

**Urinary System**

Jonathan Epstein, MD

Lara Rabih Harik, MD

Peter Humphrey, MD, PhD

# For More Information

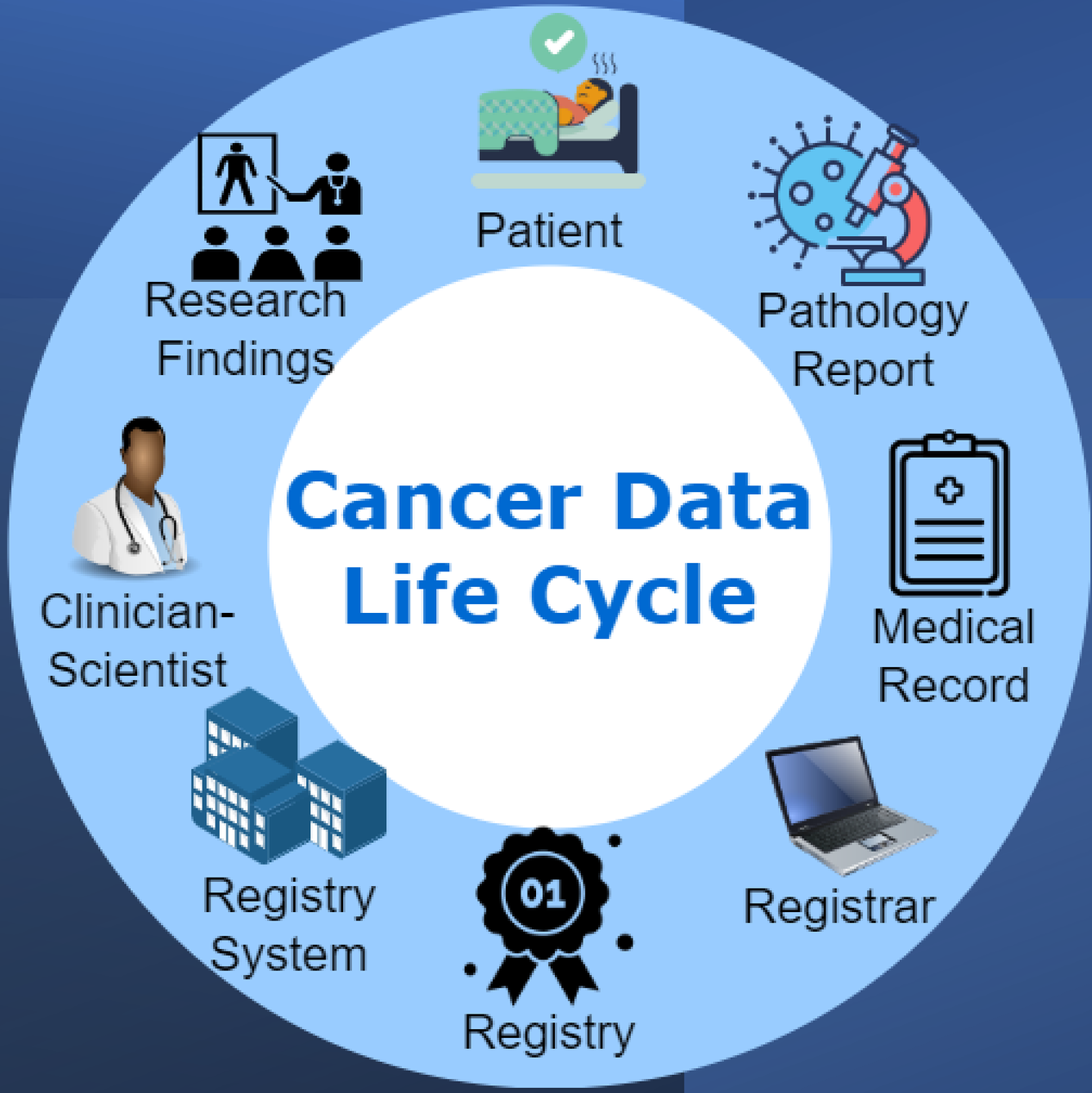
Visit the Cancer PathCHART website today!

https://seer.cancer.gov/cancerpathchart/

Contact Us at NCICancerPathCHART@mail.nih.gov

***Search tool of 2024 standards in development***

# RWE Data Examples



# Real World Evidence: Example #1

## Characterization of PLCIS on a Population Scale

| Time Period | LCIS | PLCIS |
|---|---|---|
| 2017 | 8520/2 | 8520/2 |
| 2018 forward | 8520/2 | 8519/2 |

Addressed using US Cancer Statistics (USCS)

316 PLCIS cases diagnosed in US (2018 to 2020)

Analysis of

- differences in demographic & tumor characteristics by cancer type
- risk of subsequent breast cancers after PLCIS vs. DCIS or LCIS
- differences between location & laterality of initial primary & second primary

| Characteristic | PLCIS N=316 | LCIS N=13,179 | DCIS N=104,834 | Invasive N=609,143 | P-value (PLCIS vs LCIS) | P-value (PLCIS vs DCIS) | P-value (PLCIS vs. Invasive) |
|---|---|---|---|---|---|---|---|
| ***Age (years) - Median (IQR)*** | 61 (52, 68) | 53 (47, 62) | 61 (51, 69) | 62 (52, 71) | **<0.0001** | 0.99 | 0.08 |
| ***Race/Ethnicity - N (%)*** | | | | | | | |
| Non-Hispanic White | 227 (75) | 9,168 (73) | 69,383 (69) | 425,553 (72) | 0.09 | **0.009** | 0.07 |
| Non-Hispanic Black | 24 (8) | 1,323 (11) | 14,272 (14) | 71,749 (12) | | | |
| Hispanic | 28 (9) | 1,466 (12) | 9,977 (10) | 58,828 (10) | | | |
| Other | 23 (8) | 663 (5) | 7,648 (8) | 34,032 (6) | | | |
| ***Location-initial – N (%)*** | | | | | | | |
| Nipple – C50.0 | 0 (0) | 35 (0.3) | 506 (0.5) | 2,327 (0.4) | **0.006** | 0.15 | **0.01** |
| Central - C50.1 | 13 (4) | 725 (6) | 6,659 (6) | 27,512 (5) | | | |
| UIQ - C50.2 | 21 (7) | 1,048 (8) | 9,826 (9) | 77,257 (13) | | | |
| LIQ - C50.3 | 17 (5) | 443 (3) | 6,802 (7) | 33,383 (6) | | | |
| UOQ - C50.4 | 115 (36) | 4,441 (34) | 34,227 (33) | 214,118 (35) | | | |
| LOQ - C50.5 | 27 (8) | 808 (6) | 8,244 (8) | 46,934 (8) | | | |
| Axillary Tail - C50.6 | 0 (0) | 25 (0.2) | 95 (0.1) | 2,396 (0.4) | | | |
| Overlapping - C50.8 | 80 (25) | 2,913 (22) | 25,459 (24) | 140,231 (23) | | | |
| Breast, NOS - C50.9 | 43 (14) | 2,741 (21) | 13,016 (12) | 64,985 (11) | | | |

# PLCIS Cases: Morphology of Subsequent Primary

| Morphology | N (%) |
|---|---|
| *Any second breast primary (/2 or /3)* | **N=19** |
| Invasive Breast Cancer (/3) | 8 (42) |
| DCIS (8500/2) | 8 (42) |
| PLCIS (8519/2) | 2 (11) |
| LCIS (8520/2) | 1 (5) |
| | |
| *Invasive second breast primary (/3)* | **N=8** |
| Infiltrating duct carcinoma, NOS (8500) | 3 (38) |
| Invasive lobular carcinoma (8520) | 4 (50) |
| Infiltrating duct & lobular carcinoma (8522) | 1 (12) |

# Initial Comparisons & Future Analysis

## PLCIS with vs. without subsequent breast primary

- No differences in age, race/ethnicity, or location of the initial cancer

## With additional follow up time & diagnosis years

- Repeat analysis of demographic characteristics
- Risk of subsequent breast cancers comparing patients with
  - PLCIS and DCIS
  - PLCIS and LCIS
- Location of initial vs. subsequent primary among PLCIS vs. DCIS patients

