NATIONAL CANCER INSTITUTE
Division of Cancer Control & Population Sciences

Pathology report to cancer research: ensuring data quality in the AI era

Alison Van Dyke, MD, PhD, FCAP

October 18, 2024

Disclosures & Funding

I have no conflicts of interest to disclose

This work is funded &/or performed by the

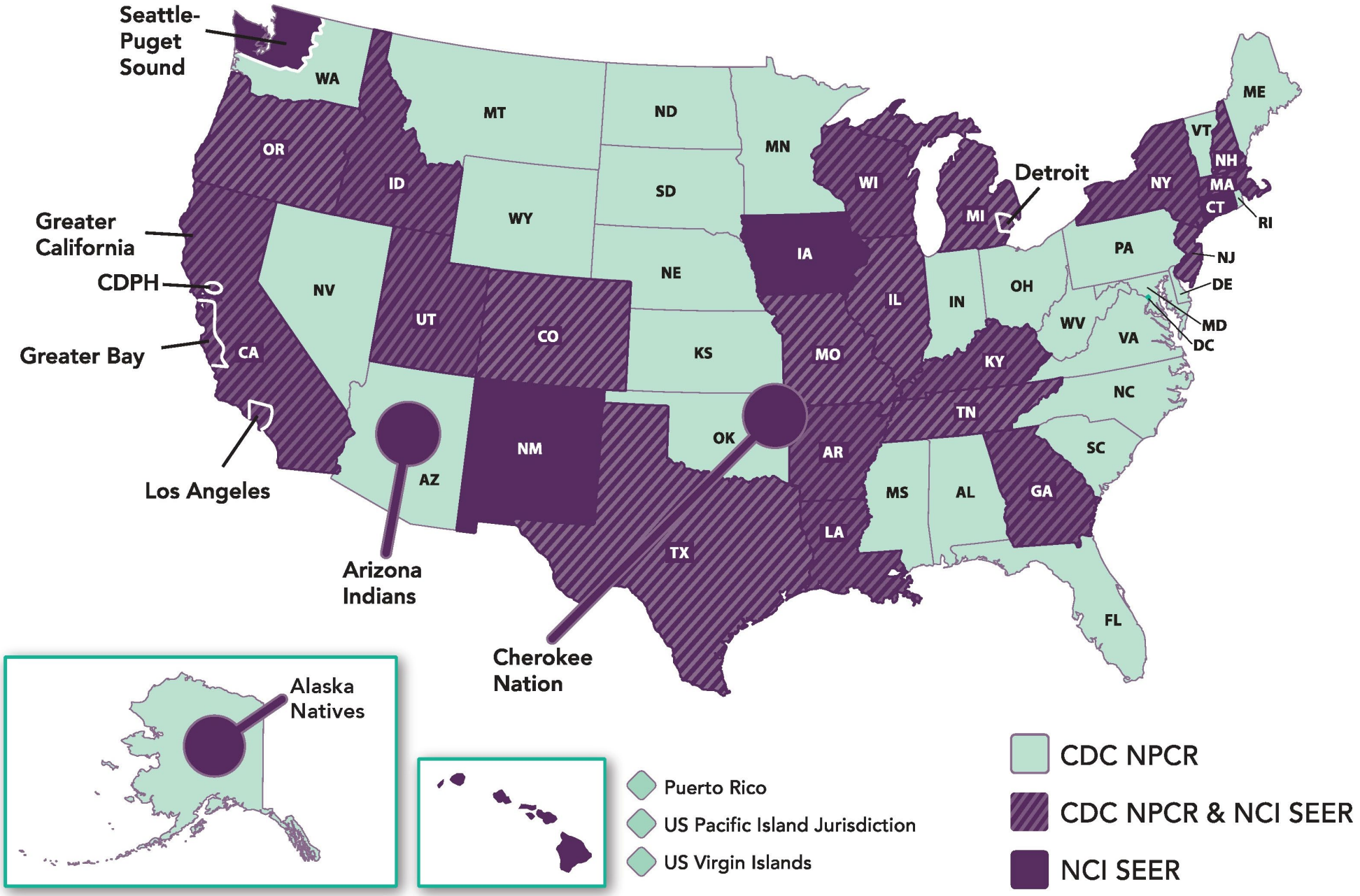
- Surveillance Research Program in the Division of Cancer Control and Population Sciences of the National Cancer Institute, National Institutes of Health
- Cancer Moonshot℠
- US Department of Energy Oak Ridge National Laboratory

Topics Exploring

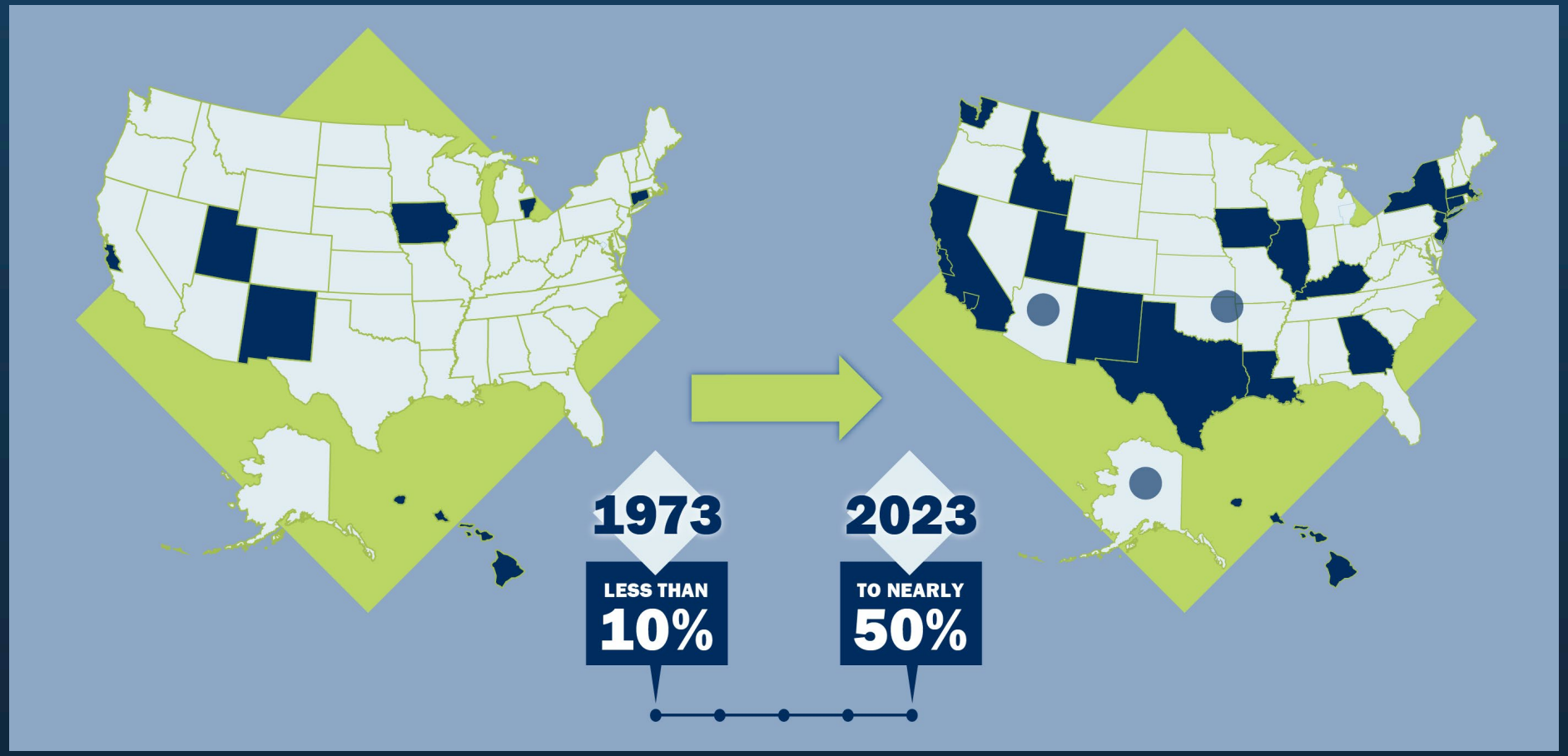
- Cancer surveillance system acquisition of pathology report data items
- Data quality assessments by data acquisition method
- How CAP is improving the foundation of cancer surveillance data quality
- Data quality standards & automated data capture
- Future of cancer surveillance data capture & reporting

Cancer surveillance systems & data acquisition methods

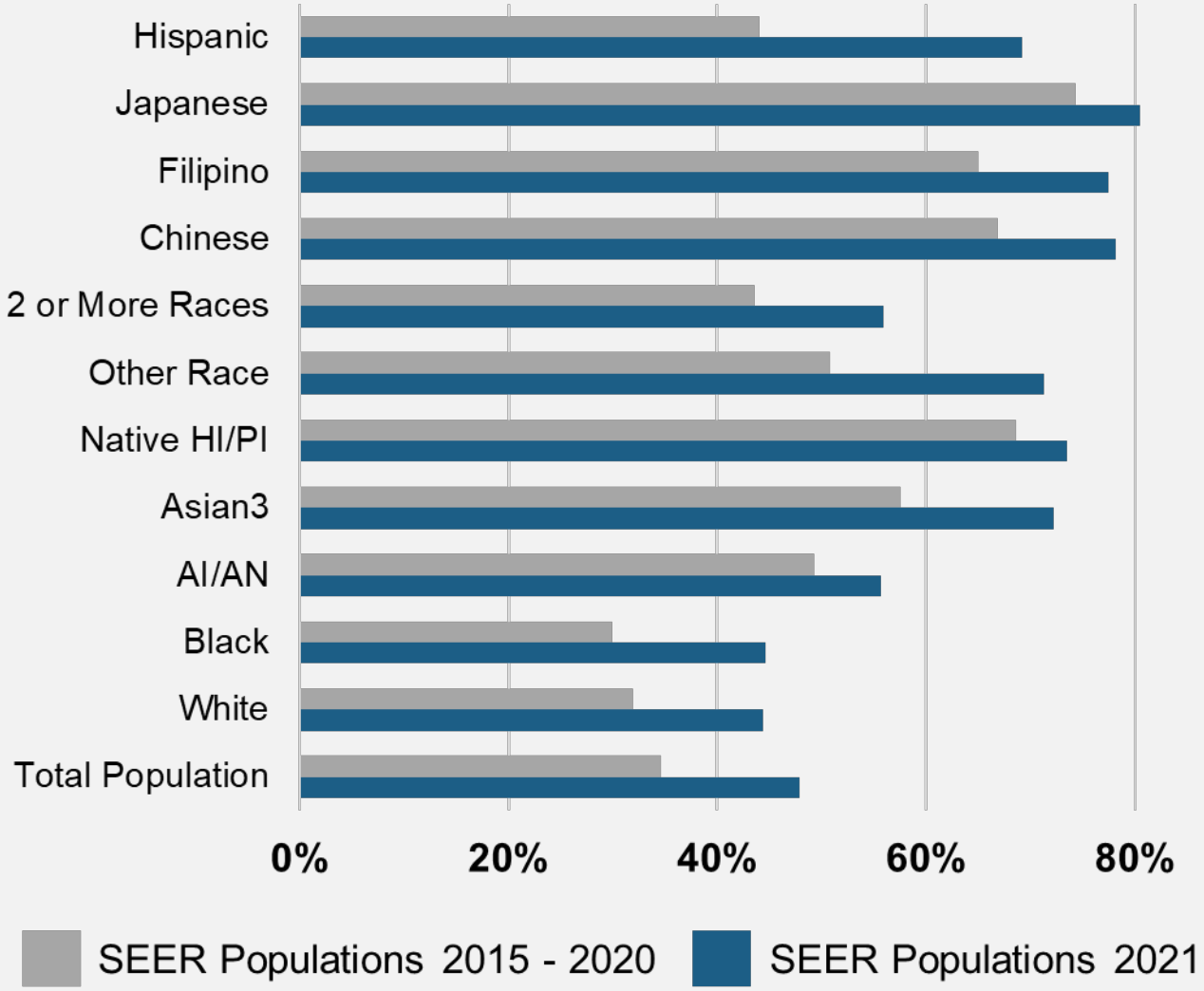
Federally-funded Cancer Registries Cover 100% of the US Population



Surveillance, Epidemiology, and End Results (SEER)



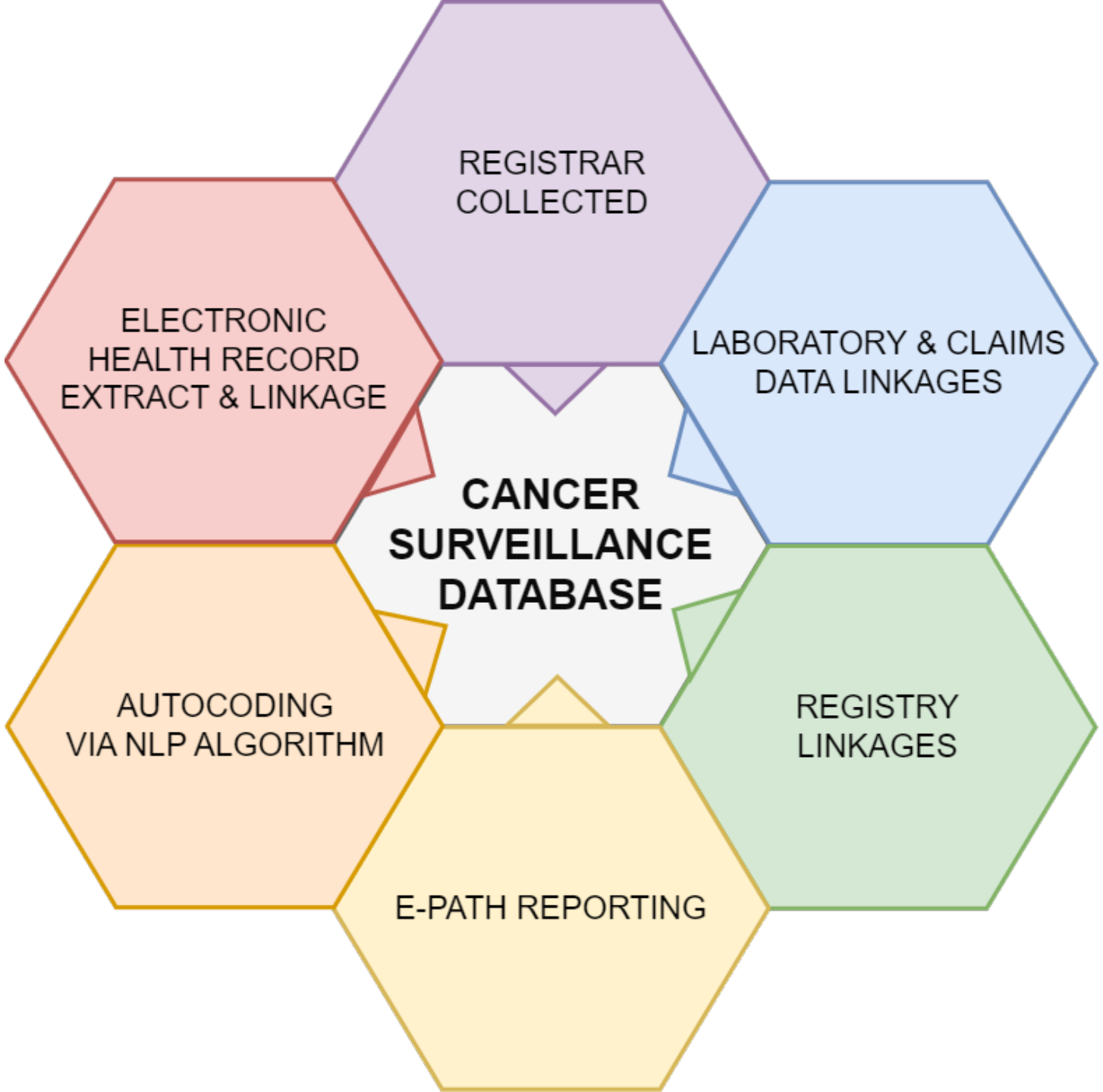
Increase in Representation of Population Subgroups with SEER Expansion

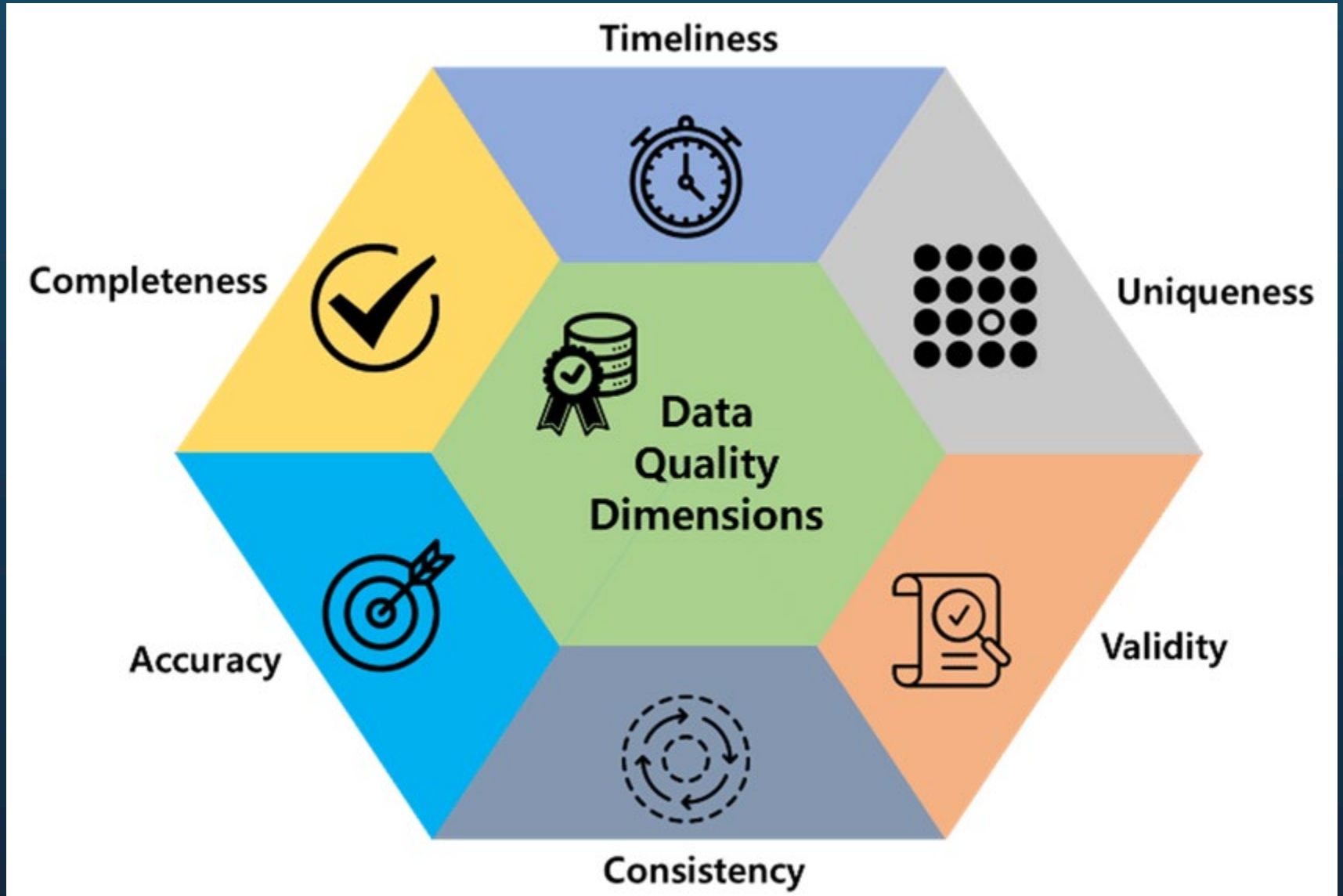


Percent increase for US population subgroups with SEER expansion 2021

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

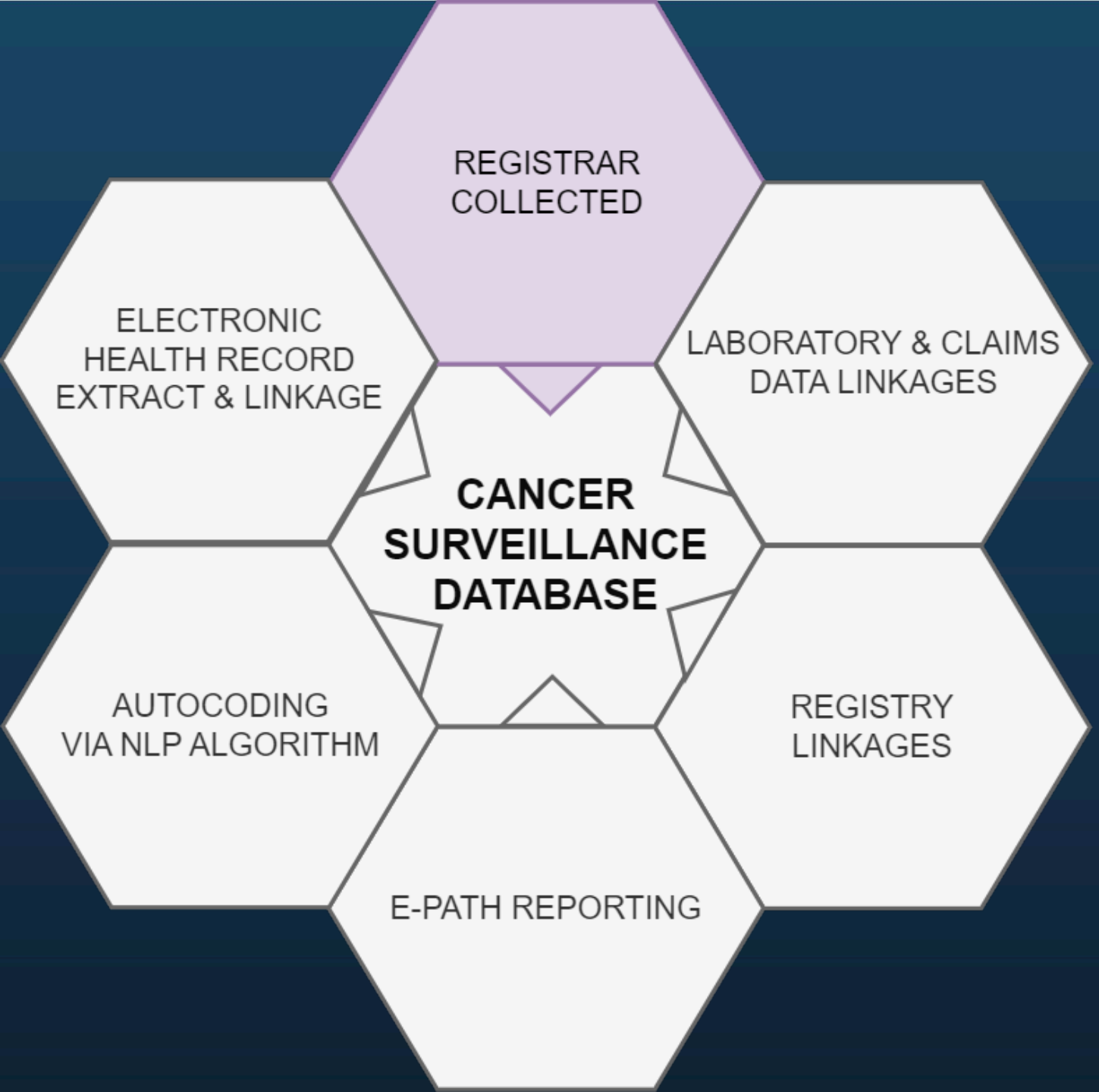
Mechanisms of Cancer Surveillance Data Acquisition





Citations: Bhandari, S.; Ranjan, N.; Kim, Y.-C.; Park, J.-D.; Hwang, K.-I.; Kim, W.-H.; Hong, Y.-S.; Kim, H. An Automatic Data Completeness Check Framework for Open Government Data. Appl. Sci. 2021, 11, 9270. https://doi.org/10.3390/app11199270. Accessed October 8, 2024.

Van Nederpelt, P. & Black, A. May 9, 2020. "Dimensions of Data Quality (DDQ) Research Paper" DAMA NL Foundation. https://www.dama-nl.org/wp-content/uploads/2020/09/DDQ-Dimensions-of-Data-Quality-Research-Paper-version-1.2-d.d.-3-Sept-2020.pdf. Accessed October 8, 2024.



Medical Record

Image: American Institute for Economic Research n.d. "The Workings of the Gold Standard" https://aier.org/article/the-workings-of-the-gold-standard/. Accessed October 13, 2024.

Oncology Data Specialists capture & maintain

ODS
USE BEGINS ON JANUARY 1, 2024

- Demographics
- Site & Morphology
- Stage at Diagnosis
- Grade
- Biomarkers
- Treatments
- Sequencing & Personalized Medicine
- Long-term follow up & Recurrence

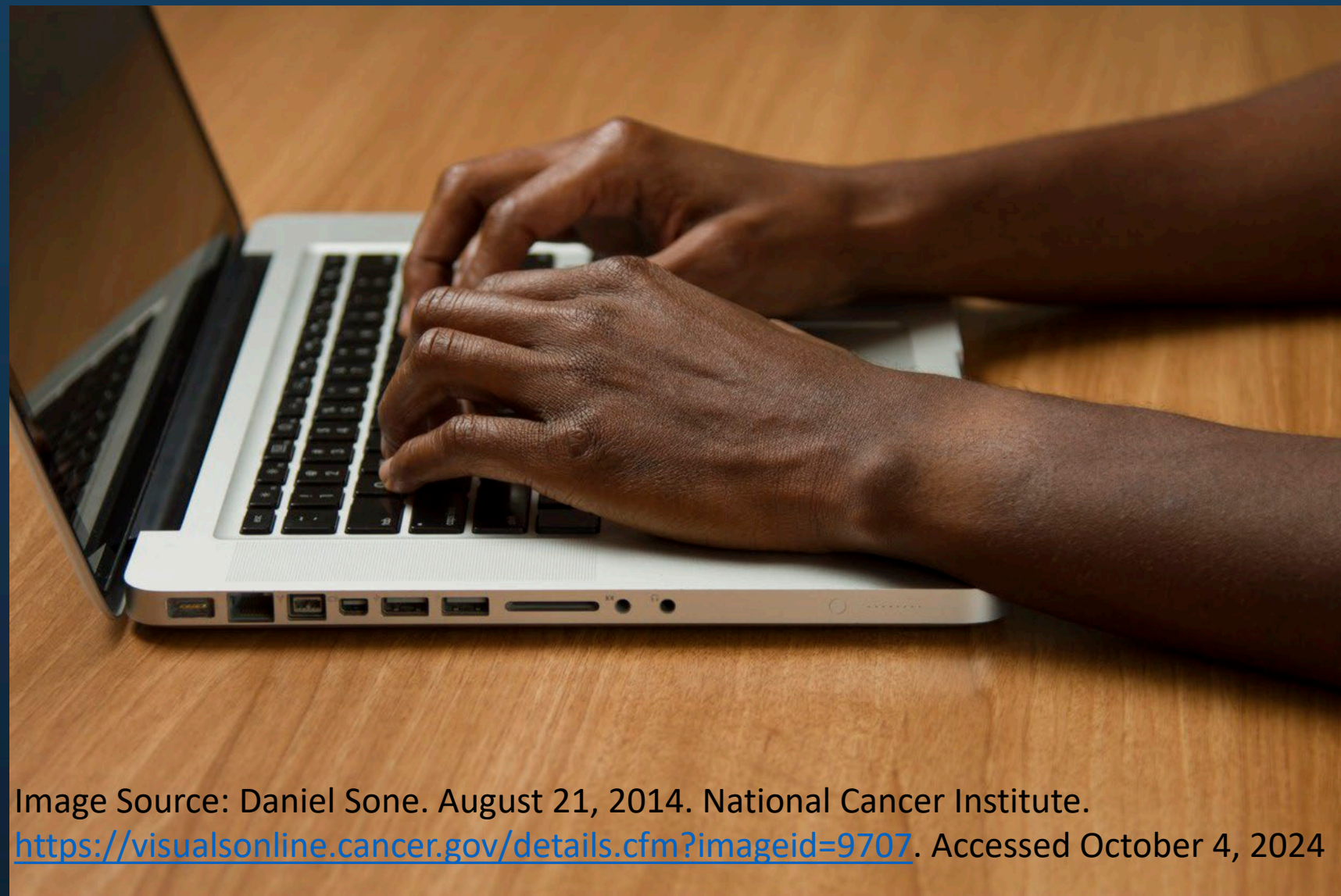
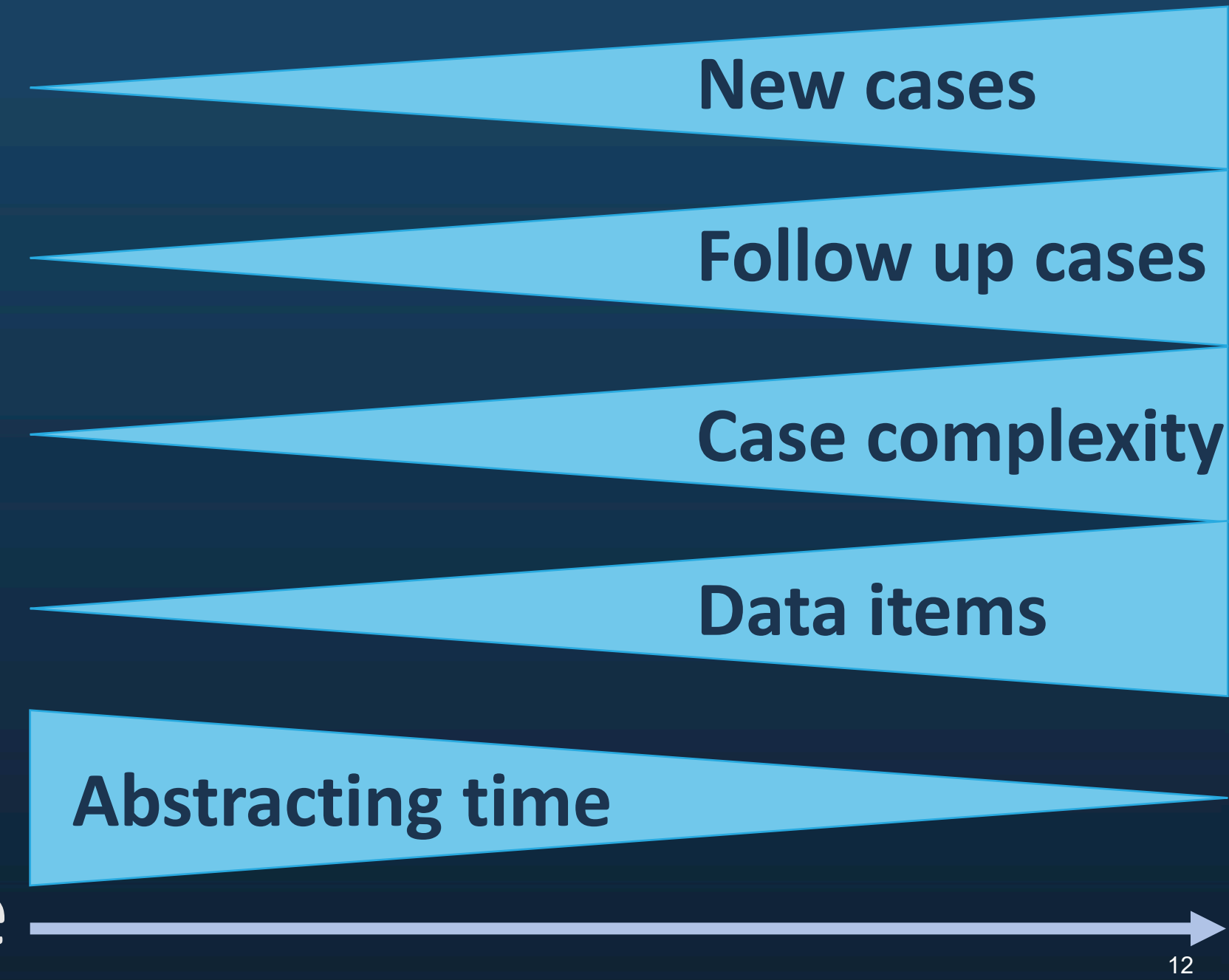


Image Source: Daniel Sone. August 21, 2014. National Cancer Institute. https://visualsonline.cancer.gov/details.cfm?imageid=9707. Accessed October 4, 2024

USE BEGINS ON JANUARY 1, 2024

NCRA Registrar Workload & Staffing Studies (2011 & 2022)



Assessing ODS Data Quality via Benchmarking

Median/Multiple Outlier Testing (MMOT)

Calculate	proportion unknown
Find	the median of the data
Calculate	deviance from the median
Standardize	median of absolute deviance
Calculate	standardized deviation

$$\text{Proportion of unknown} = \frac{Unknown}{Total}$$

$$\tilde{y}(\text{y-tilde}) = median\left\{y_{ij}, i = 1, \dots, I, j = 1, \dots J\right\}$$

$$(y_{ij} - \tilde{y})$$

$$m_0 = median\{|y_{ij} - \tilde{y}|, i = 1, \dots, I, j = 1, \dots, J\}.$$

$$d_{ij} = \frac{y_{ij} - \tilde{y}}{1.48\, m_0}$$

Unknown	R1	R2	R3	R4	R5
2010	1	0	1	13	0
2011	0	2	0	24	1
2012	0	4	0	25	0
2013	1	5	0	36	0
2014	0	5	0	32	0

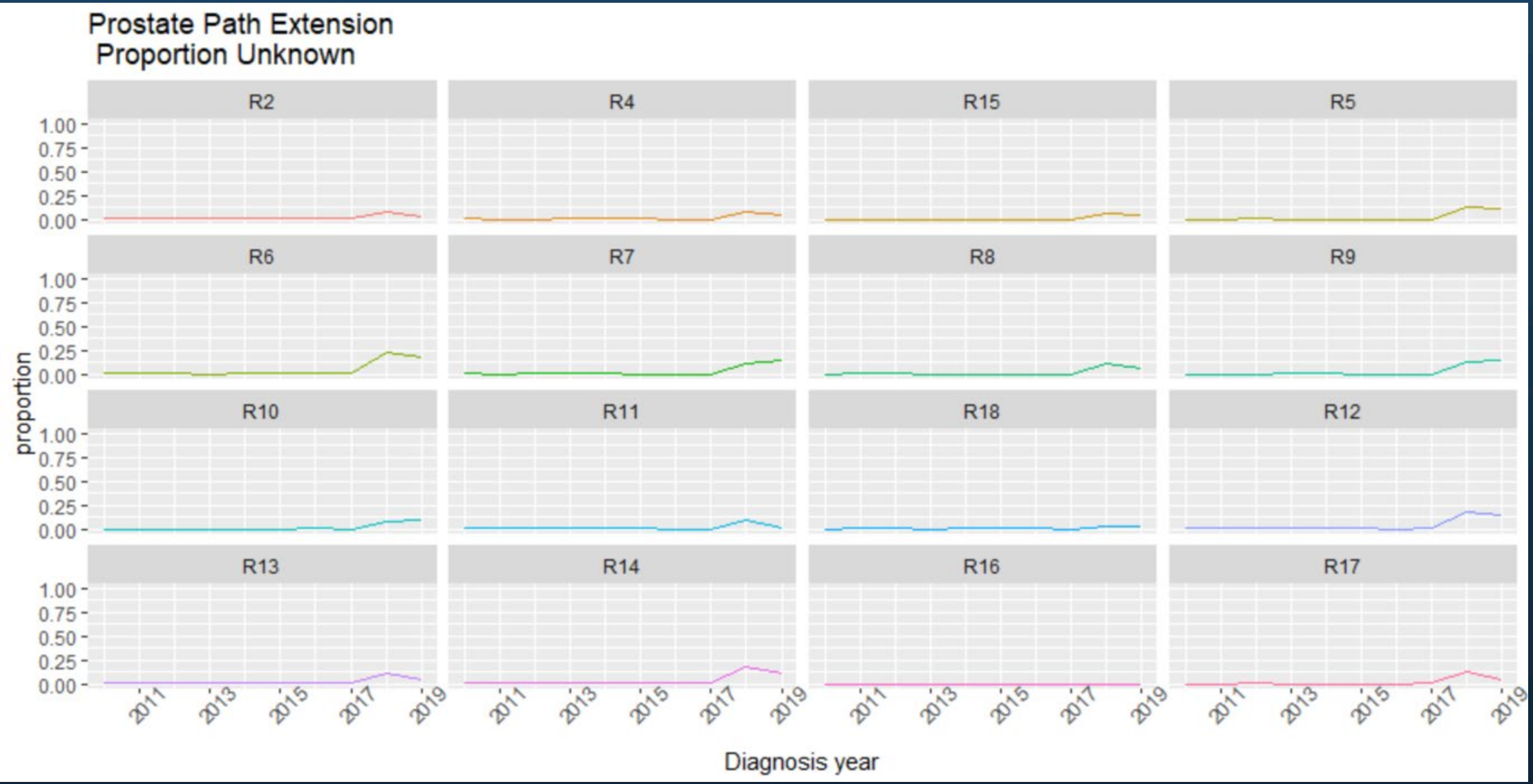
Total	R1	R2	R3	R4	R5
2010	2850	1734	1870	732	1685
2011	2730	1654	1815	736	1676
2012	2656	1740	1879	695	1635
2013	2601	1580	1839	712	1602
2014	2757	1582	1948	723	1736