# Acknowledgements – Breast PLCIS Analysis


NCI/SEER
Lois Dickie
Serban Negoita
Annie Noone
Alison Van Dyke

CDC/NPCR
Trevor Thompson
Manxia Wu

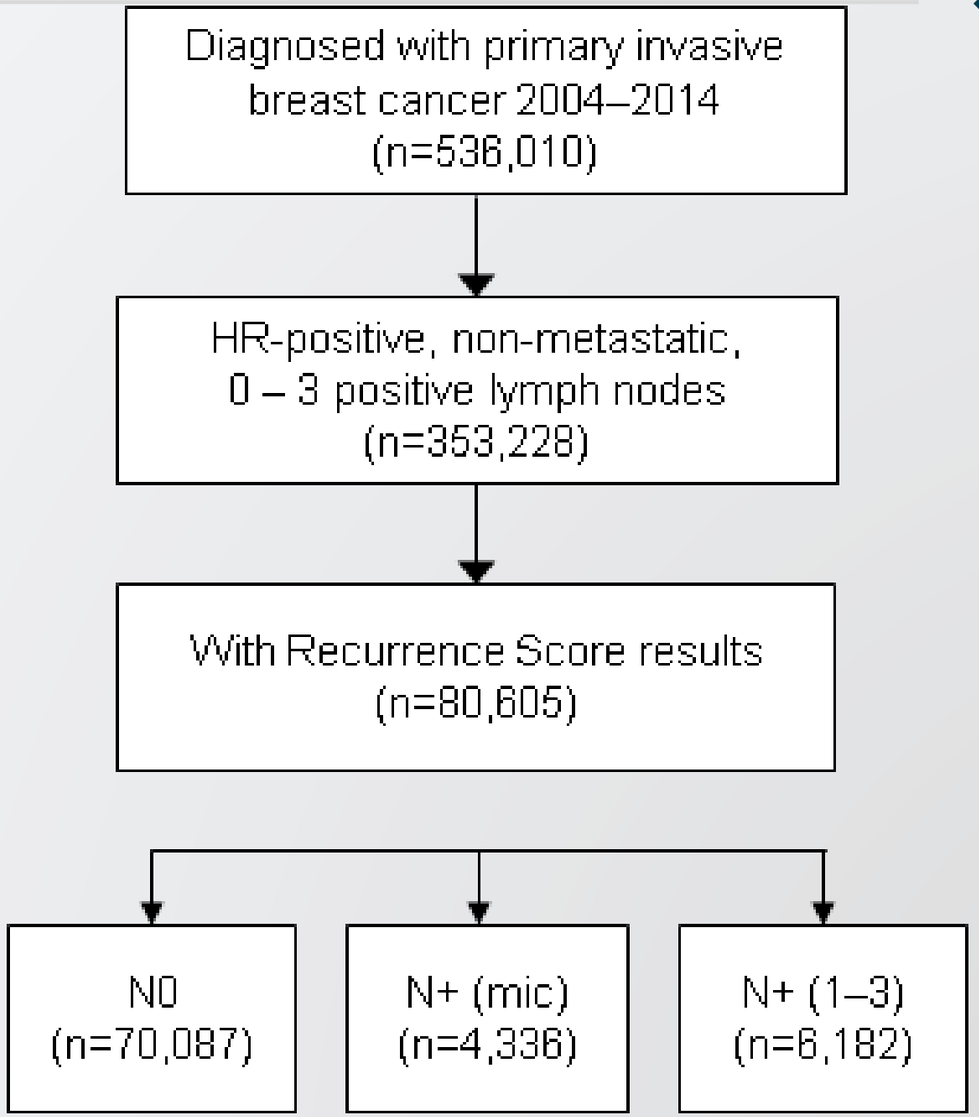
Breast Pathologists
Mary Edgerton
Uma Krishnamurti
Kate Serdy

Breast Oncologist
Allison Kurian

# Real World Evidence: Example #2

Chemotherapy benefit in Oncotype DX Breast Recurrence Score® (RS)-tested patients w/ N0 disease

Prognosis in RS-tested patients w/ N0, N1mic, & N1 disease treated without adjuvant chemotherapy

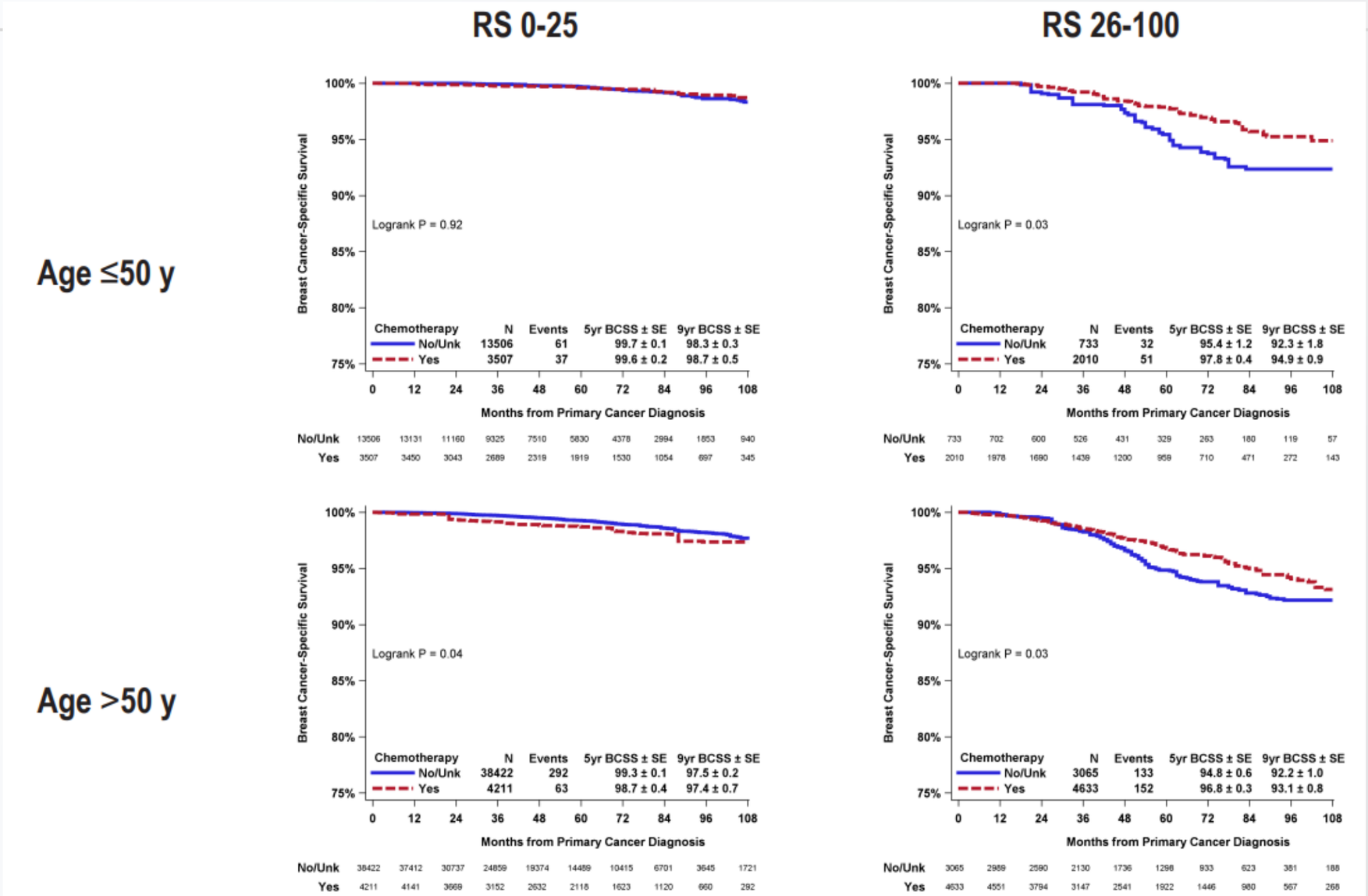



Valentina Petkov, MD, MPH

Gabriel Hortobagyi, MD

Hortobagyi et al. San Antonio Breast Cancer Symposium (December 2018) Abstract P3-11-02

# SEER-Genomic Population-Based Findings




Valentina Petkov, MD, MPH

Gabriel Hortobagyi, MD

**Hortobagyi et al. San Antonio Breast Cancer Symposium (December 2018) Abstract P3-11-02**

# SEER-Genomic Population-Based Findings

- RS predictive of chemotherapy benefit in patients with HR+ N0 disease & RS 26-100, supporting the cutoff at RS 26 for chemotherapy benefit
- 9-year BCSS >97% without chemotherapy in patients w/ RS <18 regardless of nodal status
- Insufficient events at analysis to estimate chemotherapy benefit among women with LN-positive disease