Huann-Sheng Chen, PhD
NCI/SEER

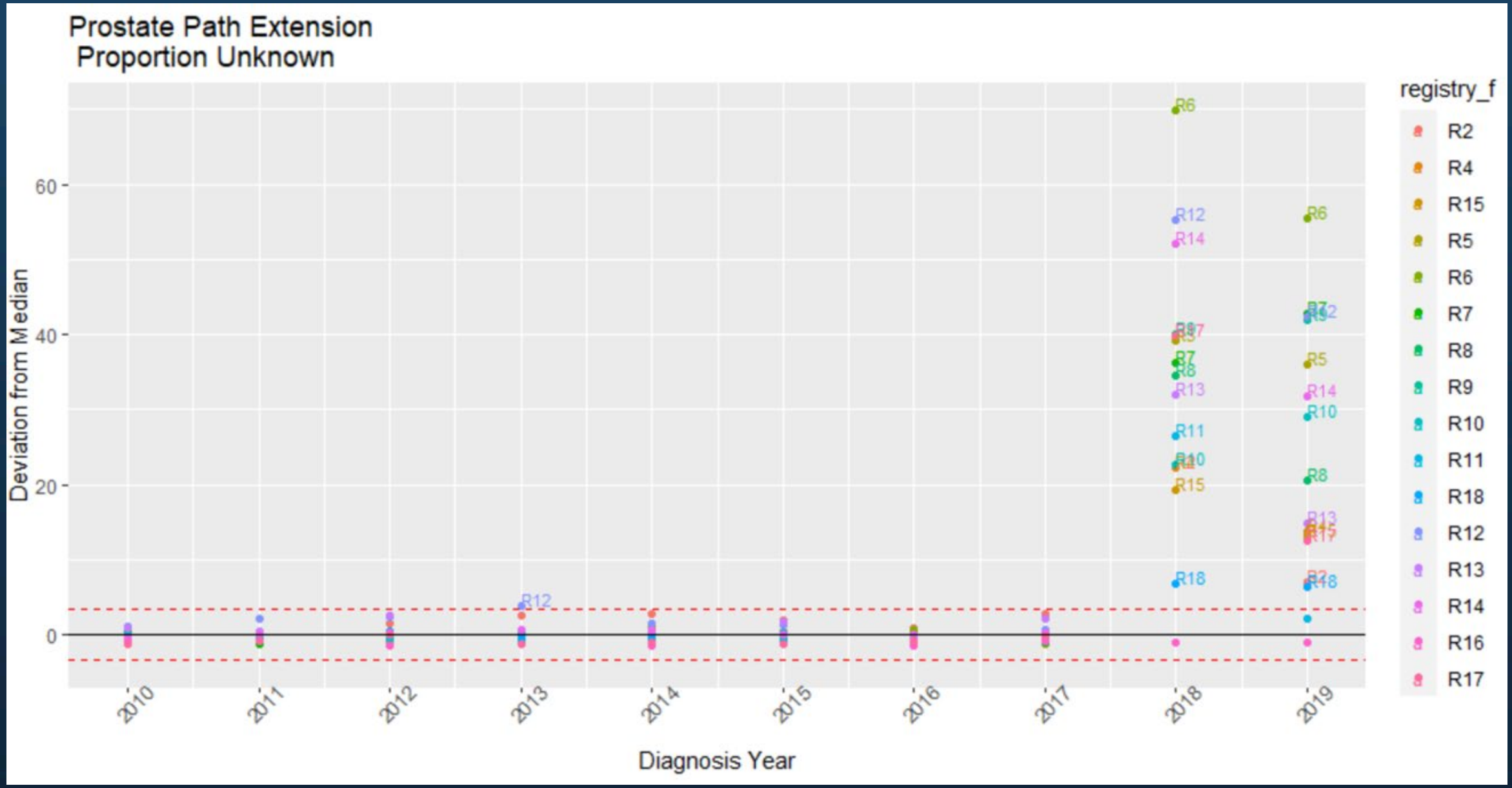
Extent of disease prostate pathologic extension

Code	Description	Summary Stage T*
000	Noninvasive	In situ
300	Confined to prostate without ECE	Localized
350	EPE without invasion of seminal vesicles	Regional
400	Invasion into seminal vesicles	
500	EPE without invasion into adjacent structures	
600	Invasion of bladder, external sphincter, EP urethra, rectum, muscle, ureter	
700	Extension to bone, penis, sigmoid colon, soft tissue other than periprostatic, other organs	Distant
800	No evidence of primary tumor	Unknown
900	No prostatectomy/autopsy	
950	No prostatectomy or surgery after disease progression	
999	No documentation or unknown if surgery	

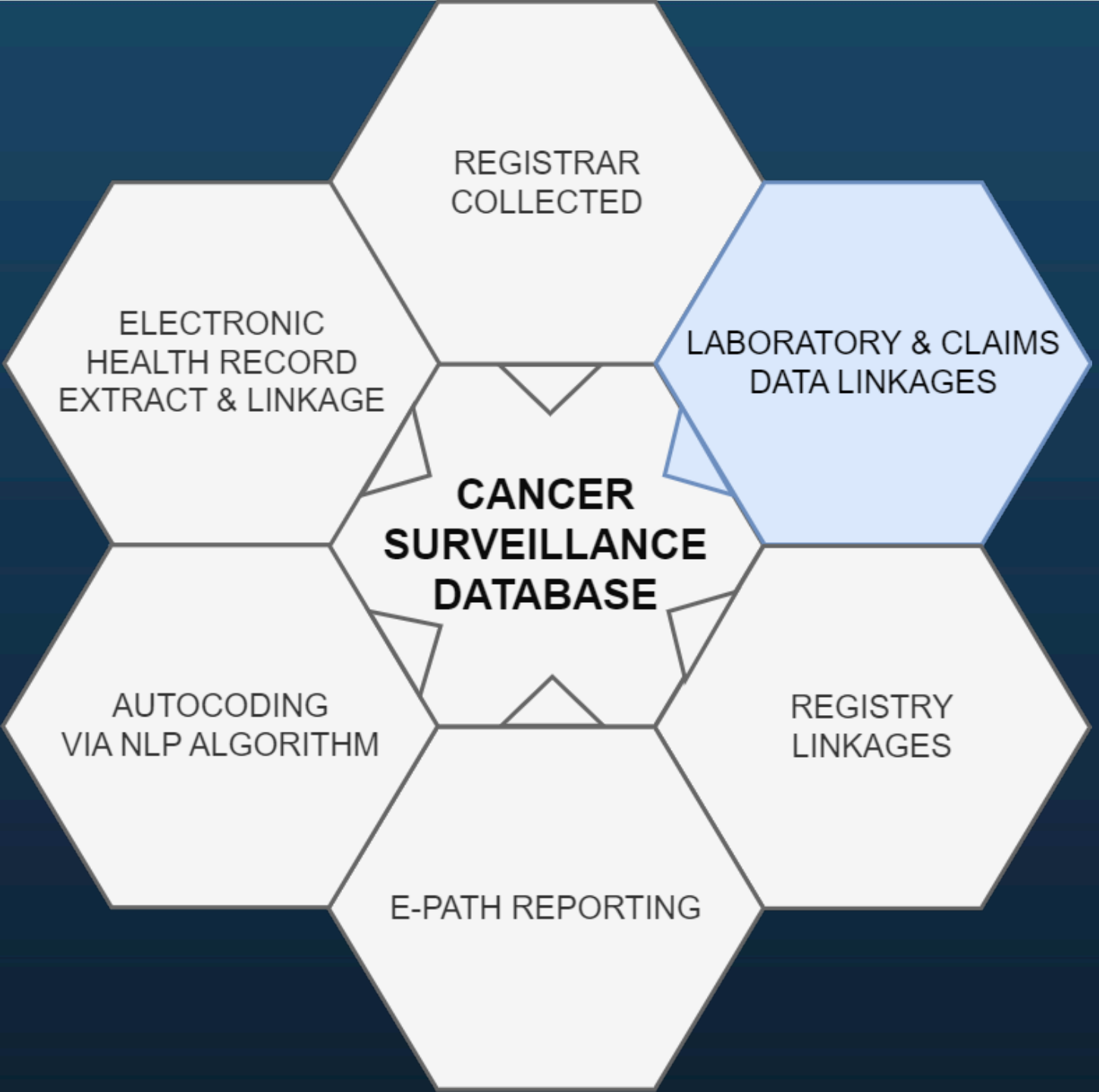
https://staging.seer.cancer.gov/eod_public/schema/3.1/prostate/?breadcrumbs=(~schema_list~)



Pamela Sanchez, MPH, ODS-C
NCI/SEER


Prostate Path Extension
Proportion Unknown
Deviation from Median
60
40
20
0
2010
2011
2012
2013
2014
2015
2016
2017
2018
2019
Diagnosis Year
registry_f
R2
R4
R15
R5
R6
R7
R8
R9
R10
R11
R18
R12
R13
R14
R16
R17

Modernization of cancer surveillance data acquisition & monitoring data quality



Laboratory Data

Image: American Institute for Economic Research n.d. "The Workings of the Gold Standard" https://aier.org/article/the-workings-of-the-gold-standard/. Accessed October 13, 2024.

SEER-cancer genomic & genetic data linkages

Valentina Petkov,
MD, MPH
NCI/SEER

Yao Yuan,
PhD, MPH
NCI/SEER
(former fellow)

OXFORD

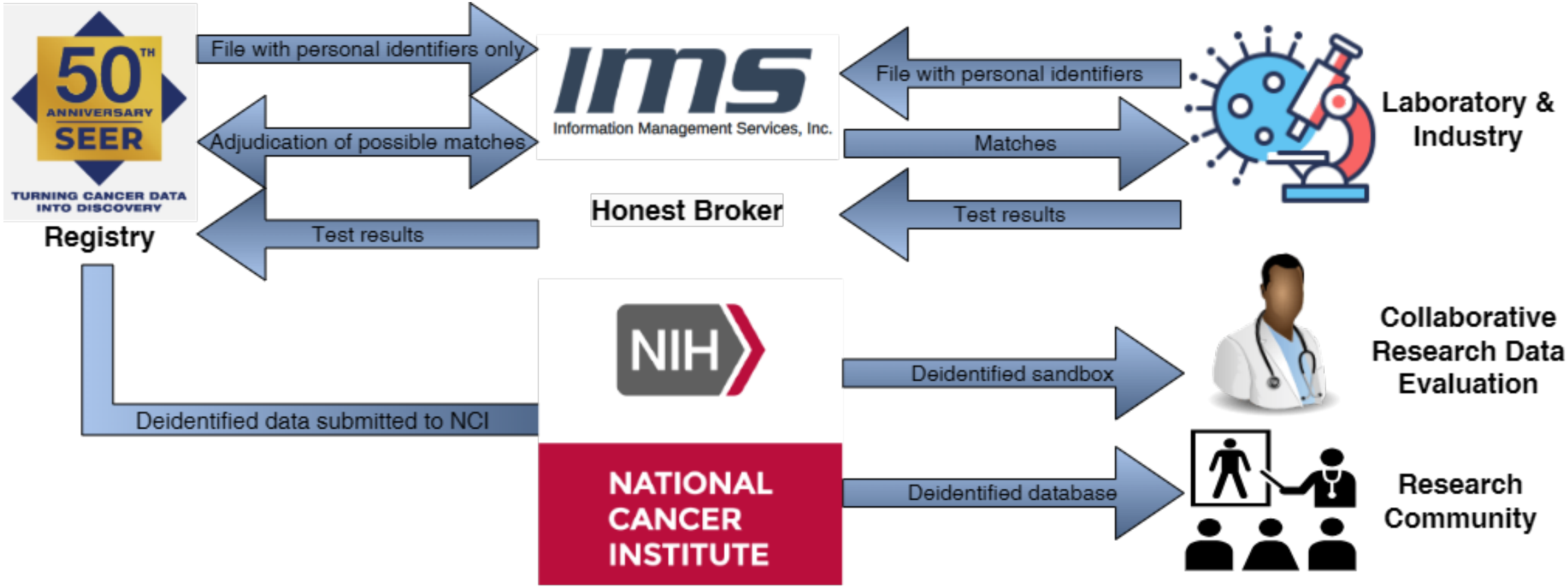
Journal of the National Cancer Institute Monographs, 2024, **2024(65)**, 168–179
https://doi.org/10.1093/jncimonographs/lgae013
Monograph

Reporting tumor genomic test results to SEER registries via linkages

Valentina I. Petkov, MD, MPH,[1,*] Jung S. Byun, PhD, MPH,[1] Kevin C. Ward, PhD, MPH,[2] Nicola C. Schussler, BS,[3] Natalie P. Archer, PhD, MS,[4] Suzanne Bentler, PhD,[5] Jennifer A. Doherty, MS, PhD,[6,7] Eric B. Durbin, DrPH, MS,[8] Susan T. Gershman, PhD, MPH,[9] Iona Cheng, MPH, PhD,[10] Tabassum Insaf, PhD, MPH, MBBS,[11] Lou Gonsalves, PhD,[12] Brenda Y. Hernandez, PhD,[13] Lori Koch,[14] Lihua Liu, PhD,[15] Alain Monnereau, MD, PhD,[16] Bozena M. Morawski, PhD, MPH,[17] Stephen M. Schwartz, PhD, MPH,[18] Antoinette Stroup, MD, PhD,[19] Charles Wiggins, PhD,[20] Xiao-Cheng Wu, MD, MPH,[21] Sarah Bonds, MS,[1] Serban Negoita, MD, DrPH,[1] Lynne Penberthy, MD, MPH[1]

Rationale for Lab Test Linkage

- **More efficient** way for **data collection** by centralizing data acquisition the Honest Broker between SEER registries & 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 & treated at community specialty practices


50TH ANNIVERSARY SEER
TURNING CANCER DATA INTO DISCOVERY
Registry
File with personal identifiers only
Adjudication of possible matches
Test results
IMS
Information Management Services, Inc.
Honest Broker
File with personal identifiers
Matches
Test results
Laboratory & Industry
NIH
NATIONAL CANCER INSTITUTE
Deidentified data submitted to NCI
Deidentified sandbox
Deidentified database
Collaborative Research Data Evaluation
Research Community

Industry Partners

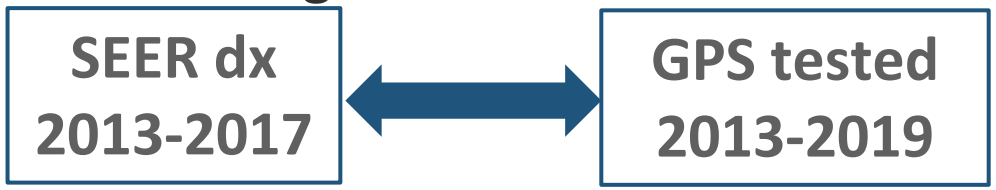
Clinical purposes	Active linkages	Under negotiation
Cancer risk prediction	EXACT SCIENCES, DECIPHER BIOSCIENCES, myriad, CASTLE BIOSCIENCES INCORPORATED	Prolaris-prostate: EndoPredict-breast, AGENDIA, MammaPrint
Cancer genomics testing by multigene-panel next generation sequencing	myriad, CARIS LIFE SCIENCES	TEMPUS, NEO GENOMICS, FOUNDATION MEDICINE, GUARDANT, Quest Diagnostics
Hereditary cancer genetic testing	myriad, Ambry Genetics A Konica Minolta Company, BioReference LABORATORIES an OPKO Health Company, INVITAE	color
Others (BMs, radiology, images)	syapse	AMBRA

Exact Sciences linkage

Establishing data release process for specialized database

OncotypeDX Genomic Prostate Score (GPS)

- Recommended in guidelines for treatment decisions & prediction of adverse pathology, on market since 2013
- Latest linkage in 2022



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

OncotypeDX Invasive Breast Recurrence Score (RS)

- 2004-2019
- 40% tests results reported were not coded by registrar (2010-2012)
- Agreement 94%
- Risk group misclassification <2%
- Released as specialized database

OncotypeDX DCIS

First time linkages

Caris

All SEER cases
SEER dx 2000-2018
Tests 2013-2019
IHC (MMR/PD-L1)
>300 tumor genes NGS

- Used in oncology practice
- Mutated/normal/VUS
- Pathologic variant
- TMB, MSI
- Testing method (NGS, RNA, IHC)
- Date

Castle Biosciences

Cutaneous melanoma
(DecisionDx-CM)

- 31 gene expression assay, suggested for stage I-III
- Predicts metastasis & level of follow-up
- SEER dx 2010-2019
- Risk group, test date, other path features

Uveal melanoma
(DecisionDx-UM)

- 15 gene expression assay for nonmetastatic disease
- 37% of cases linked to results

Decipher/Veracyte

Prostate cancer

- Decipher Bx/RP
- 22 gene assay, indicated at initial dx & after RP
- Prognostic & predictive
- Recommended in NCCN guidelines
- SEER dx 2010-2018
- Score, risk groups, & other path features

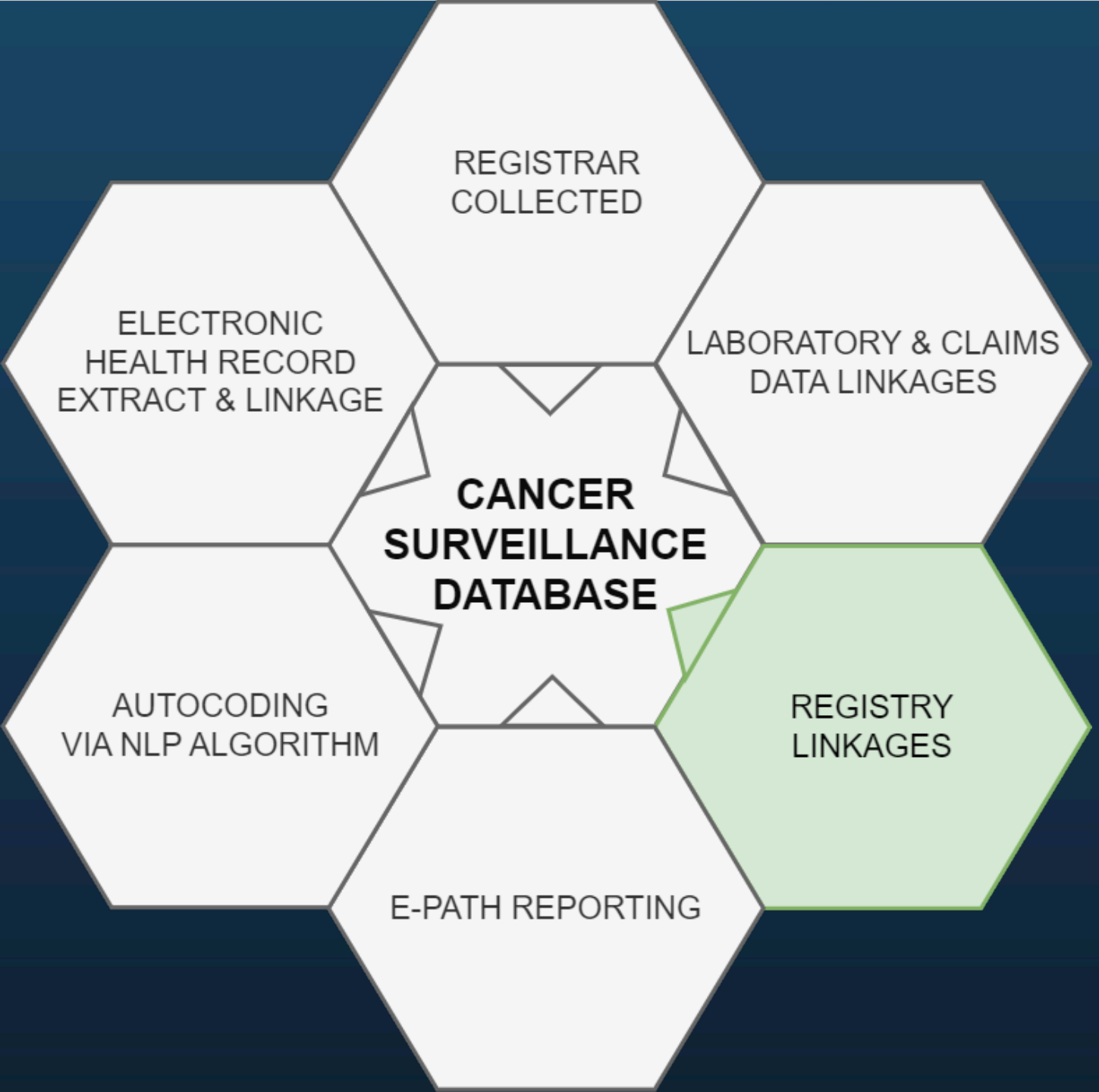
Top Limitations of Laboratory-SEER Linkages

False-negative linkage results

- PII incomplete or discrepant
- Diagnostic workup and/or treatment received out of state/catchment area
- Prior restrictions on results when all care received within VA/DoD health systems

Imprecise linkage results to specific cancer

- Multiple cancers (e.g., breast, cutaneous melanoma, lung, etc.)

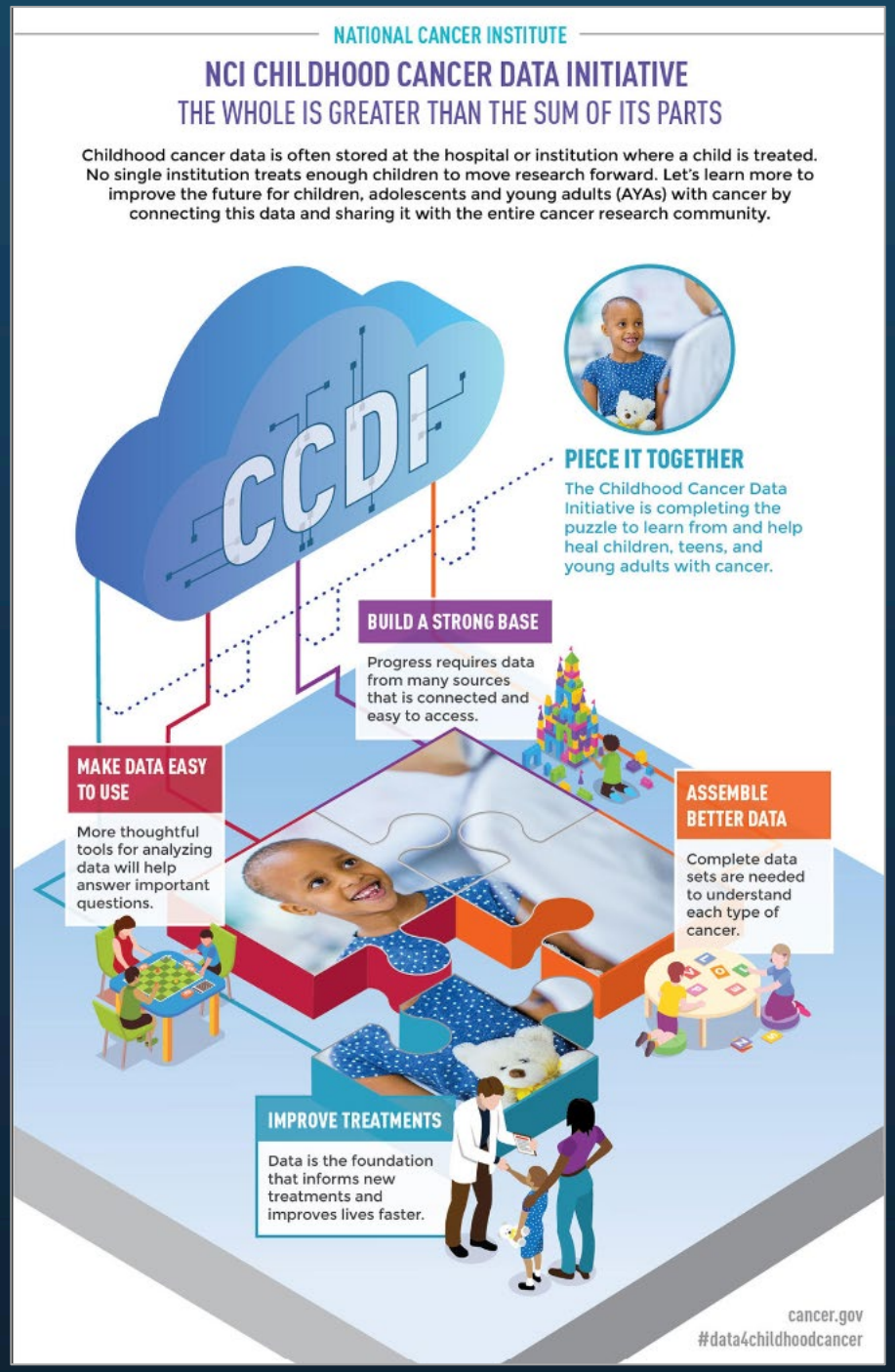


Registrar Coding

Image: American Institute for Economic Research n.d. "The Workings of the Gold Standard" https://aier.org/article/the-workings-of-the-gold-standard/. Accessed October 13, 2024.

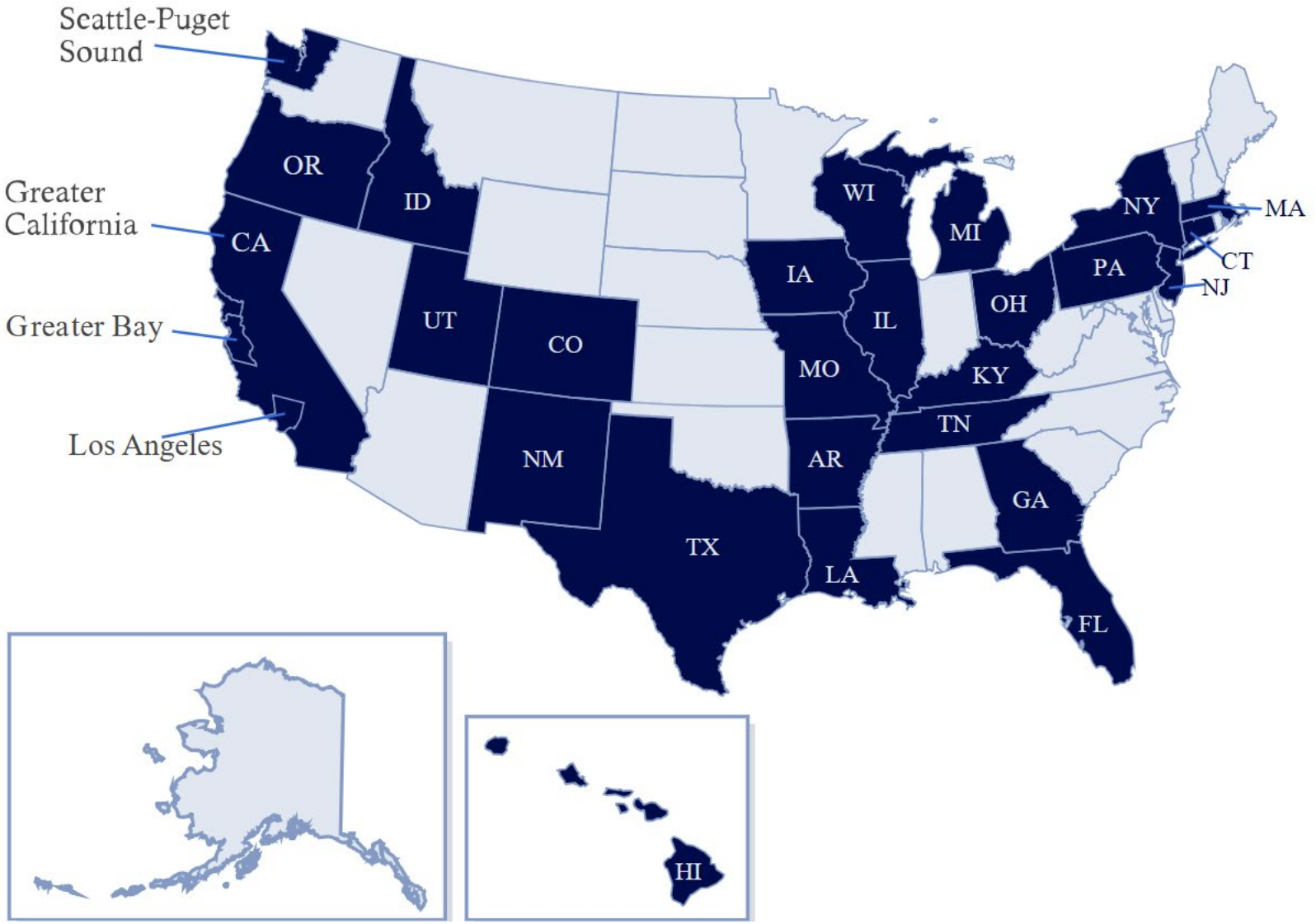
National Childhood Cancer Data Initiative

- 1 to 3% of cancers in US per year
- All rare cancers
- Established the National Childhood Cancer Registry (NCCR) cohort



National Cancer Institute. May 31, 2024. "About the Childhood Cancer Data Initiative (CCDI). https://www.cancer.gov/research/areas/childhood/childhood-cancer-data-initiative/about Accessed October 4, 2024.

National Childhood Cancer Registry (NCCR)



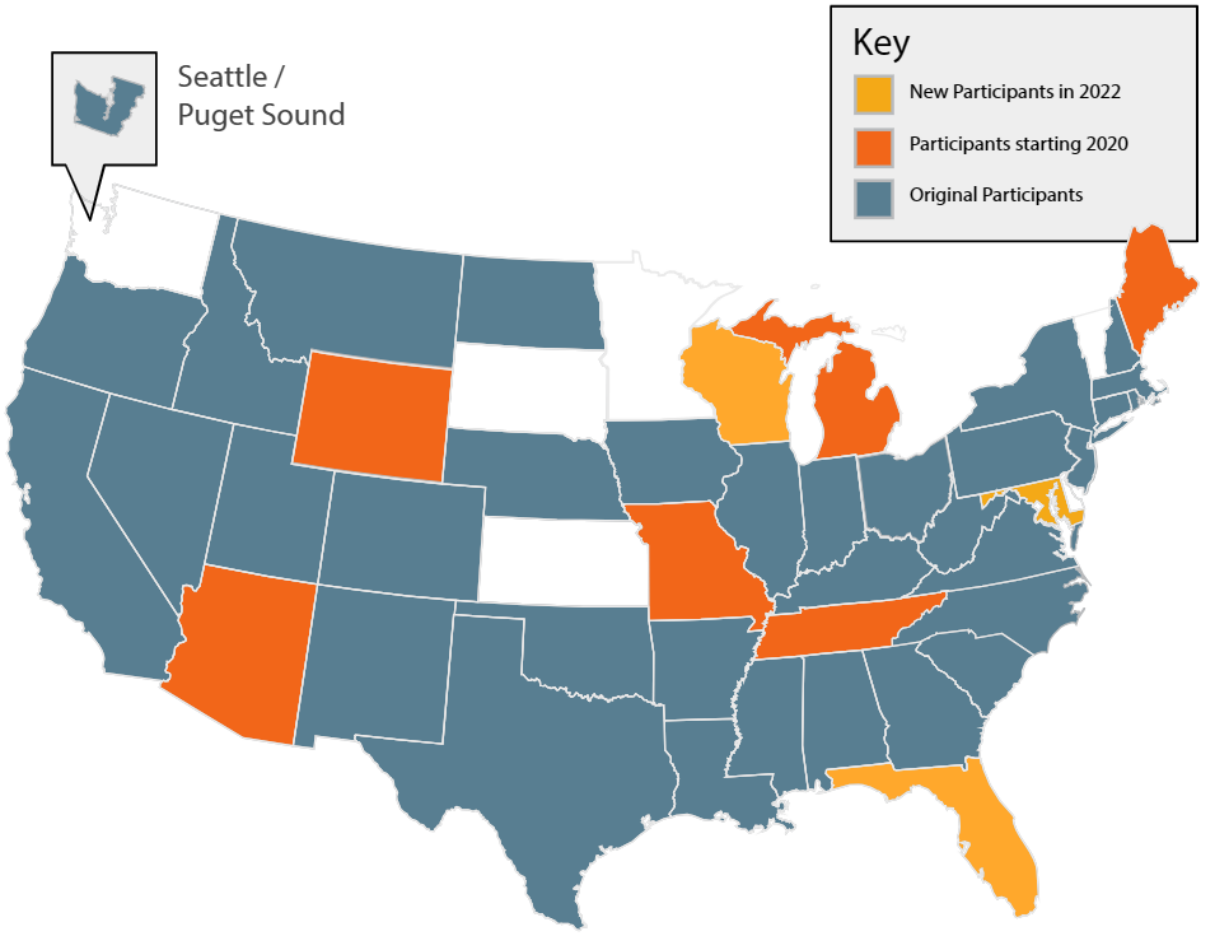
National Cancer Institute. n.d. "National Childhood Cancer Registry." https://nccrexplorer.ccdi.cancer.gov/about/nccr.html. Accessed October 4, 2024.