**Hortobagyi et al. San Antonio Breast Cancer Symposium (December 2018) Abstract P3-11-02**

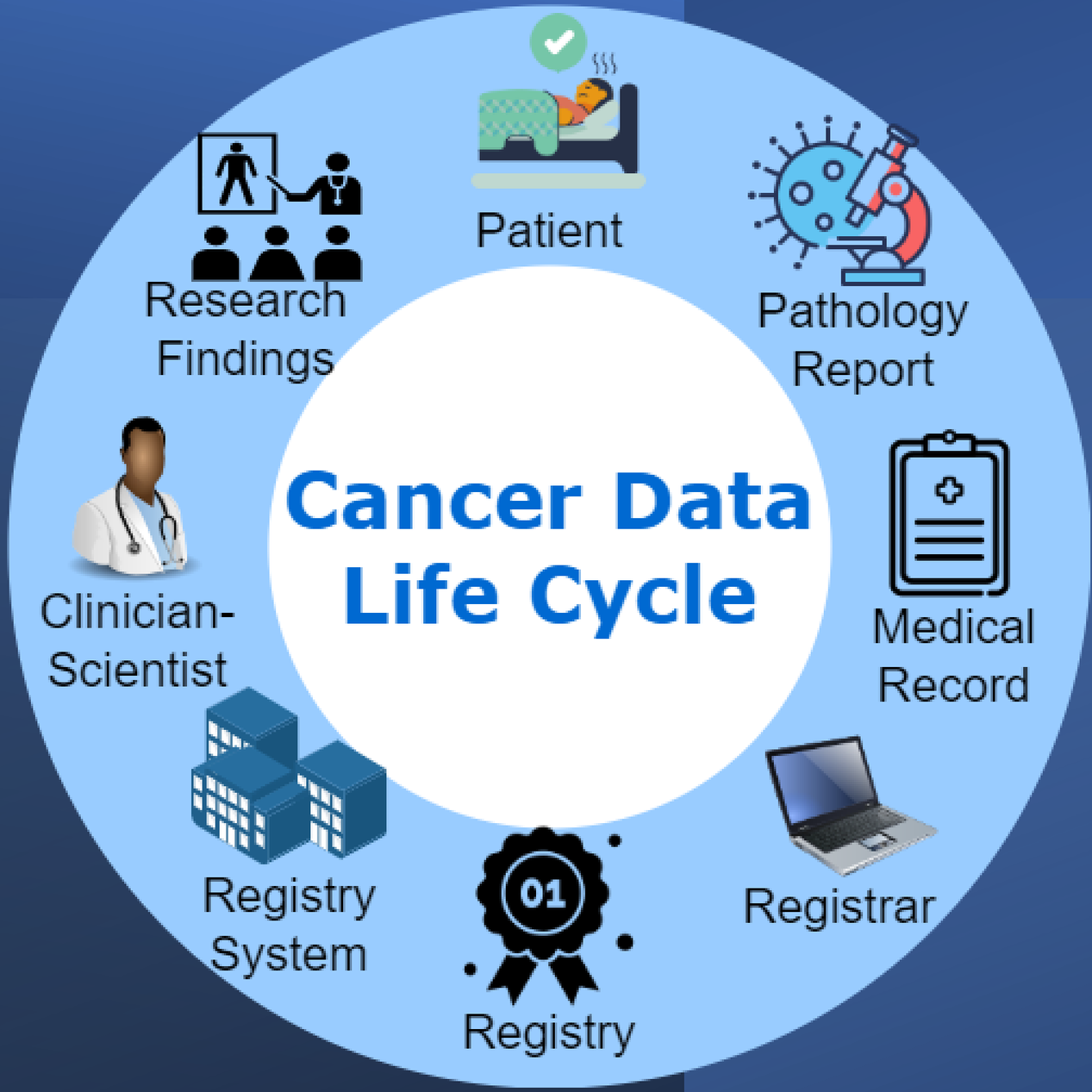
Valentina Petkov, MD, MPH

Gabriel Hortobagyi, MD

# Reference for RWE Example #2

Hortobagyi, G. N. et al. “Breast cancer-specific mortality (BCSM) in patients (pts) with  node-negative (N0) and node-positive (N plus) breast cancer (BC) guided by the 21-gene assay: A SEER-genomic population-based study.” *Cancer Research* 2019;79(4S): Meeting Abstract P3-11-02.

# New & Future Data Resources



# Data Linkages

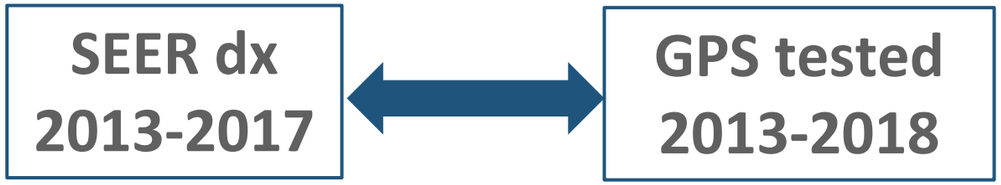
## Why link to genomic/germline testing data?

- More ***efficient*** way for ***data collection*** by centralizing data acquisition the Honest Broker b/n SEER registries and industry
- ***Difficulties*** in training registrars ***in coding*** genomic/genetic data due to complicated, rapidly changing clinical practice
- ***Assure completeness and quality*** of data
- ***Case finding*** source, especially for cancer patients diagnosed and treated at community specialty practices

# Exact Sciences Linkage

## Oncotype DX Genomic Prostate Score (GPS)

- Recommended in guidelines for treatment decisions & prediction of adverse pathology, on market since 2013
- 1st time linkage with this test




- Case finding study (~20% of tested cases with no matching in SEER)

## Oncotype DX IBC

- 4th linkage: 2004-2017
- Linkage methodological improvements

## Oncotype DX DCIS

**Establishing data release process for specialized database**

# Virtual Pooled Registry Cancer Linkage System VPR-CLS

NCI/SEER-funded & NAACCR-managed

Launched in February 2022

Linkages between research studies & U.S. registries

- Single linkage software & standard matching criteria
- Aggregate match counts to inform selection of registries for requests

Minimize burden & cost to researchers, registries, & IRBs; Increase ease of access and timely use of registry data

# Virtual Pooled Registry Cancer Linkage System VPR-CLS

45 participating registries (95% of U.S. population)

22 study linkages with 6 initial pilot test studies

1.1M matches among over 13.5M cohort members

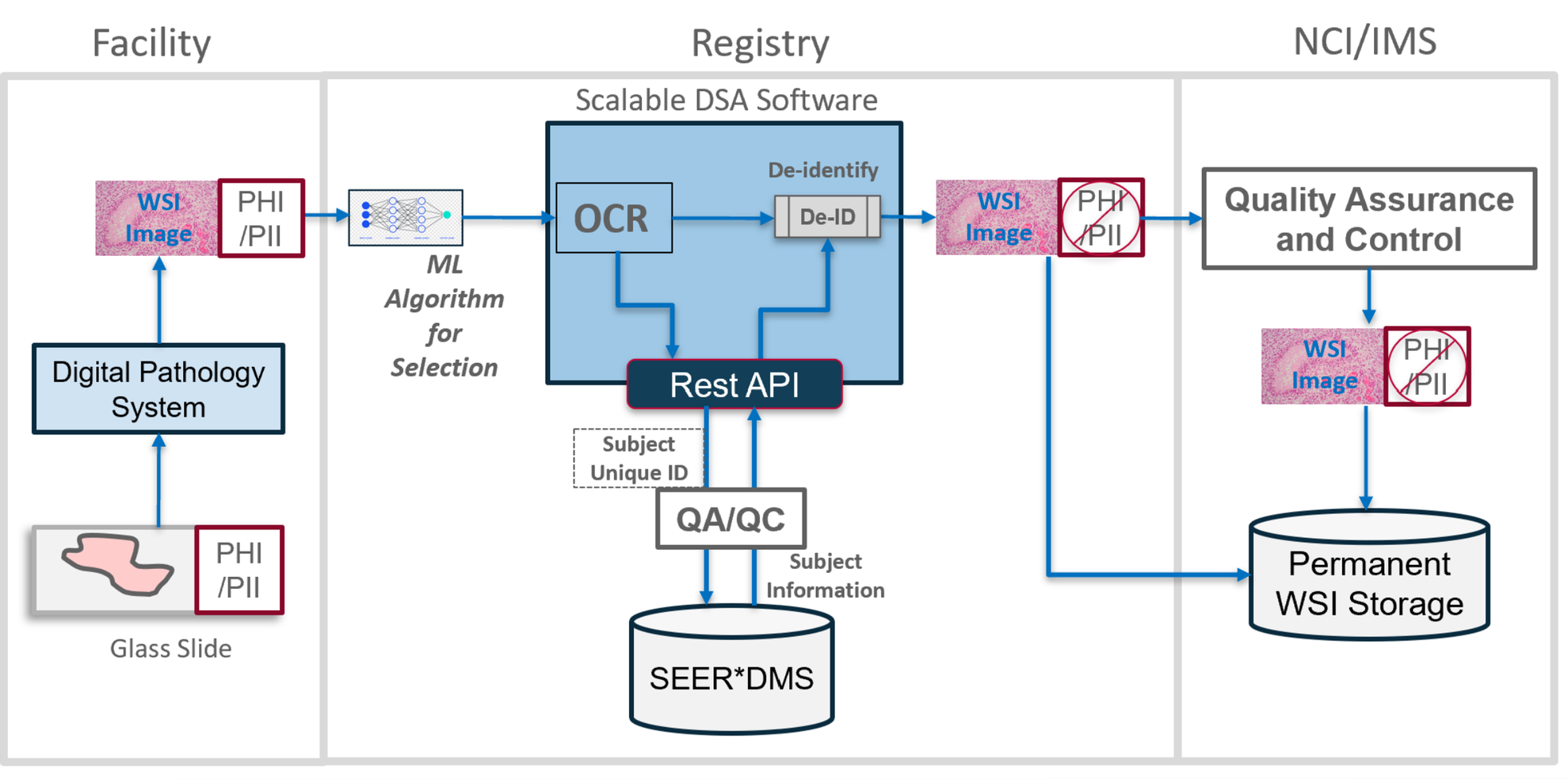
Fact sheets and webinars regarding the Common Rule changes & secondary data sharing: https://www.naaccr.org/vpr-fact-sheets/