Virtual Pooled Registry Cancer Linkage System (VPR-CLS)

45 registries covering 95% of US population

Matching for

- Cohort studies
- Post-marketing surveillance
- Registries (multiple primaries, deduplication & outcomes & treatments sharing)



https://www.naaccr.org/about-vpr-cls/

Assessing completeness of registry data for pediatric vs. adult cancer cases using NAACCR CiNA data

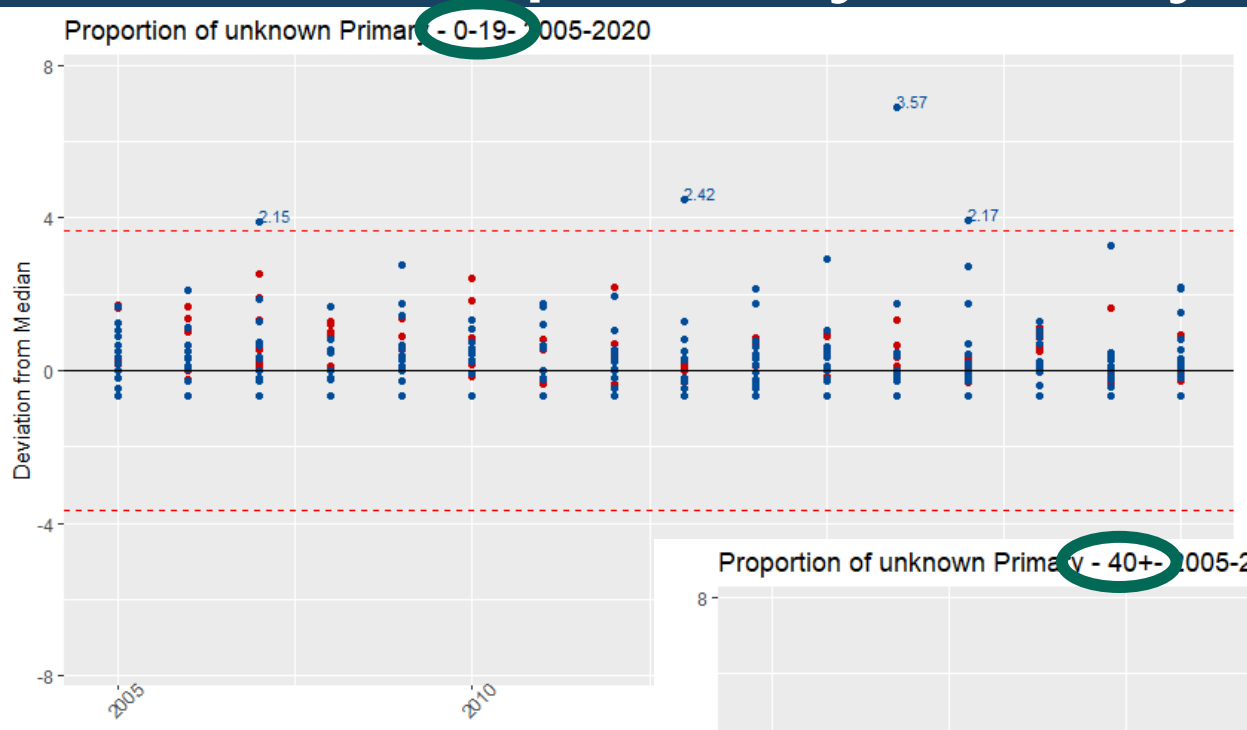
Fernanda Michels, MSc, PhD, ODS-C
NAACCR

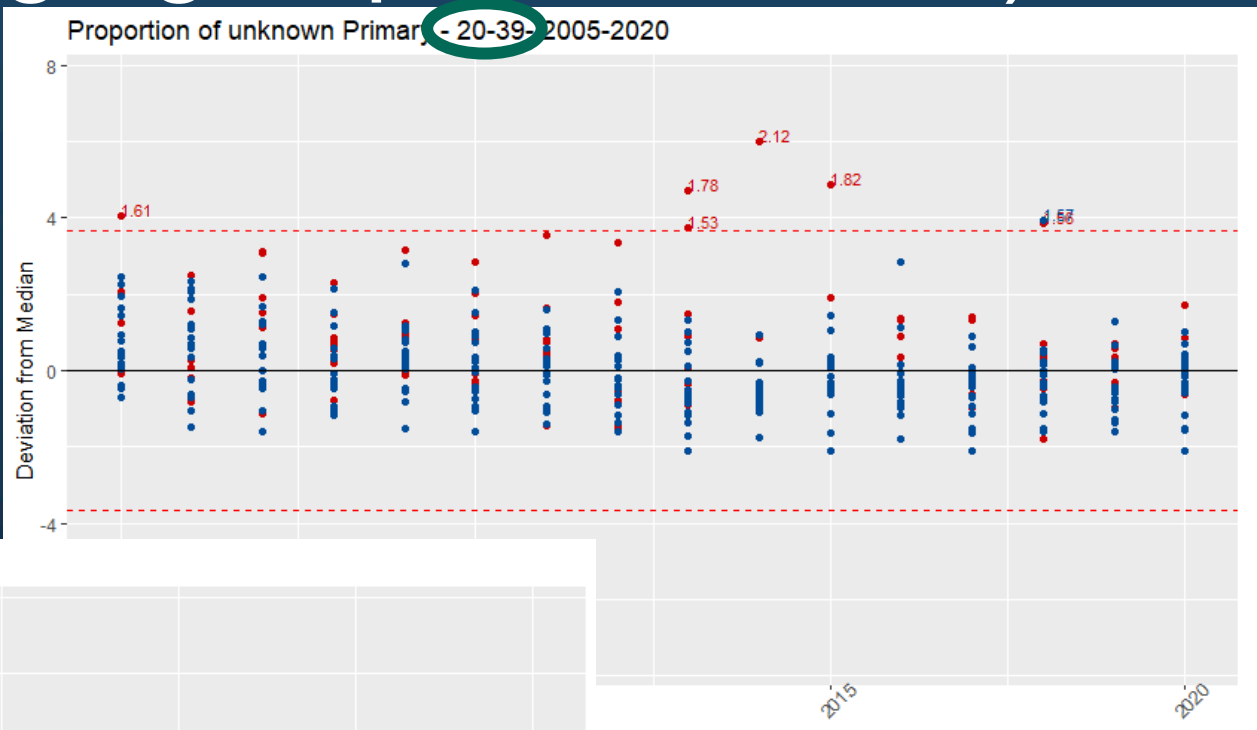
Stephanie Hill, MPH, ODS-C
NAACCR

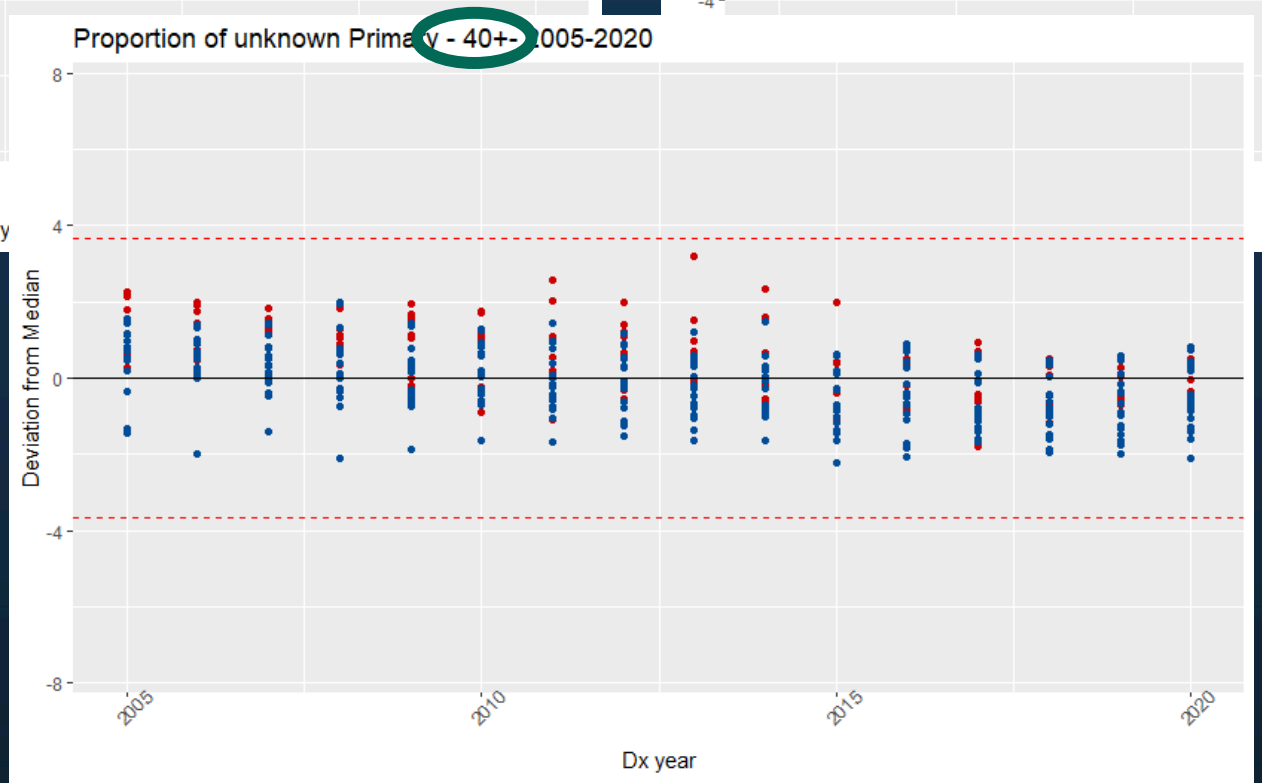
Betsy Kohler, MPH, ODS-C
NAACCR

Goncalo Forjaz, DVM, MSc,ODS-C
NCI/SEER

Unknown primary site by age group (2005-2020)

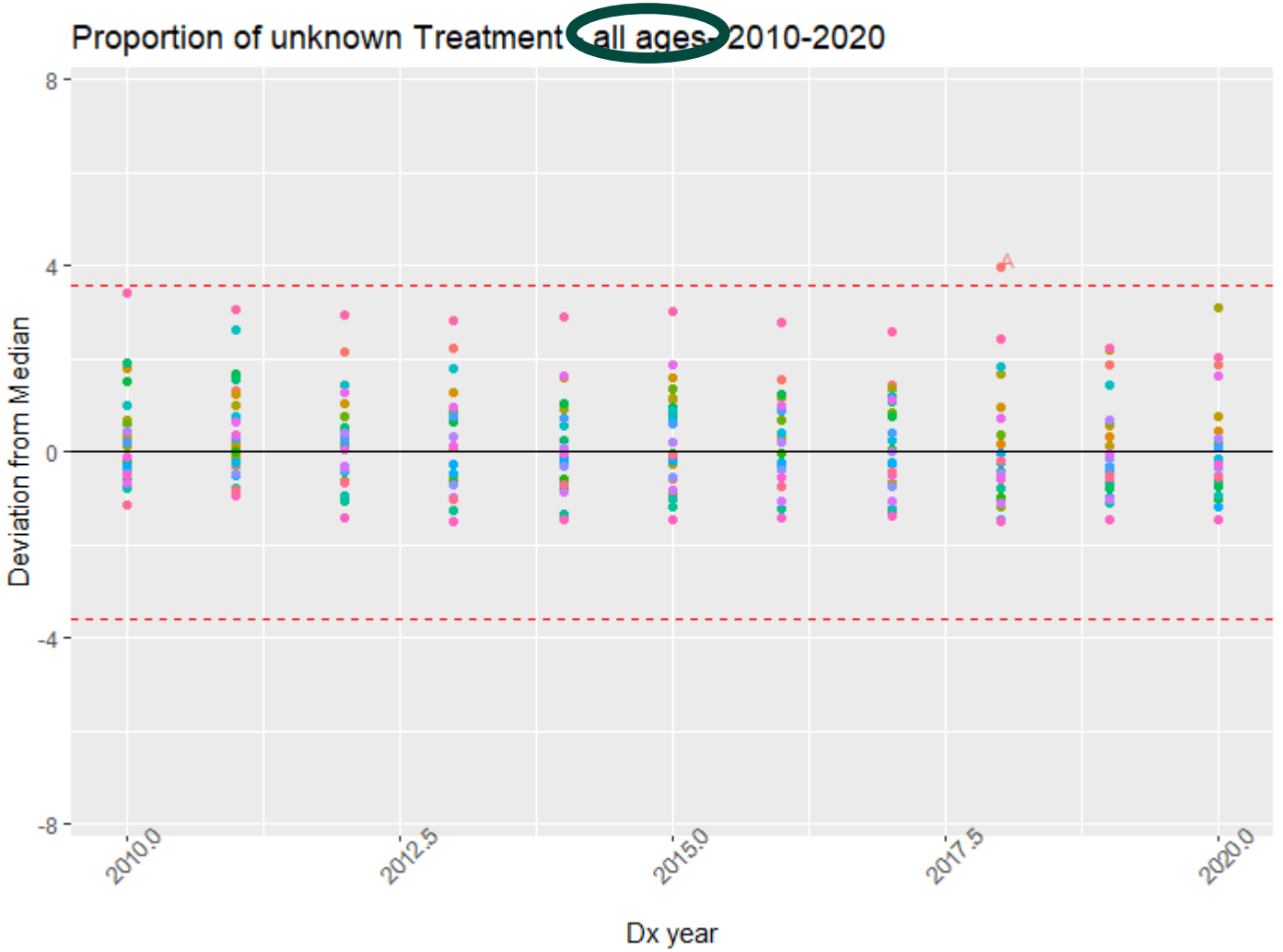


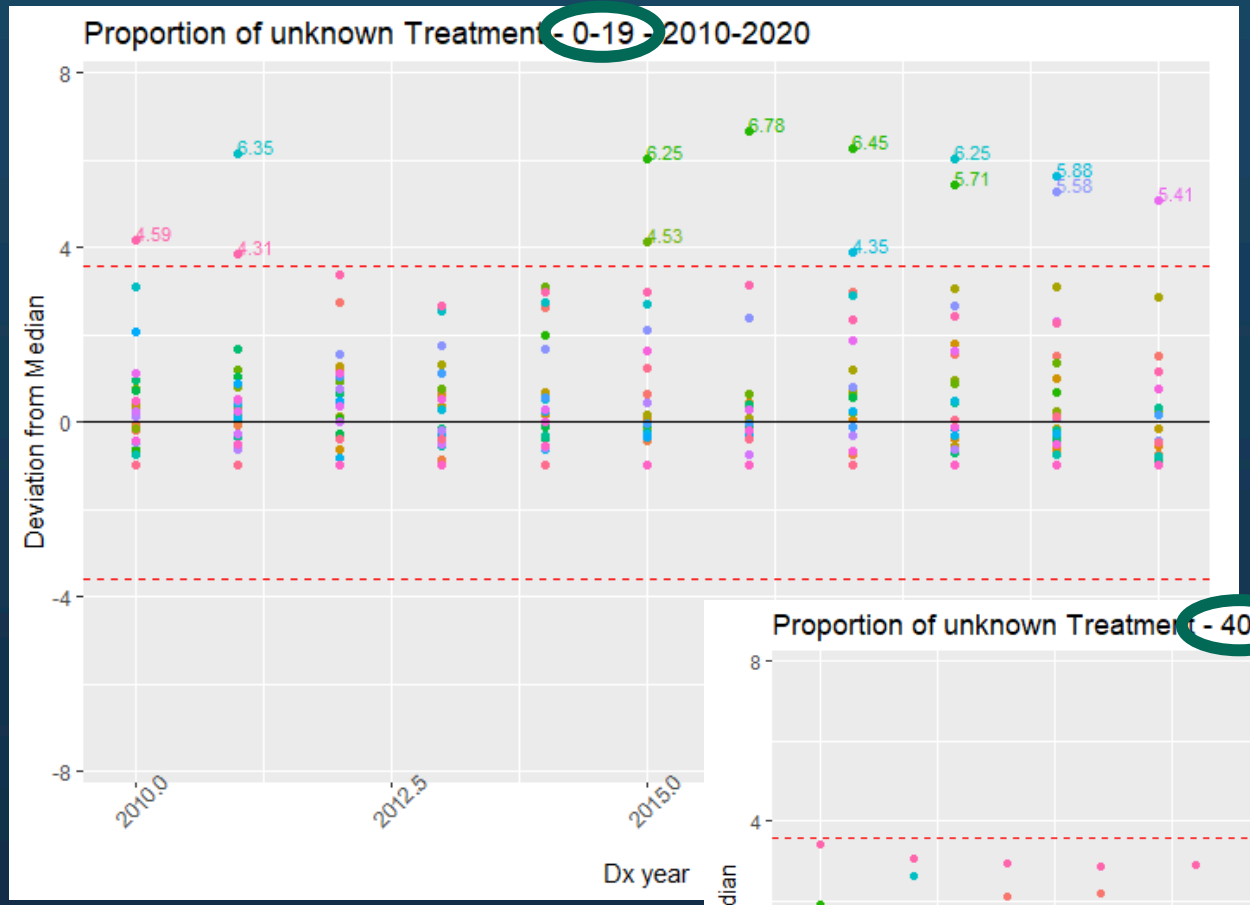


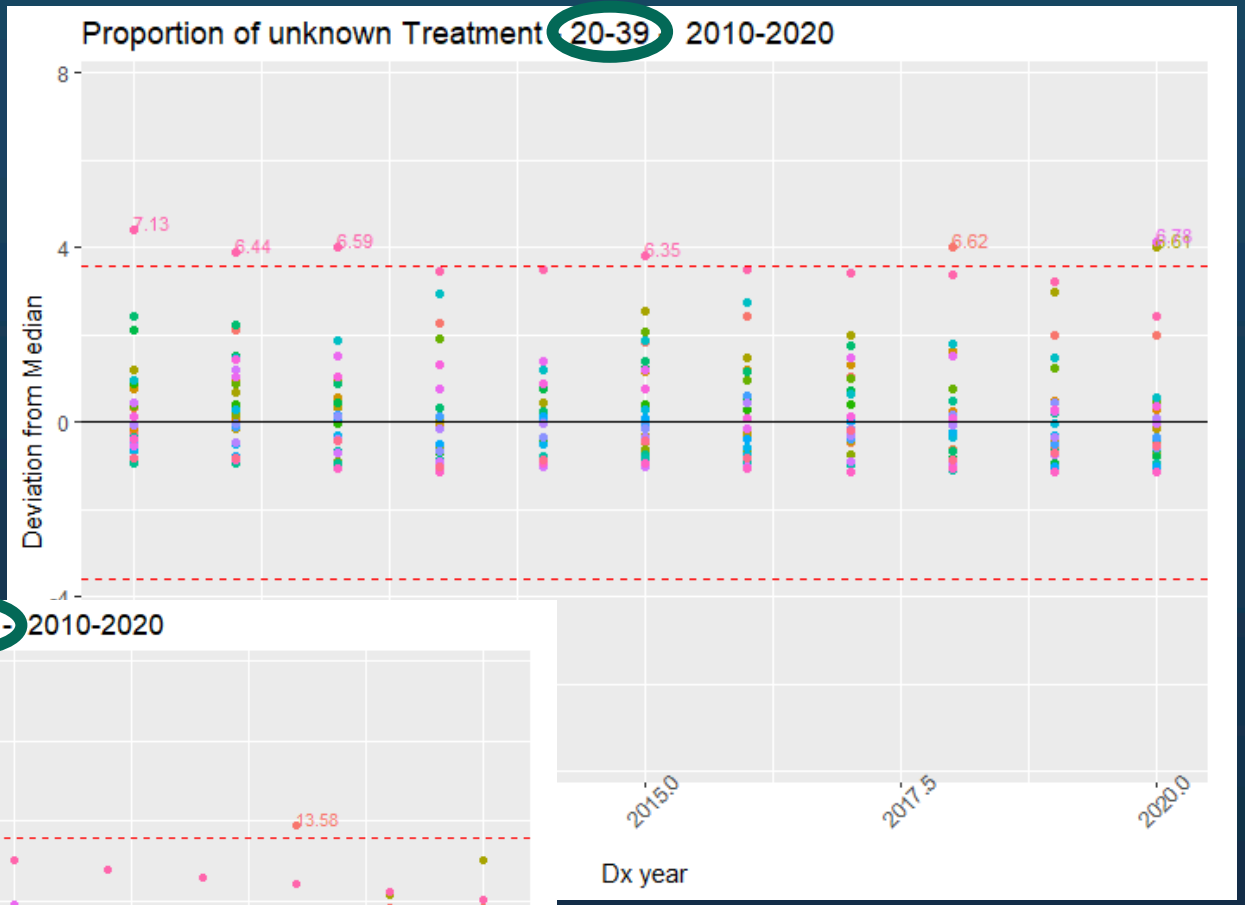


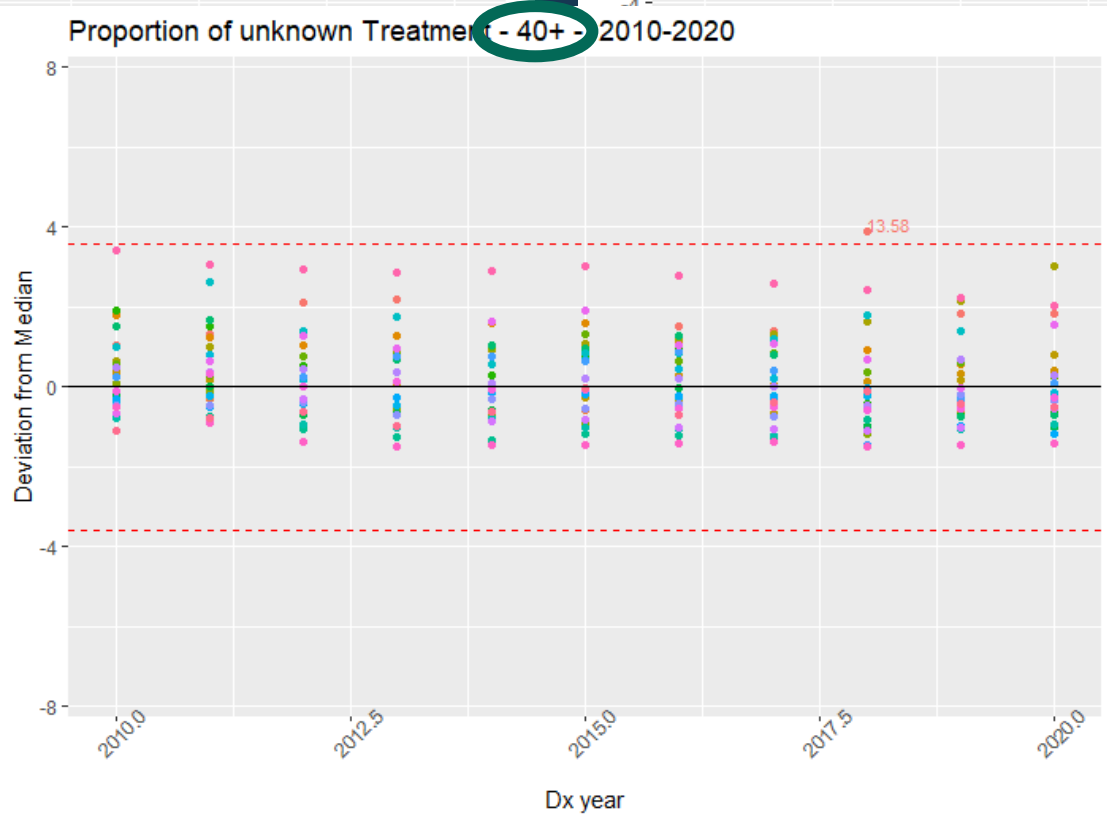
Proportion unknown for treatment status – all ages

Code	Treatment Status Description
0	No treatment given
1	Treatment given
2	Active surveillance (watchful waiting)
9	Unknown if given

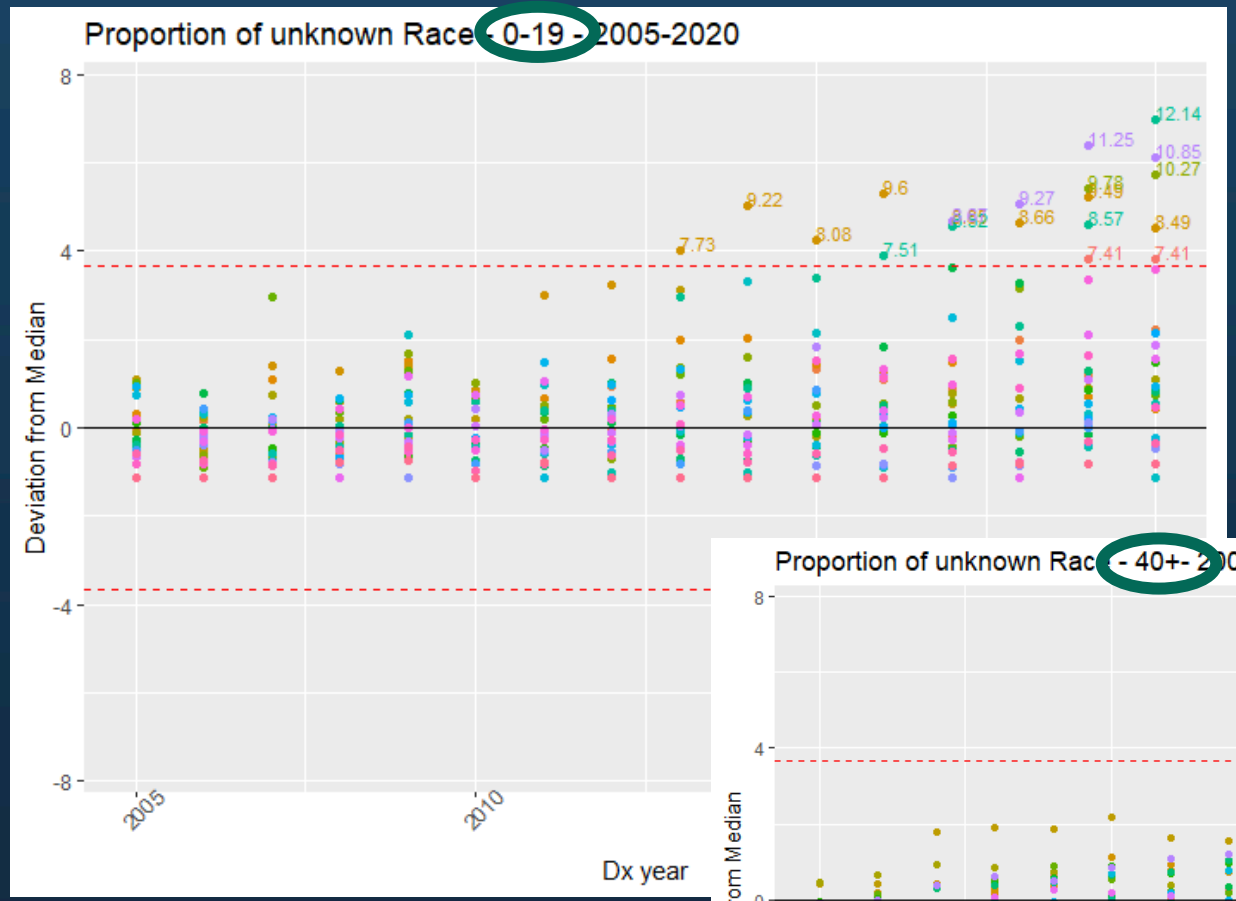


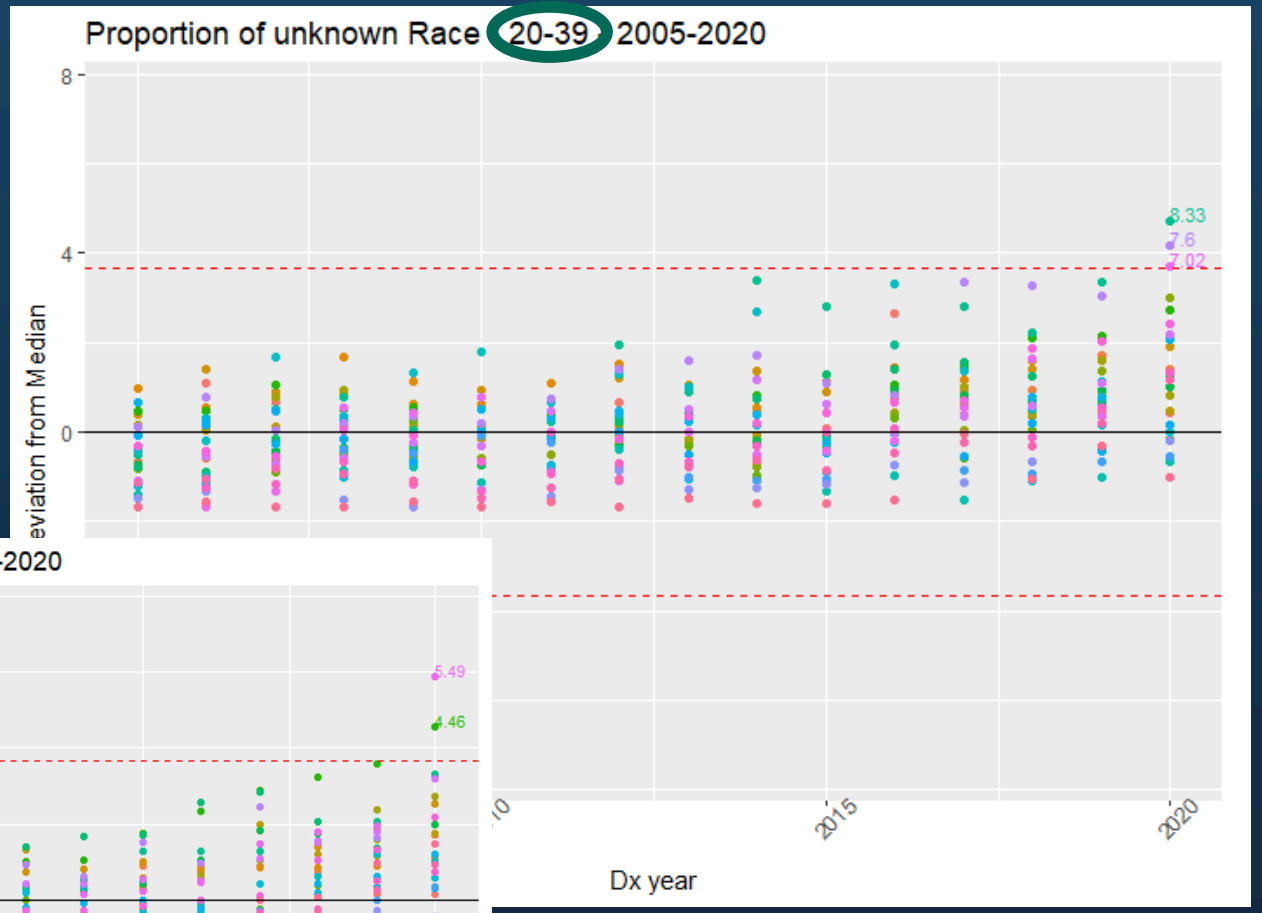

Proportion of unknown Treatment - 0-19 - 2010-2020
Deviation from Median
Dx year
8
4
0
-4
-8
2010.0
2012.5
2015.0
6.35
6.78
6.25
6.45
6.25
5.71
5.88
5.58
5.41
4.59
4.31
4.53
4.35

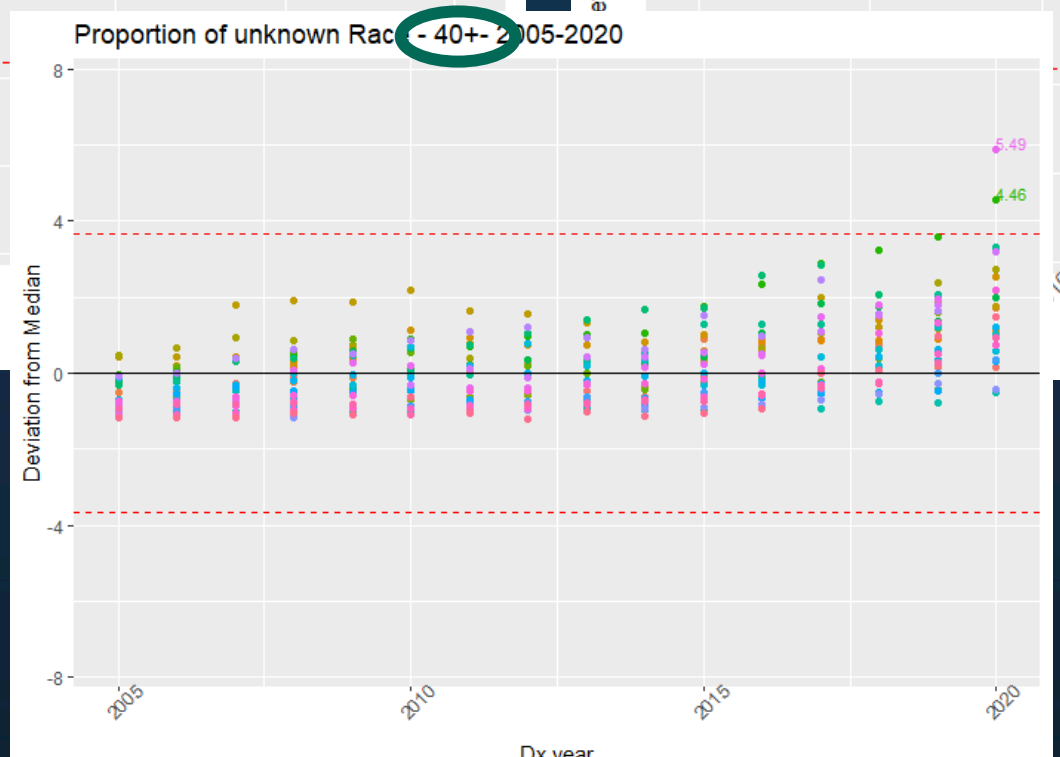

Proportion of unknown Treatment - 20-39 - 2010-2020
Deviation from Median
Dx year
8
4
0
-4
2015.0
2017.5
2020.0
7.13
6.44
6.59
6.35
6.62
6.78
6.61


Proportion of unknown Treatment - 40+ - 2010-2020
Deviation from Median
Dx year
8
4
0
-4
-8
2010.0
2012.5
2015.0
2017.5
2020.0
13.58

Proportion unknown for race by age group







Manually ODS coded Consolidated Tumor Case

Image: American Institute for Economic Research n.d. "The Workings of the Gold Standard" https://aier.org/article/the-workings-of-the-gold-standard/. Accessed October 13, 2024.

Modeling Outcomes using Surveillance data & Scalable AI for Cancer (MOSSAIC)

50TH ANNIVERSARY SEER
TURNING CANCER DATA INTO DISCOVERY

OAK RIDGE National Laboratory

OXFORD

Journal of the National Cancer Institute Monographs, 2024, **2024(65)**, 145–151
https://doi.org/10.1093/jncimonographs/lgae018
Monograph

Machine learning and deep learning tools for the automated capture of cancer surveillance data

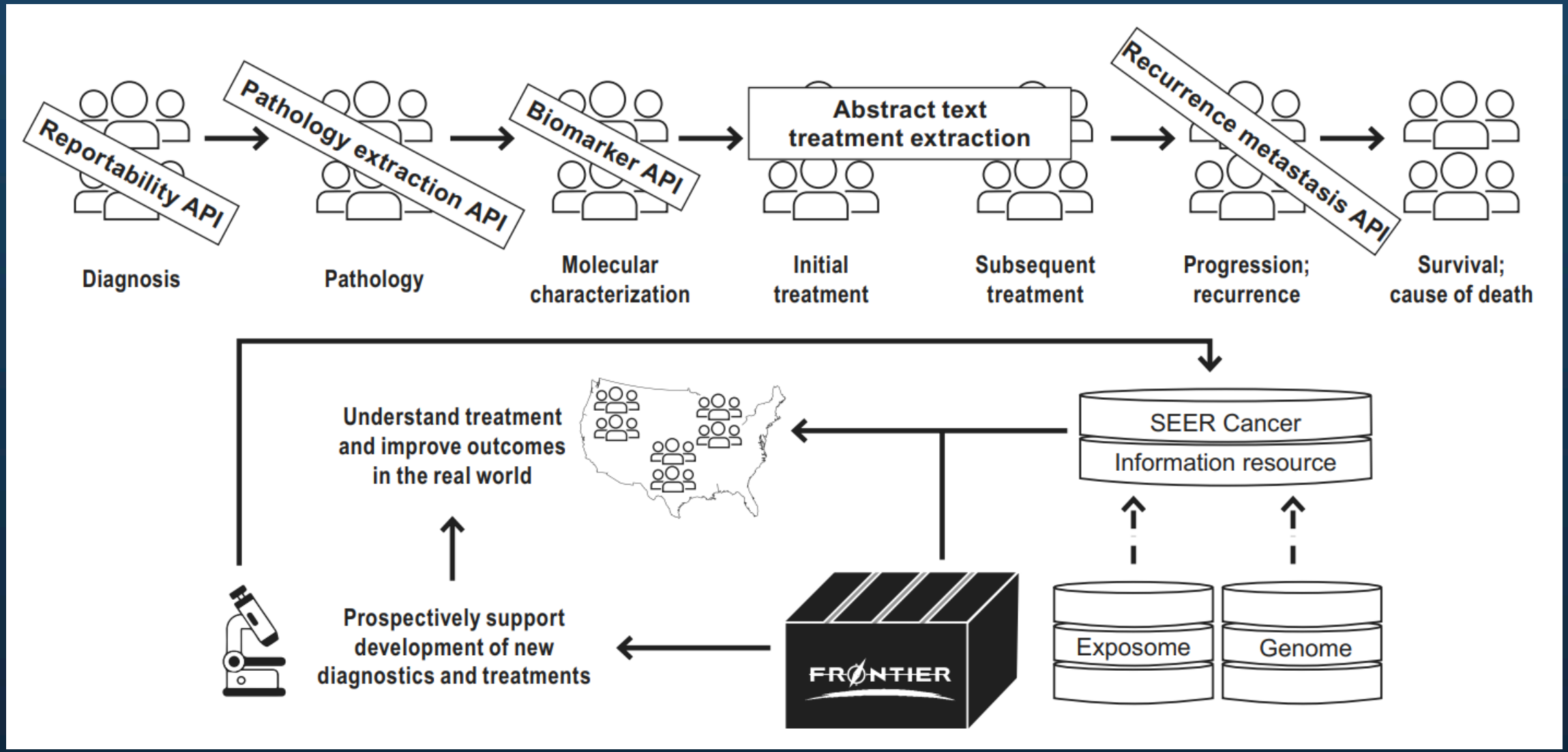
Elizabeth Hsu, PhD, MPH,[1,*] Heidi Hanson, PhD,[2] Linda Coyle, BS,[3] Jennifer Stevens, BS,[3] Georgia Tourassi, PhD,[4] Lynne Penberthy, MD, MPH[1]

[1]Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, Bethesda, MD, USA
[2]Advanced Computing for Health Sciences, Computing and Computational Sciences Directorate, Oak Ridge National Laboratory, Oak Ridge, TN, USA
[3]Information Management Services Inc, Calverton, MD, USA
[4]Computing and Computational Sciences Directorate, Oak Ridge National Laboratory, Oak Ridge, TN, USA

*Correspondence to: Elizabeth Hsu, PhD, MPH, Surveillance Research Program, Division of Cancer Control and Population Sciences, National Cancer Institute, 9609 Medical Center Dr, Bethesda, MD 20852, USA (e-mail: hsuel@mail.nih.gov).