VPR Program Manager: Castine Clerkin, cclerkin@naaccr.org

# Data Linkages in NCCR Data Platform at 2024 Launch

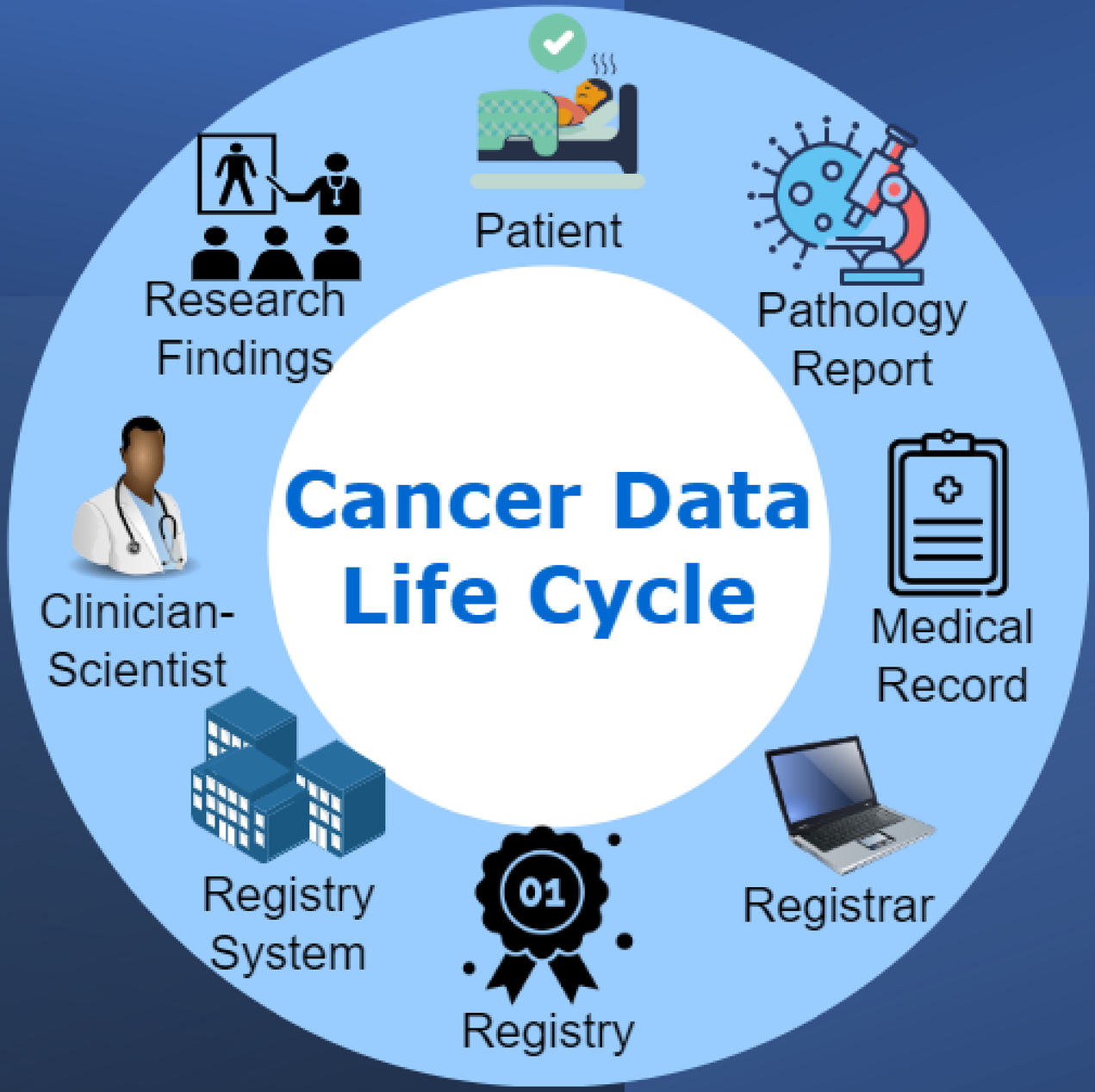
- Registry-abstracted data
- Social Determinants of Health
- Results of data linkages
  - Children's Oncology Group (COG)
  - Pediatric Proton/Photon Consortium Registry (PPCR)
  - Virtual Pooled Registry (VPR)
  - Medical and Pharmacy claims from multiple data sources

# Deidentified WSIs Linked to Data in Future



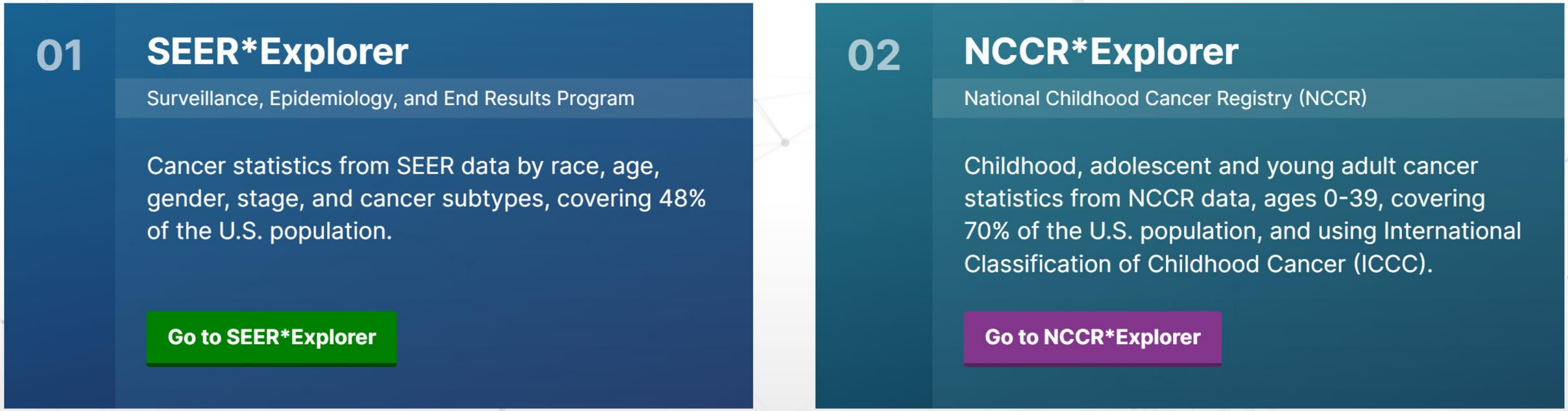

https://digitalslidearchive.github.io/DSA-WSI-DeID/

# Cancer Surveillance Data Access



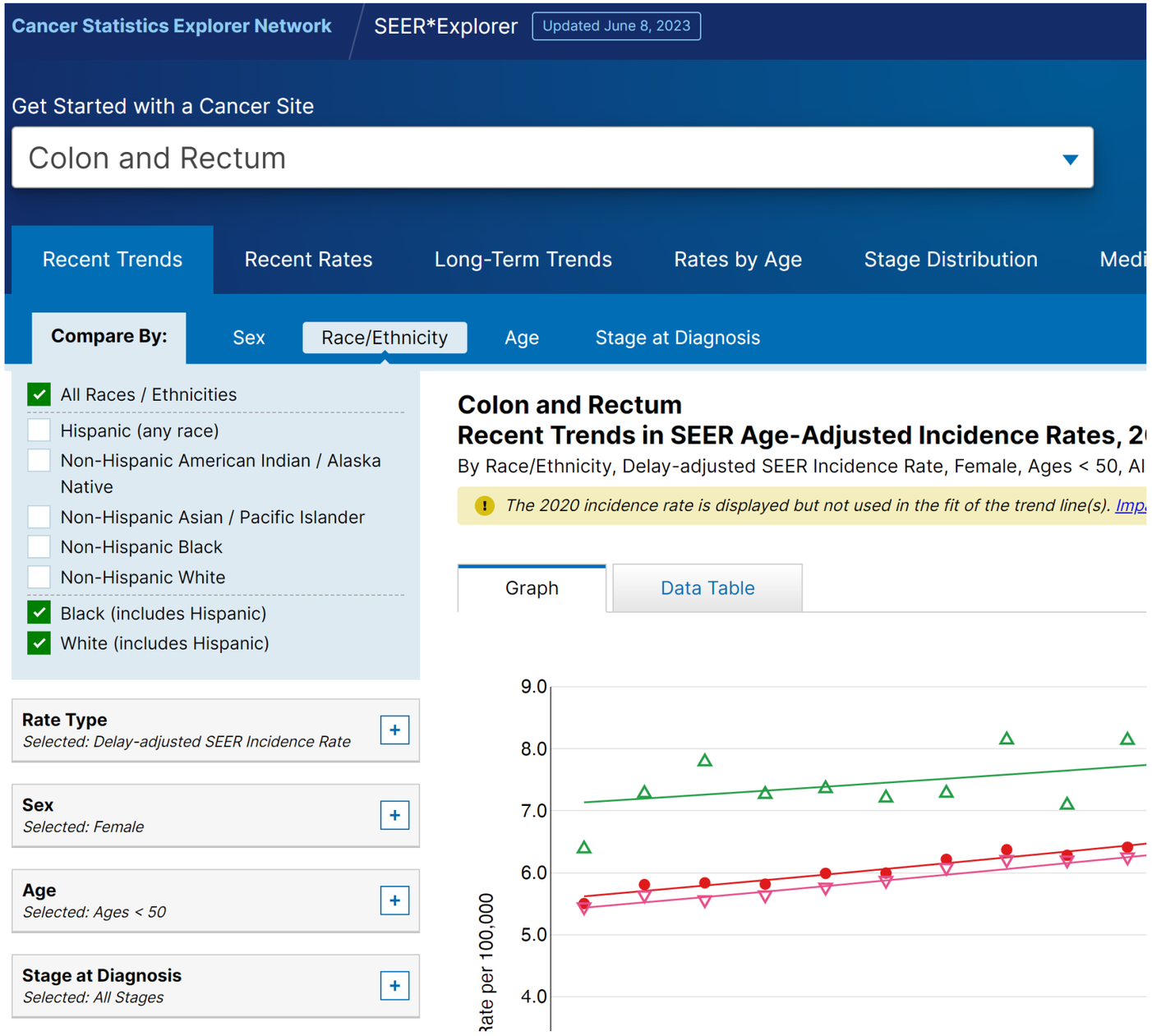
# Welcome to the Cancer Statistics Explorer Network

**Find the cancer statistics tool best suited to your needs.**




**https://seer.cancer.gov/statistics-network/**

# SEER*Explorer



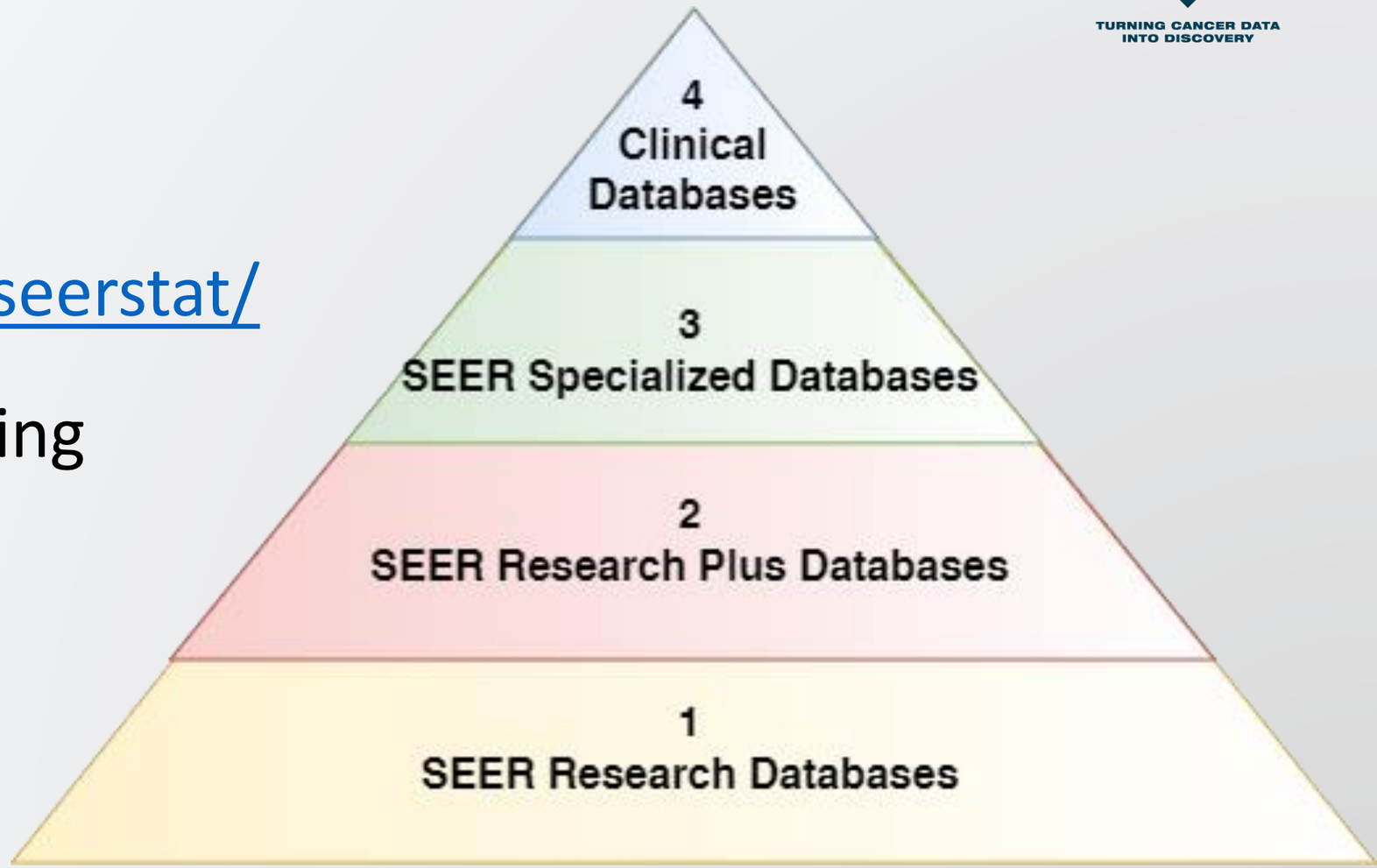
# SEER Tiered Data Release

- Current data release via SEER*Stat

  https://seer.cancer.gov/seerstat/

- Tier 4 data resources being developed

- Future plans for data aggregation & access

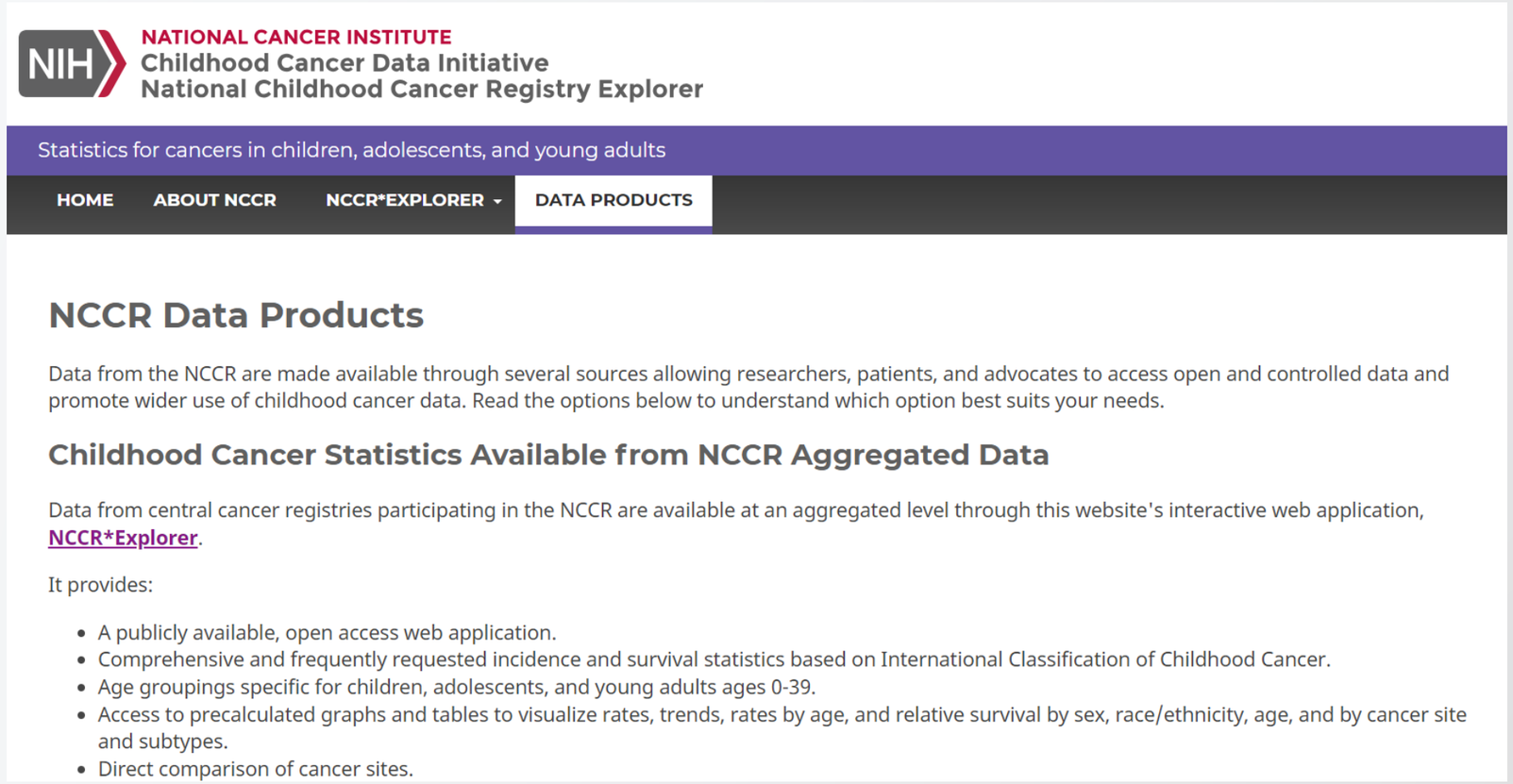


# Information about NCCR Data Products

**Open Data Access**
NCCR*Explorer

**Registered Data Access**
NCCR data in SEER*Stat

**Registered & Controlled Data Access**
NCCR Data Platform




https://nccrexplorer.ccdi.cancer.gov/data-products.html