Lynne Penberthy, MD, MPH
NCI/SEER

Elizabeth Hsu, PhD, MPH
NCI/SEER

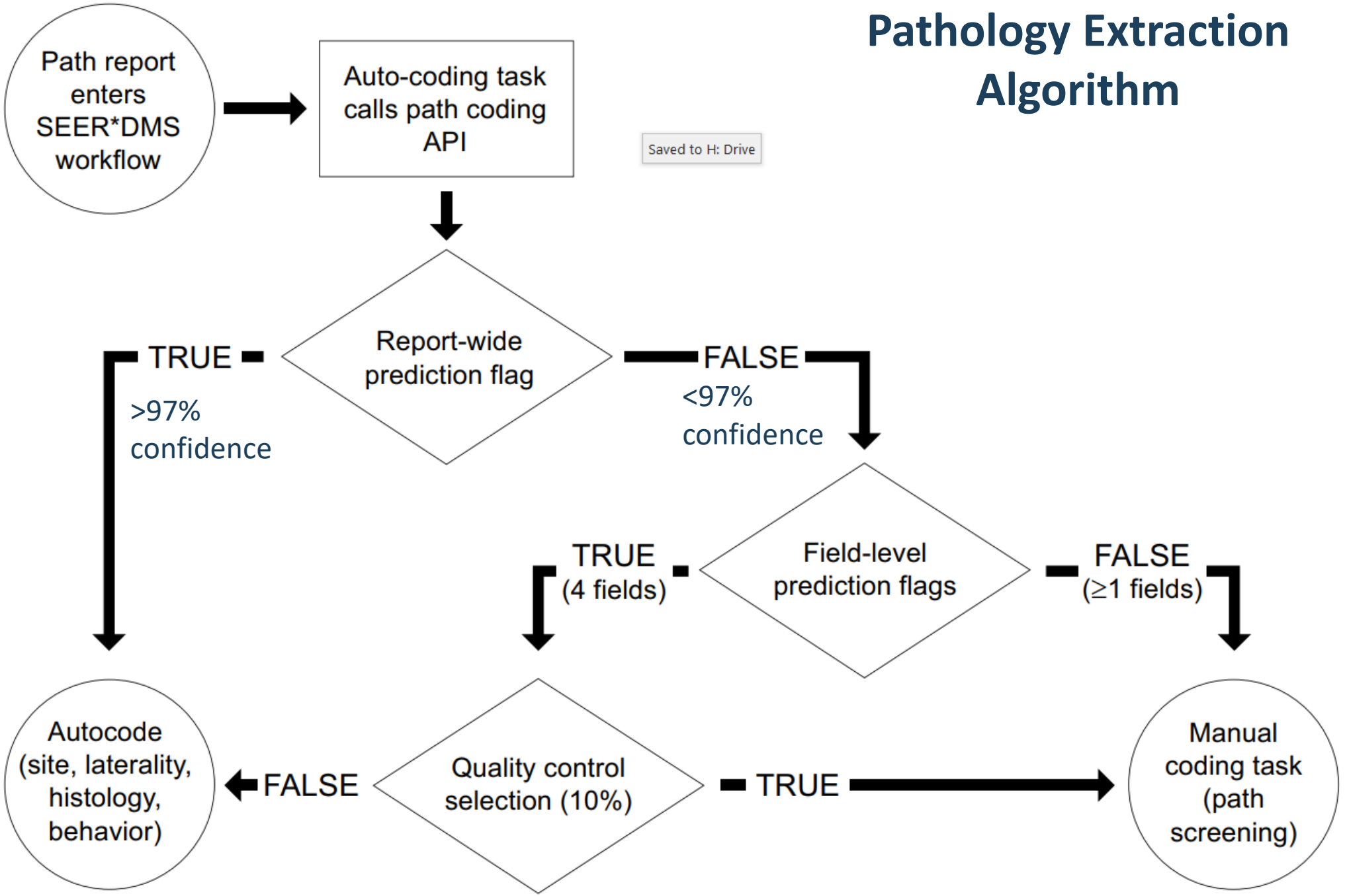


Hsu, Elizabeth et al. "Machine learning and deep learning tools for the automated capture of cancer surveillance data." *Journal of the National Cancer Institute. Monographs* vol. 2024,65 (2024): 145-151. doi:10.1093/jncimonographs/lgae018

Algorithms & APIs in MOSSAIC

Algorithm & API	Primary prediction task(s)	Status
Pathology extraction	Tumor site &/or subsite Laterality Histology Behavior	In production in SEER*DMS
Reportability	Reportable or Nonreportable	Testing & validation
Biomarkers	ER, PR, HER2 (breast cancer) *KRAS* mutation for (CRC)	Development to extend to more biomarkers
Recurrence & metastasis	Yes, no, unknown	Development

Pathology Extraction Algorithm



Hsu, Elizabeth et al. "Machine learning and deep learning tools for the automated capture of cancer surveillance data." *Journal of the National Cancer Institute. Monographs* vol. 2024,65 (2024): 145-151. doi:10.1093/jncimonographs/lgae018

Performance of Algorithms

Pathology Extraction Algorithm

Measure	Report-level version of API	Case-level Context version of API
Reports API can autocode	17.5%	23 to 27%
Accuracy (field or report)	98% (range: 97.1 to 99.4%)	>98%
Speed relative to human	~18,000 times faster	

Knowledge Transfer of Biomarker Algorithm

From KRAS (CRC) & HER2 (breast)

KRAS (lung) & HER2 (stomach, esophagus, lung) all >98% accuracy with confidence >97%

Timeliness of

- Data acquisition
- Reporting statistics
- Data release for cancer research & control
- Data standards

https://img.freepik.com/free-photo/hourglass-with-sand-middle-word-sand-it_123827-23414.jpg?size=626&ext=jpg. Accessed October 9, 2024

Moving to more timely data acquisition & incidence reporting

OXFORD

Journal of the National Cancer Institute Monographs, 2024, **2024(65)**, 123–131

https://doi.org/10.1093/jncimonographs/lgae024

Monograph

Toward real-time reporting of cancer incidence: methodology, pilot study, and SEER Program implementation

Huann-Sheng Chen, PhD[1], Serban Negoita, MD[1,*], Steve Schwartz, PhD, MPH[2], Elizabeth Hsu, PhD[1], Jennifer Hafterson, CTR[2], Linda Coyle, BS[3], Jennifer Stevens, BS[3], Anna Fernandez, PhD[1], Mary Potts, CTR[2], Eric J. Feuer, PhD[1]

[1]Division of Cancer Control and Population Sciences, National Cancer Institute, Rockville, MD, USA
[2]Fred Hutch Cancer Surveillance System (CSS), Seattle, WA, USA
[3]Information Management Services (IMS), Inc, Calverton, MD, USA

***Correspondence to:** Serban Negoita, MD, 9609 Medical Center Dr, Rockville, MD 20850, USA (e-mail: serban.negoita@nih.gov).

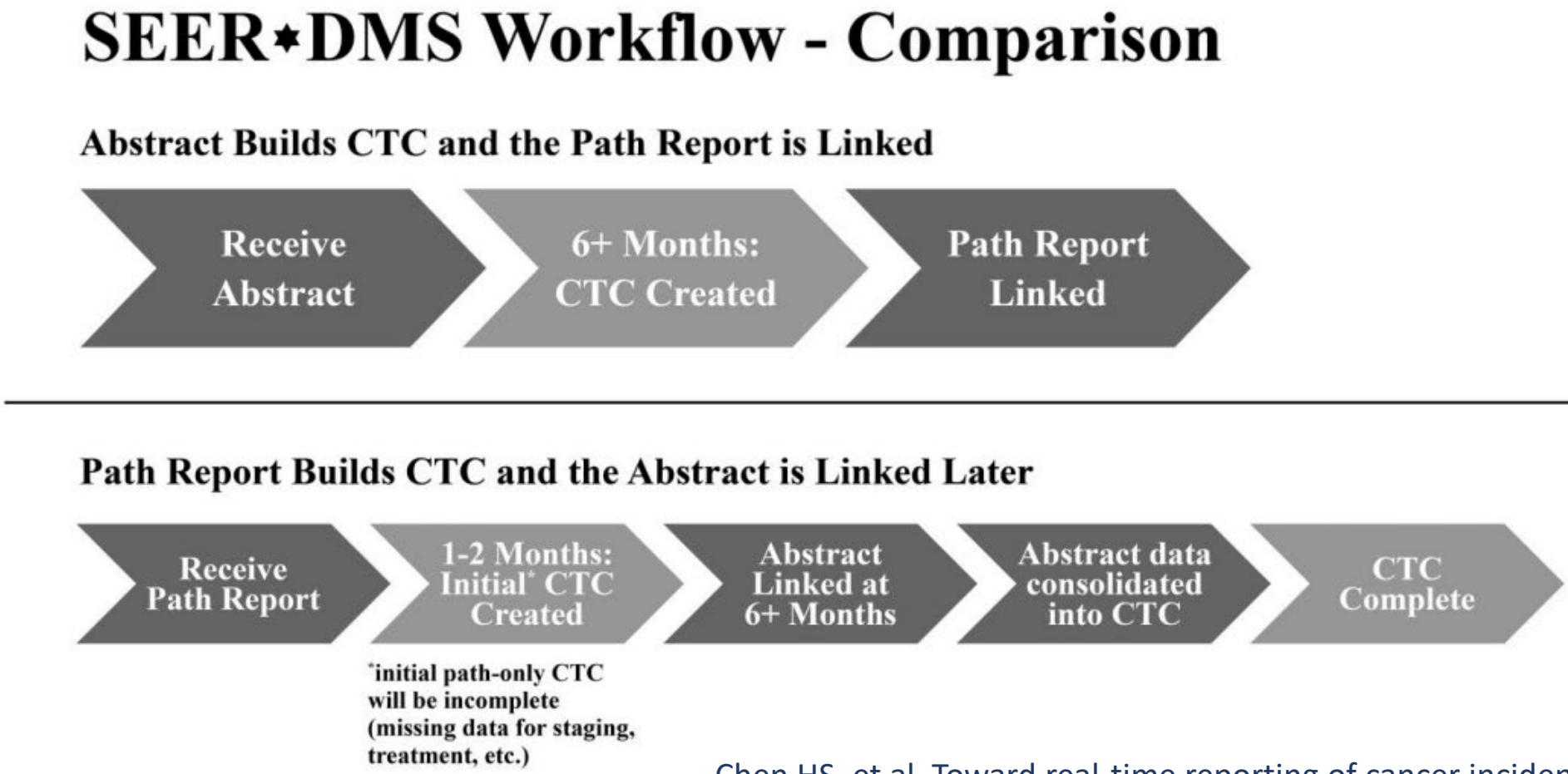
Huann-Sheng Chen, PhD
NCI/SEER

Serban Negoita, MD, DrPH
NCI/SEER

Current vs. Proposed Early Incidence Quality Standards

QUALITY STANDARDS CATEGORY	TIME OF SUBMISSION SINCE END OF DIAGNOSIS YEAR (ex: 2021)			
	22 MONTHS (November 2023)	14 MONTHS (February 2023)	10 MONTHS (November 2022)	2 MONTHS (February 2022)
Required Data Elements	Full Abstract	Full Abstract	Full Abstract of minimal data set	
Timeliness & Completeness	98% 1-yr reporting delay, DCO	95%	TBD	
Availability of Critical Data Elements	Primary site, Histology, Laterality, Stage	Primary site, Histology, Behavior, Year of diagnosis, Sex, & Age		
Accuracy of Geographic Elements	Urban geocoding , Rural geocoding	None		
Fitness for Survival Statistics	Follow-up %; missing COD	None		
Valid Values & Logical Checks	SEER Edits	Early Incidence Edits		

Data acquisition & submission



Chen HS, et al. Toward real-time reporting of cancer incidence: methodology, pilot study, and SEER Program implementation. *J Natl Cancer Inst Monogr*. 2024;2024(65):123-131. doi:10.1093/jncimonographs/lgae024

Issues with pathology report-based real-time reporting

- Sites less likely to have pathologic confirmation or are based on imaging (e.g., brain, spine, & liver tumors &/or metastatic disease)
- CAP eCPs & synoptic reporting not uniformly adopted
- Complex data fields not always available in EHR or structured pathology reports (race/ethnicity, sex at birth, & some biomarkers)

Data Standards Gap

Cancer Surveillance Standards

Pathology Report Language

n.d. *Navajo Bridge across Grand Canyon on route 89*. US Route 89 Appreciation Society, accessed December 5, 2023 https://usroute89.com/wp-content/uploads/2017/01/Navajo_Bridge_1989-07-1080x675.jpg

Cancer PathCHART (CPC)

Cancer

Pathology

Coding

Histology

And

Registration

Terminology

Statistics Canada Statistique Canada

Vision & End Goals

- One source of truth for tumor site-morphology combination standards
- Reducing differences between stakeholders
- Decreased implementation timeline
- Improved data quality

CPC Validity Standards

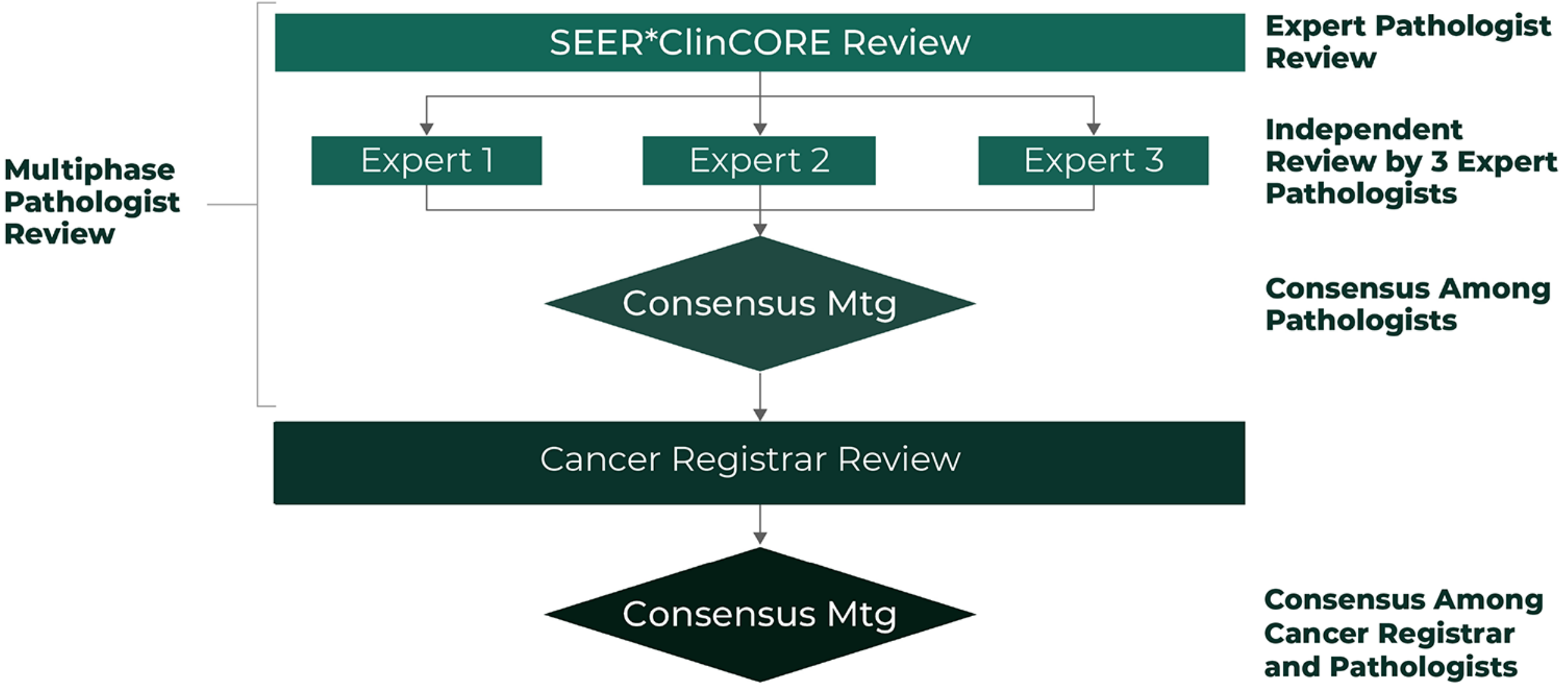
Sites Reviewed for 2024 Implementation

Validity Status	Site-Type Edit Errors	Coding in Cancer Registry Database
Valid	Will not generate edit errors	Can be coded
Impossible	Will generate an edit error	Cannot be coded
Unlikely*	Will generate an edit error	Requires manual override or correction to site and/or morphology to be coded

Sites Not Reviewed for 2024 Implementation

2024 Standard Used for Primary Sites Not Yet Reviewed	2024 Validity Status
2023 ICD-O-3 SEER Site-Histology Validation List	Valid
2023 Primary Site, Morphology-Imposs ICDO3 (SEER IF38)	Impossible
Combinations not included in 2023 Valid or Impossible standards listed above	Unlikely

Interdisciplinary Review Process



Organ Systems Reviewed

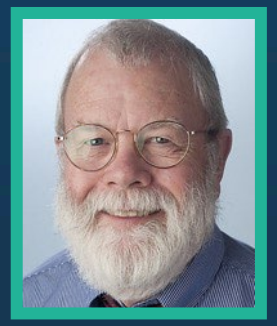
Implemented – 2024	To Be Implemented – 2025
Bone & soft tissue*	Central nervous system
Breast	Male genital*
Digestive	Respiratory
Female genital & reproductive	Soft tissue*
Male genital*	Thorax
Urinary*	Urinary*