# Official Federal Cancer Statistics

## U.S. Cancer Statistics

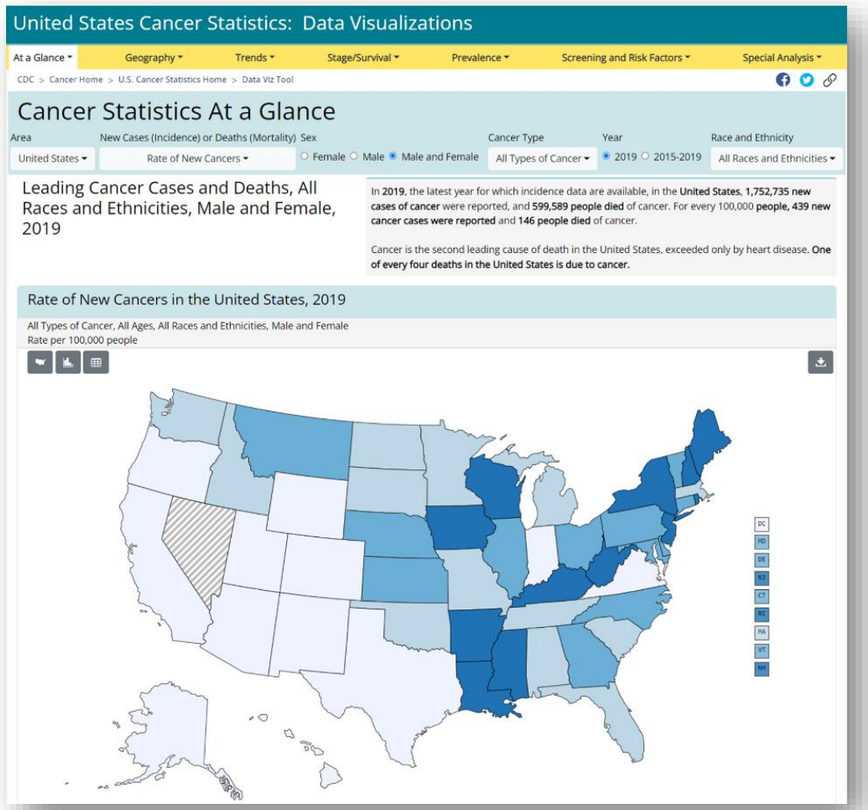
NATIONAL PROGRAM of CANCER REGISTRIES
NPCR
1992 30 YEARS 2022

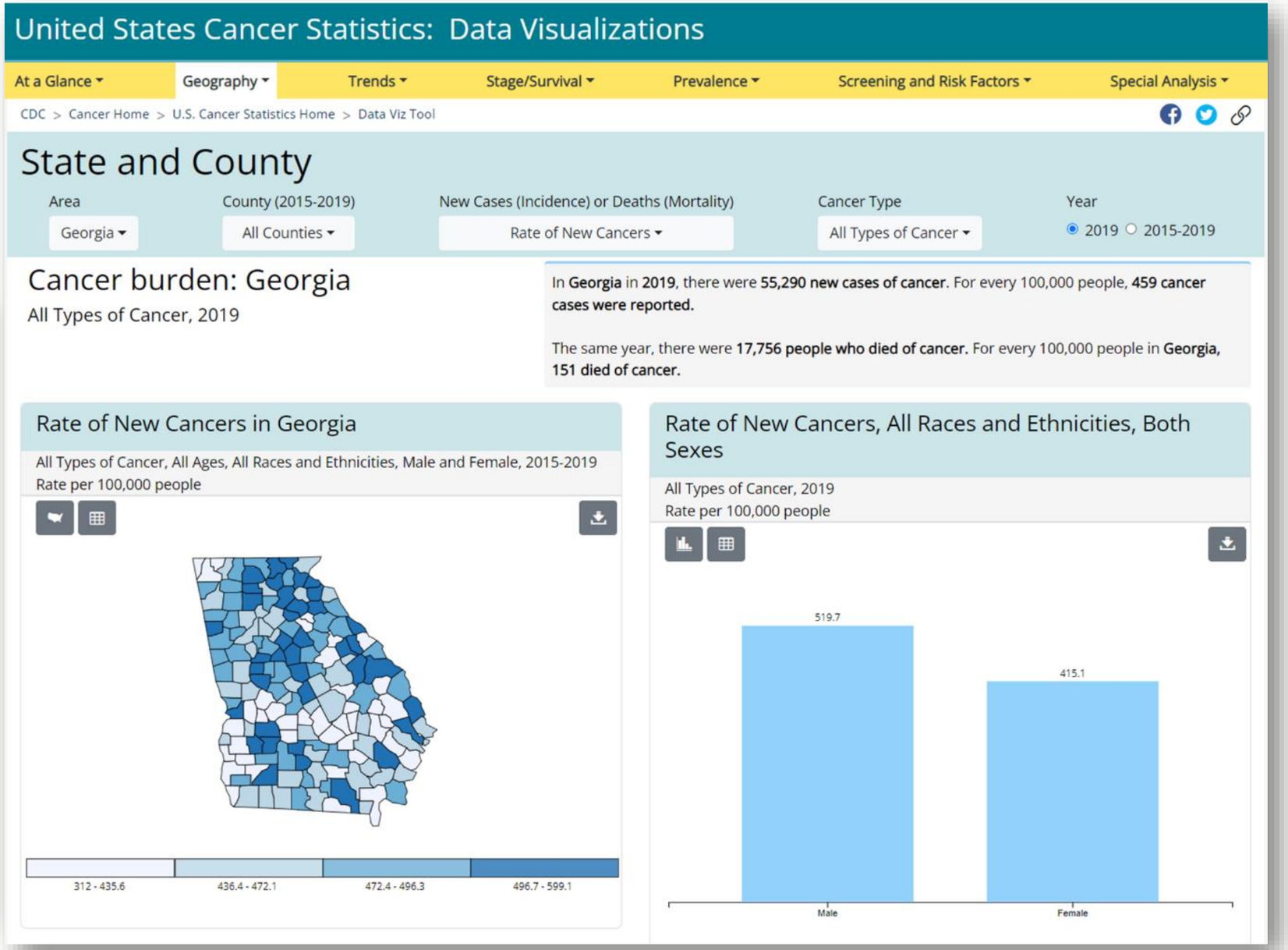
50TH ANNIVERSARY SEER
TURNING CANCER DATA INTO DISCOVERY


**www.cdc.gov/uscs**

# Data Visualizations Tool

## U.S. Cancer Statistics




**www.cdc.gov/uscs/dataviz**

# Public Use Databases

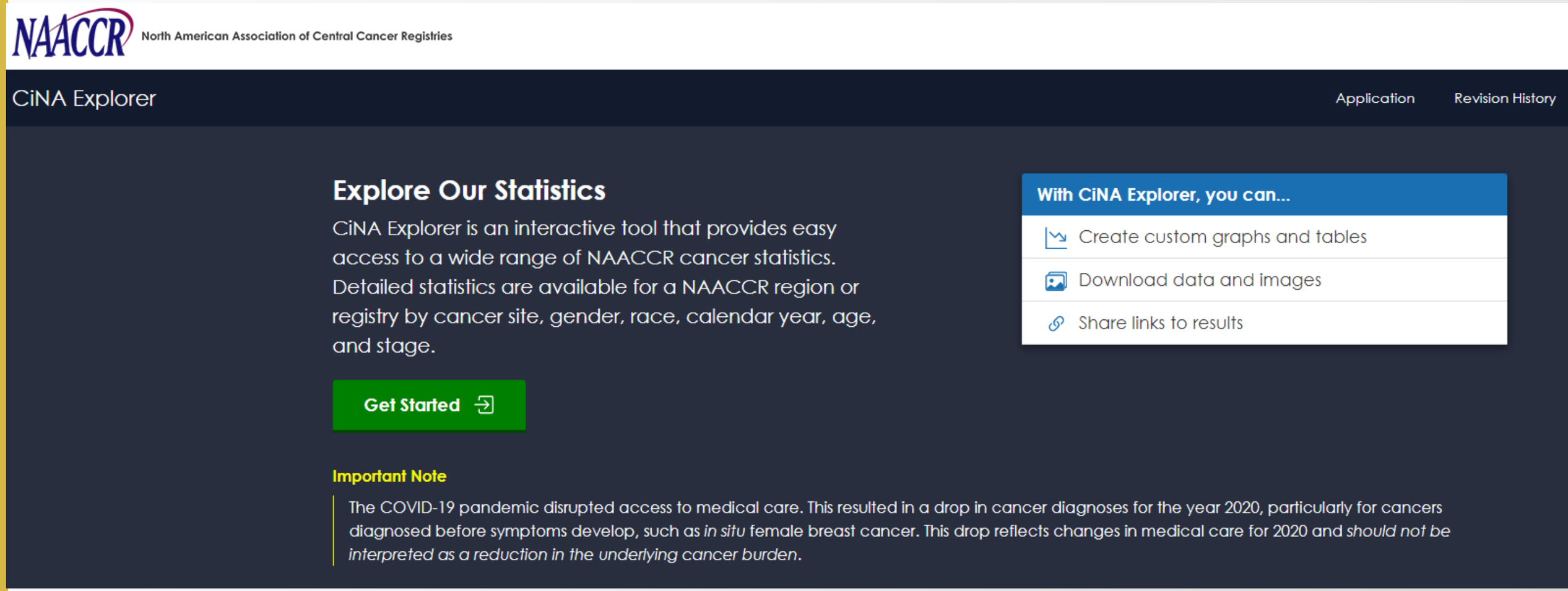
## U.S. Cancer Statistics

- Demographics data
  - age, sex, race, ethnicity, state
- Tumor identification
  - primary site, histology, grade, behavior, stage


**www.cdc.gov/uscs/public-use**

# NAACCR Online CiNA+ Interactive Tools

https://apps.naaccr.org/explorer/



# ols

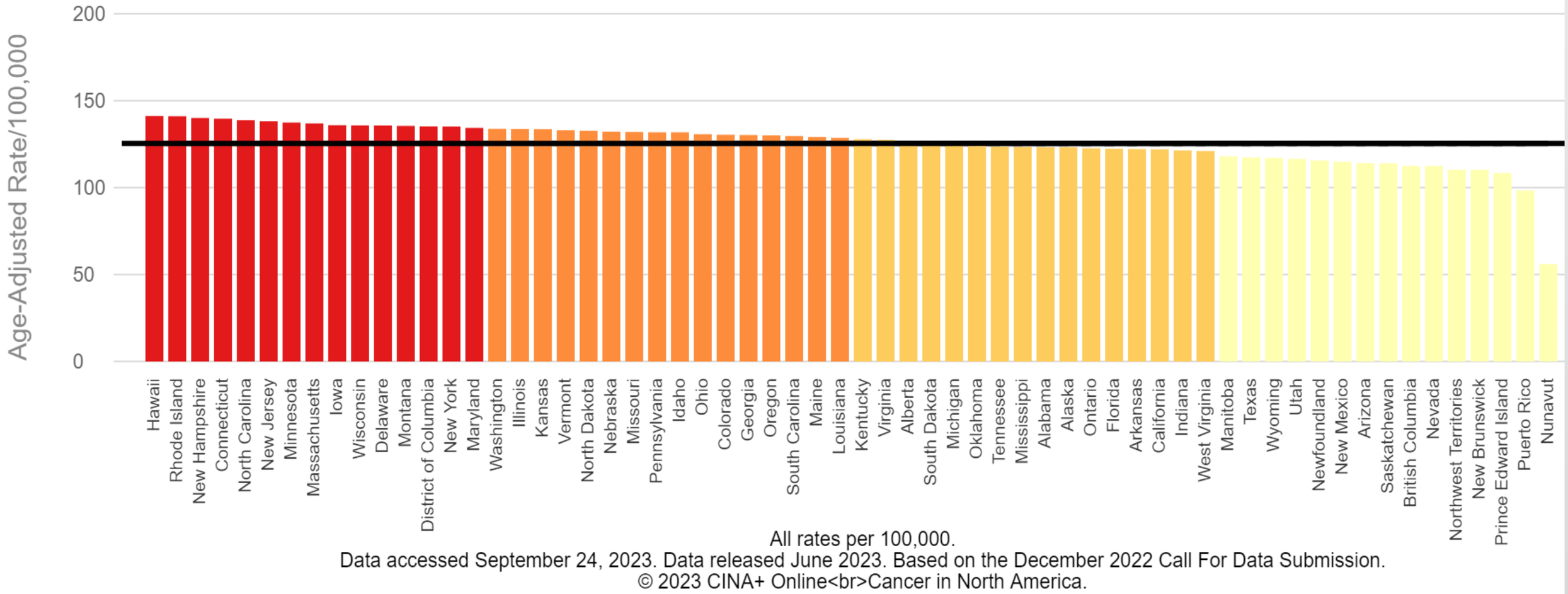
Age-Adjusted Invasive Cancer Incidence Rates in North America

Breast, Female, 2016 - 2020

By State/Province

Age-Adjusted to the 2000 U.S. Standard Population

North America Rate: 125.92 / per 100,000




All rates per 100,000.

Data accessed September 24, 2023. Data released June 2023. Based on the December 2022 Call For Data Submission.

# NAACCR Data Access

| NAACCR Data Product | Requirements | For More Information |
|---|---|---|
| CiNA Public Use Dataset | Signed DUA | https://www.naaccr.org/cina-public-use-data-set/ |
| CiNA Research Dataset | NAACCR member as PI/Co-PI | https://www.naaccr.org/cina-research/ |
| CiNA Survival & Prevalence Data | | https://www.naaccr.org/cina-survival/ |
| CiNA Special Dataset Request | | Contact NAACCR Program Manager of Data Use & Research (rsherman@naaccr.org) |

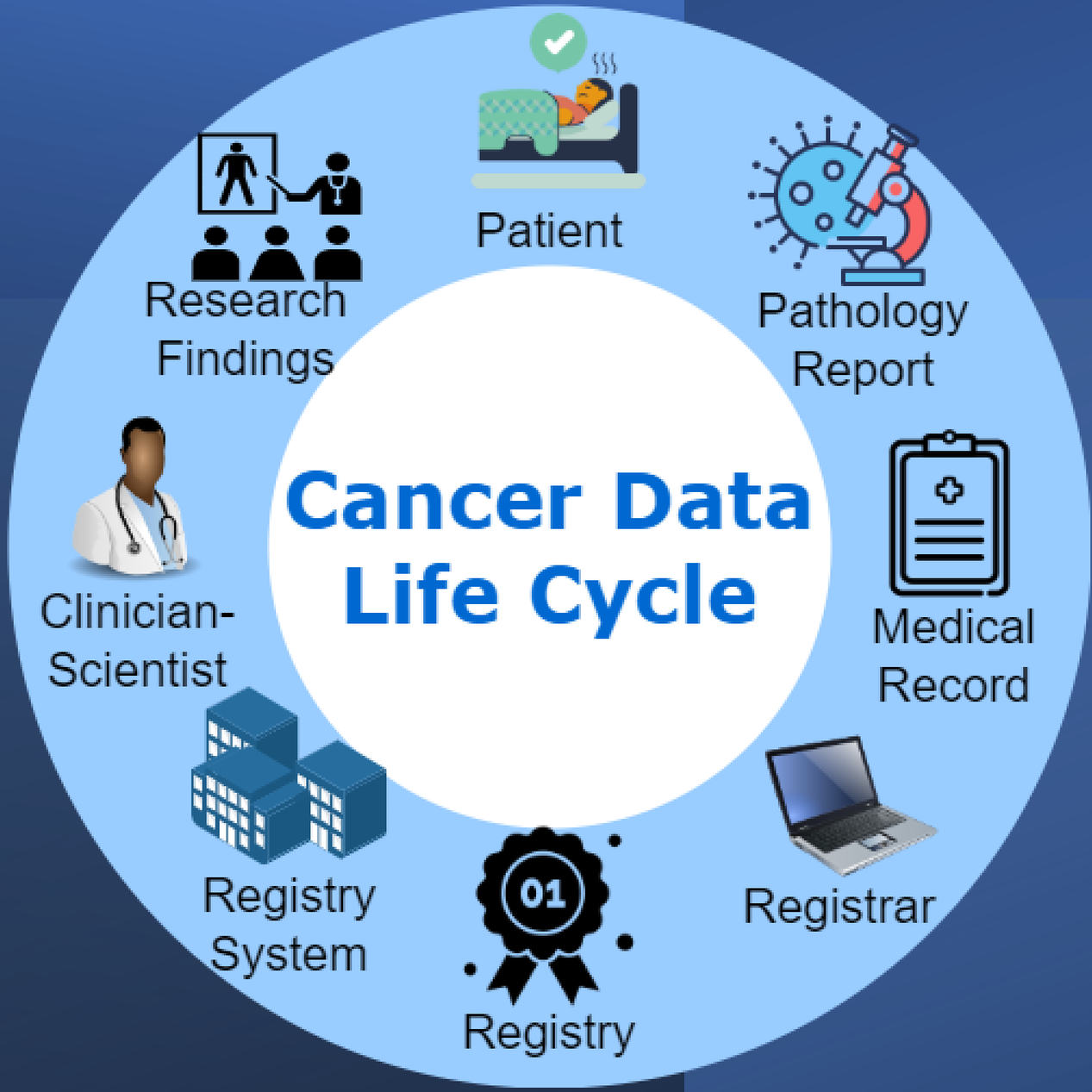
- Released via SEER*Stat: https://seer.cancer.gov/seerstat/
- Need collaborator who has experience working in SEER*Stat

# Accessing U.S. Population-Based Cancer Registry Data

| WHO | NCI - SEER | CDC & NCI - USCS | NAACCR |
|---|---|---|---|
| WHAT | NCI - SEER data | CDC-NPCR & NCI-SEER data | U.S. & Canadian registry data |
| WHERE | seer.cancer.gov/data-software/ | cdc.gov/uscs | apps.naaccr.org/dart/ |
| EXPLORE | seer.cancer.gov/statistics/interactive.html | cdc.gov/cancer/dataviz | apps.naaccr.org/explorer/ |
| HOW | **seer.cancer.gov/data/product-comparison.html**<br><br>**Research Data**<br>For those without eRA Commons or HHS accounts<br><br>**Research Plus and NCCR**<br>1. Access Data Request System (eRA commons or HHS account) & complete app<br>2. Acknowledge SEER DUA, SEER Treatment Data Limitations, Best Practices Assurance, & NCCR DUA<br>3. Gain SEER*Stat access | **cdc.gov/cancer/public-use**<br><br>**Public Use Databases**<br>1. Gain access to SEER Research Plus data<br>2. Sign USCS Research Data Agreement<br>3. Email to uscsdata@imsweb.com | **naaccr.org/cina-data-products-overview/**<br><br>**CiNA Public Use Data**<br>(counts, rates, & trends)<br>1. Request via DaRT<br>2. Sign Data Assurance Agreement<br><br>**CiNA Research Datasets**<br>(counts, rates, trends, & prevalence)<br>1. Identify NAACCR member as PI or Co-PI<br>2. Request via DaRT |

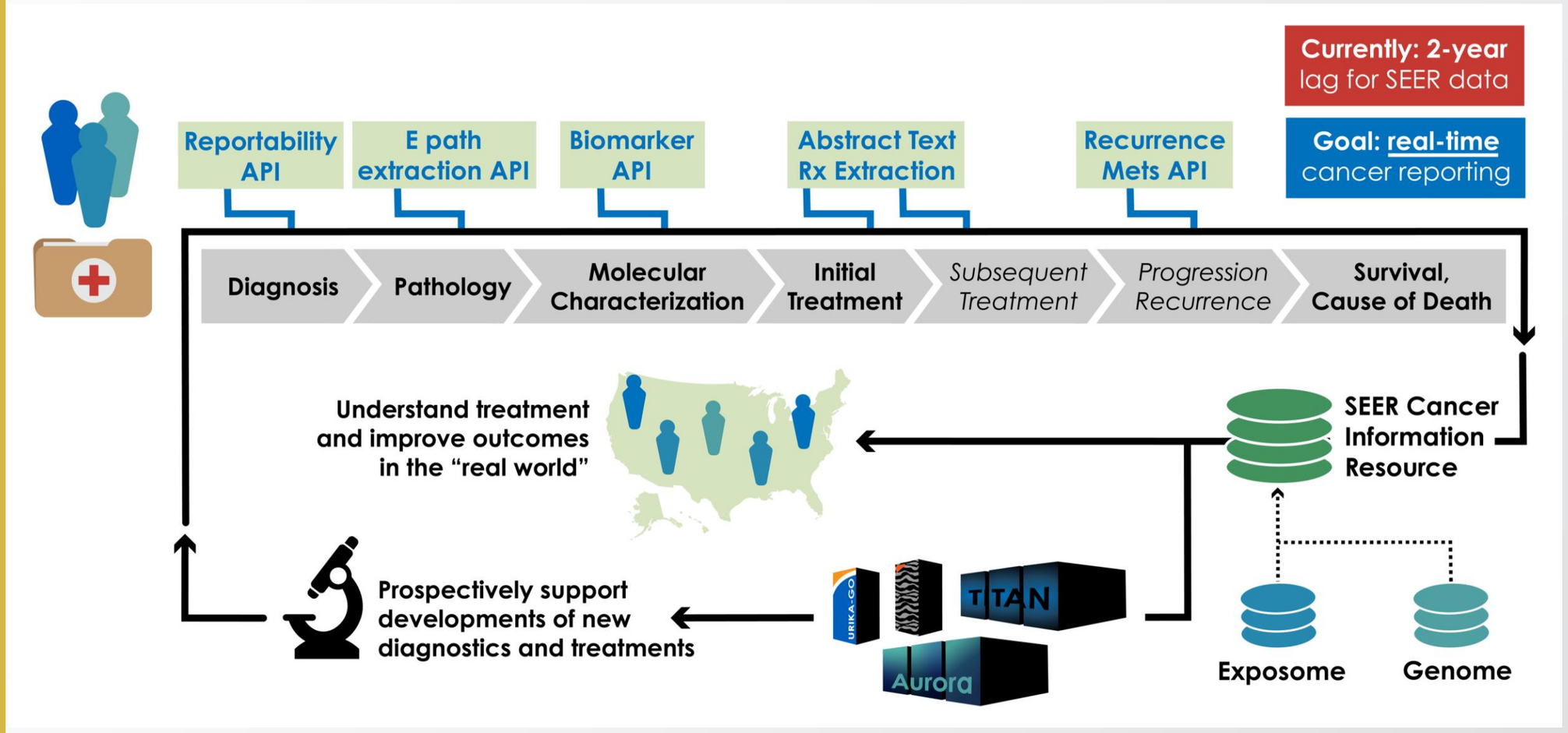
**Currently accessible via SEER*Stat**
**seer.cancer.gov/seerstat/**

# AI Tools in Cancer Surveillance



# The MOSSAIC Challenge

**Translational AI for better cancer surveillance & ultimately better cancer care.**




OAK RIDGE National Laboratory

50TH ANNIVERSARY SEER

TURNING CANCER DATA INTO DISCOVERY

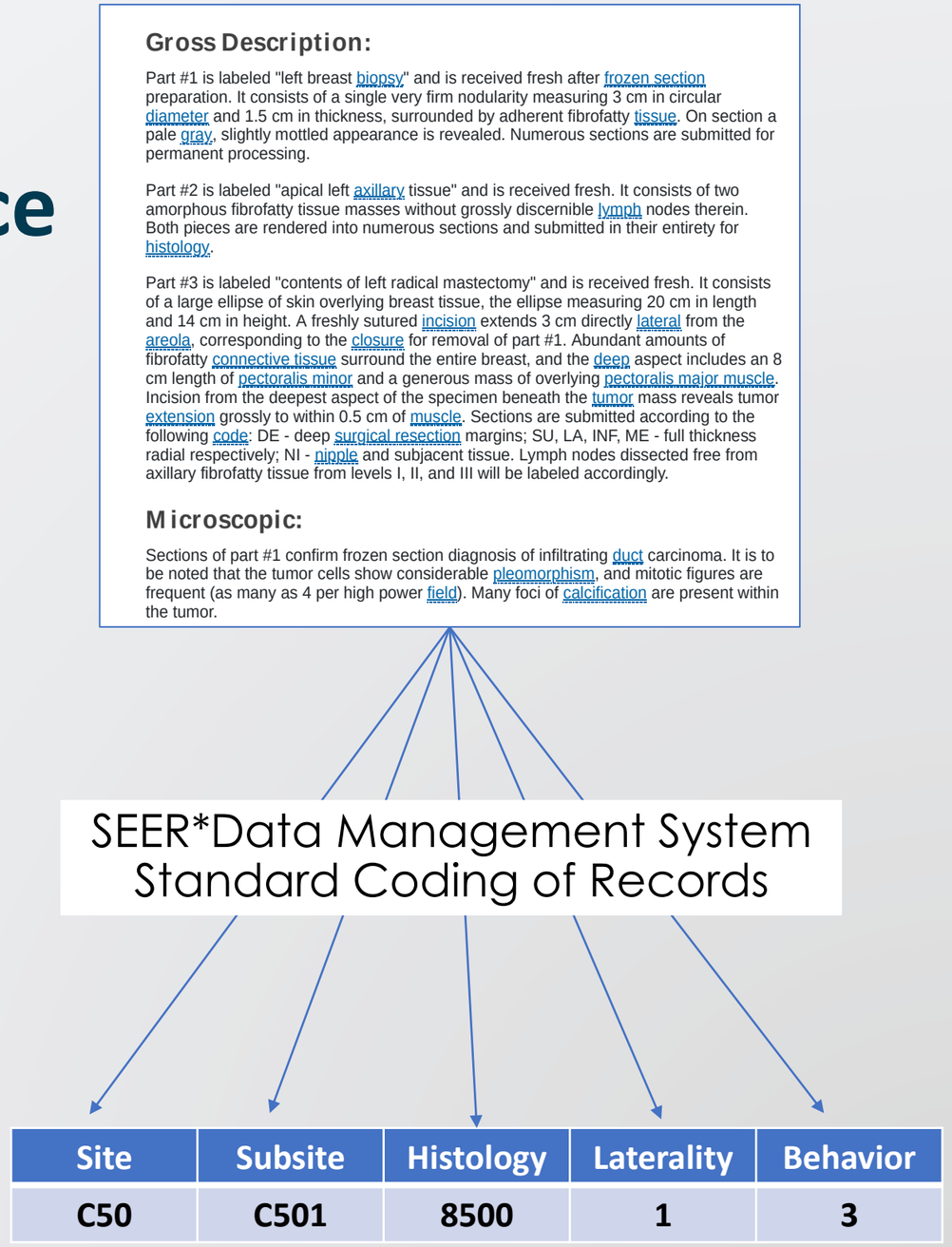
# Automating Common Data Models Deep Learning for Health Surveillance

## Auto-Extraction from Pathology Reports

Accuracy: **23-27% of path reports** with >98 accuracy across all data elements.

Auto-coding **performance can be easily tuned**

**Efficiency:** saves ~14,000 person-hours/year



# Take Home Messages

Pathologists & registrars as partners to ensure data quality

AI as an aid for pathologists & registrars instead of a replacement

Population-based data lead to more accurate answers to questions


NCICancerPathCHART@mail.nih.gov

https://seer.cancer.gov/cancerpathchart/

NIH NATIONAL CANCER INSTITUTE

# The Future of Cancer Data:
# Unlocking Insights With Pathology Reporting Summit

October 6, 2023