*Review of morphologies at these sites completed for 2025

SEER*ClinCORE Pathologists

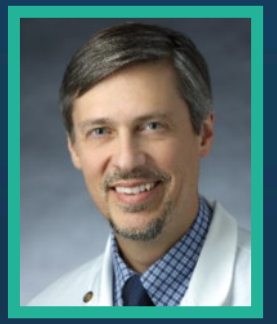
Aaron Auerbach
Hematopathology

Mary Beth Beasley
Thoracic Pathology

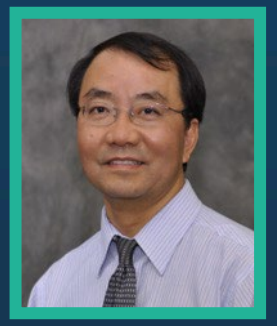


James Connolly
Breast Pathology

Jessica Davis
Bone/Soft Tissue &
Pediatric 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

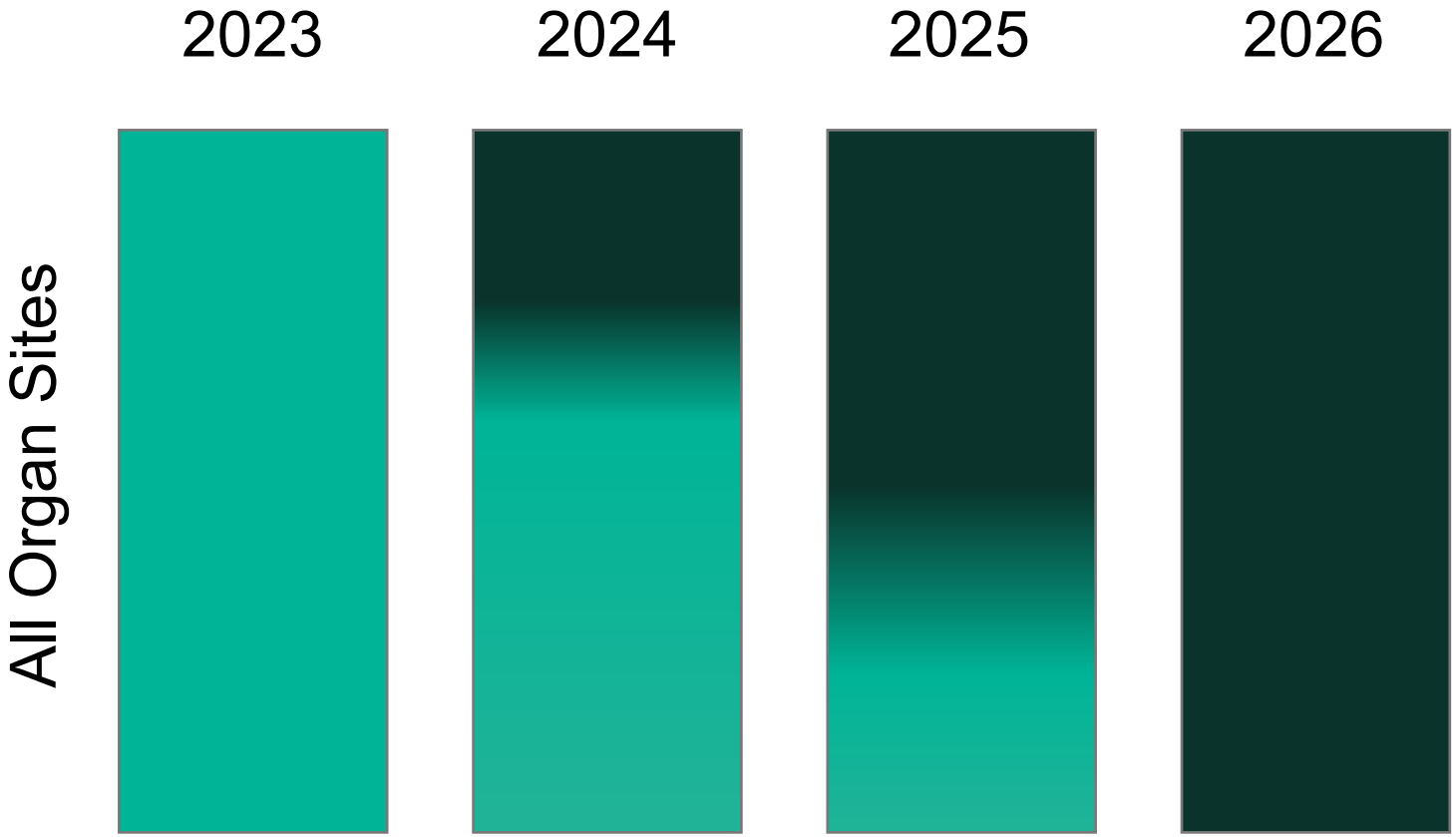
Priya Nagarajan
Dermatopathology

Kay Washington
GI Pathology

Implementation Timeline

Cancer PathCHART Updated Standards

Previous Standards



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

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

Home / Registry Operations / Cancer PathCHART

Cancer PathCHART – Tumor Site–Morphology Surveillance Standards Initiative

Last Updated: September 6, 2023

What Is Cancer PathCHART?

Cancer Pathology Coding Histology And Registration Terminology (Cancer PathCHART) is a first-of-its-kind initiative in North America and around the world to update cancer surveillance standards for tumor site, histology, and behavior code combinations and associated terminology.

Why Is It Needed?

- The foundational data items of site, histology, and behavior are the basis for all subsequent data abstraction for a tumor (e.g., stage, treatment, outcomes).
- Accurate data are essential for the evaluation, management, research, and surveillance of cancer patients.

What Will Its Impact Be?

This vital online resource will help cancer registrars, clinicians, pathologists, researchers, and developers use the same terms and coding standards, making cancer surveillance more accurately reflect medical practice without altering cancer registration workflows, all to better support the critical data necessary for public health monitoring and cancer research.

Key Collaborators

- National Cancer Institute - Surveillance, Epidemiology, and End Results Program (NCI – SEER)
- National Cancer Registrars Association (NCRA)
- North American Association of Central Cancer Registries (NAACCR)
- Centers for Disease Control and Prevention National Program of Cancer Registries (CDC – NPCR)
- International Association of Cancer Registries (IACR)
- Statistics Canada | Statisique Canada
- World Health Organization - International Agency for Research on Cancer (WHO – IARC)
- College of American Pathologists (CAP)
- American Joint Committee on Cancer (AJCC)
- American College of Surgeons - Commission on Cancer (ACS – CoC)
- International Collaboration on Cancer Reporting (ICCR)

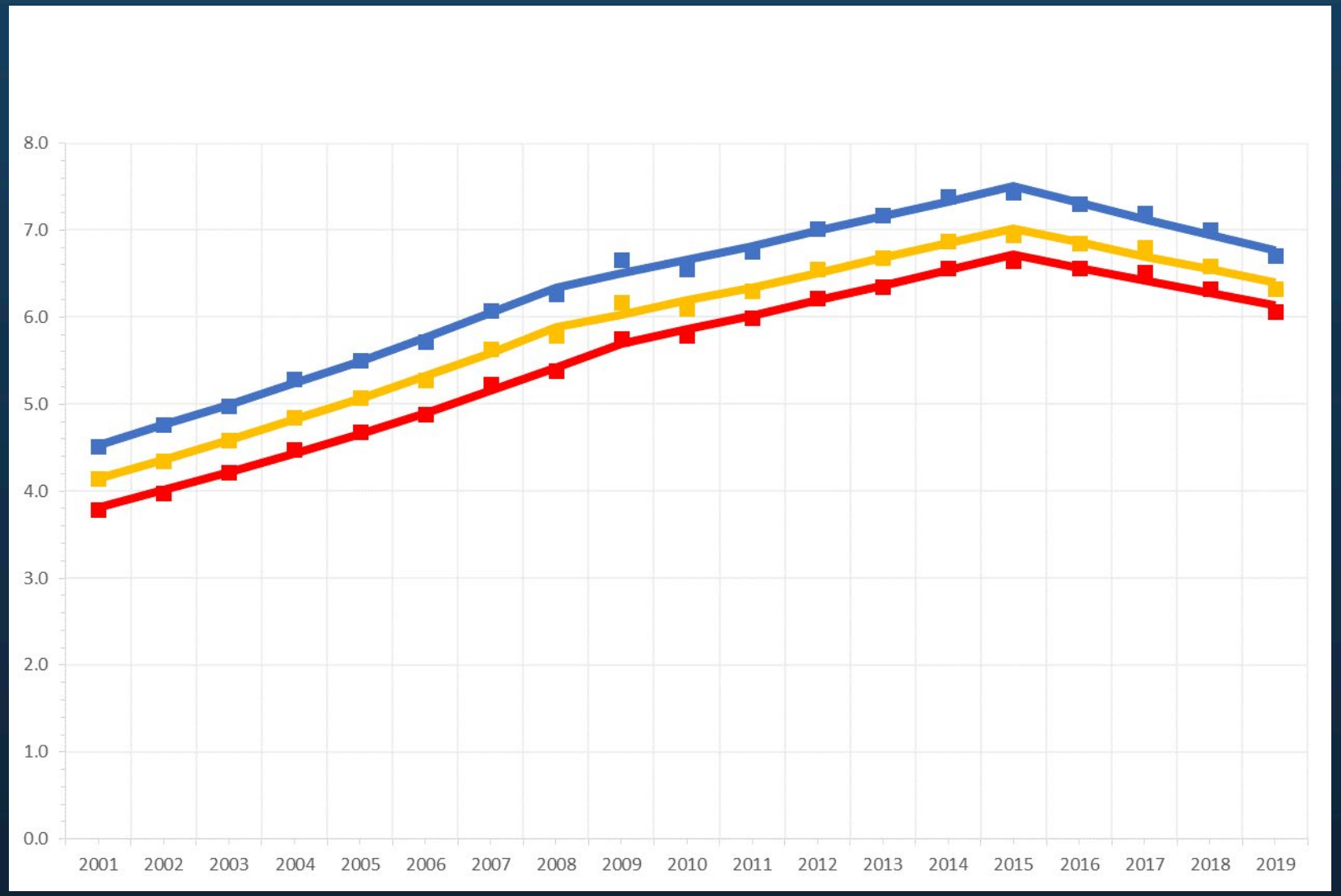
Previously Valid Combinations: Impossible as of 2024

Site Group	Code	ICD-O-3.2 Preferred Term	USCS Count 2015–2019	Consensus Comment
Endometrium	8460/3	Low grade serous carcinoma	3,486	Use 8441/3 (Serous carcinoma, NOS)
Liver	8160/3	Cholangiocarcinoma	2,998	Code to IHBD if clinical fits Biologically impossible in liver
Liver	8140/3	Adenocarcinoma, NOS	2,373	Code to IHBD if clinical fits Biologically impossible in liver
Prostate	8550/3	Acinar cell carcinoma	2,167	Use code 8140/3 (Acinar adenocarcinoma of prostate)
Breast	8510/3	Medullary carcinoma, NOS	1,224	Use 8500/3 if there's no additional characterization Not being used anymore

Liver

USCS (2001-2021)

Age Adjusted Incidence Rate per 100,000



Legend

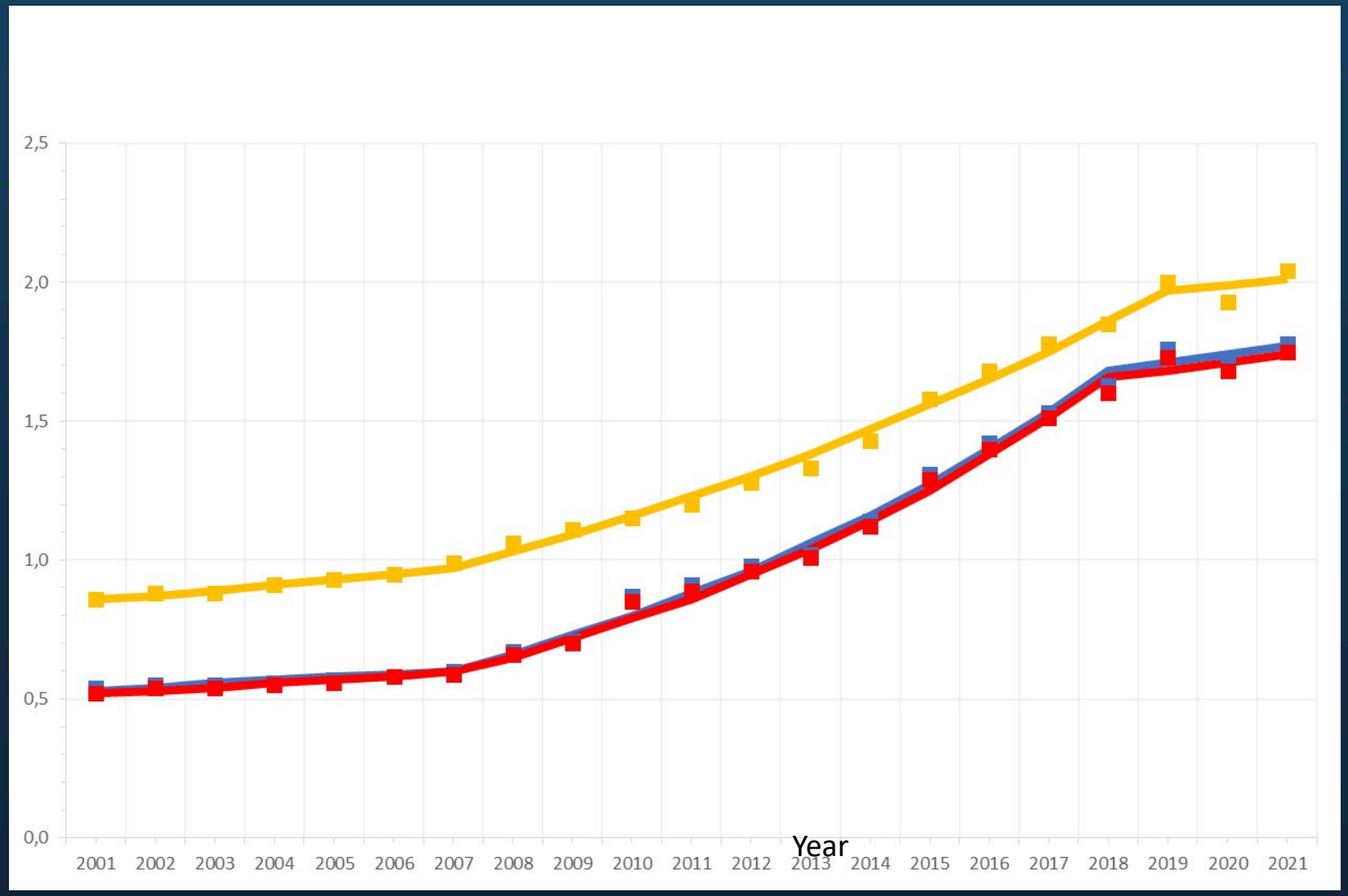
Pre-CPC valid & unlikely

CPC w/ 8140/3 & 8160/3

CPC w/o 8140/3 or 8160/3

Intrahepatic Bile Ducts

USCS (2001-2021)



Take Home Messages

Surveillance gold standards need to align with medical practice

Prioritize completeness using the correct gold standard for accuracy

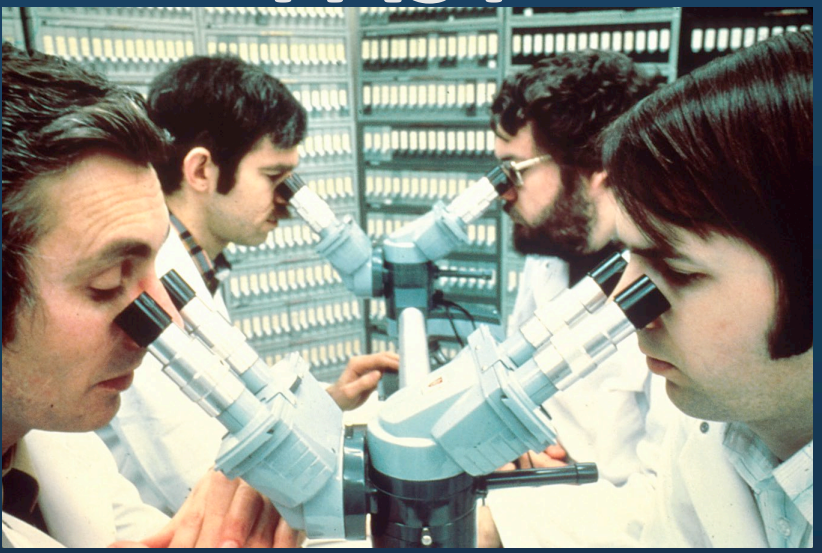
Increased eCP synoptic reporting ensures cancer registry accuracy of critical data items

Timelier & more accurate data acquisition & reporting feasible with

- e-pathology report submission
- data linkages
- autocoding through NLP algorithms

Both registrars & pathologists are becoming the great integrators of automated data

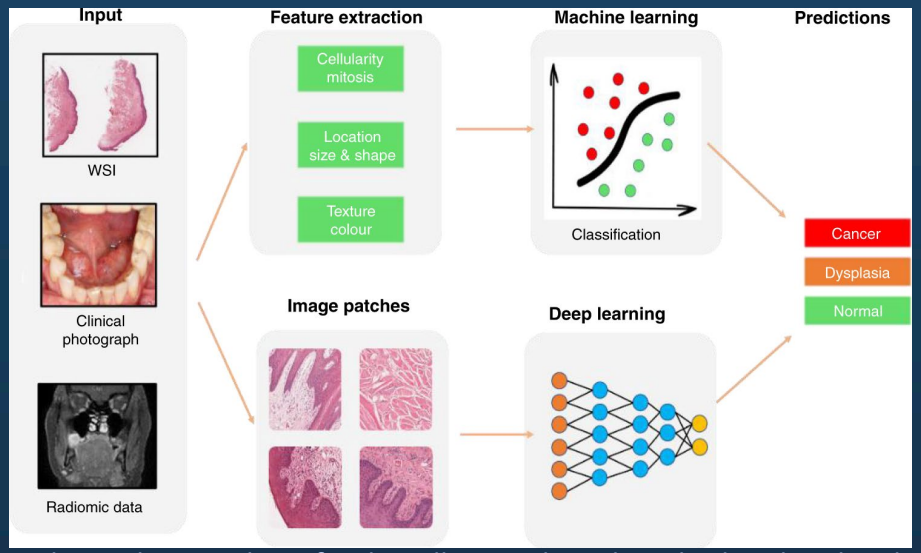
PAST



Linda Bartlett. 1980. NCI Visuals Online Pathologists Looking into Microscopes. Jan 1, 2001

Document Management n.d. https://documentarchiving.com/medical-records-storage/.

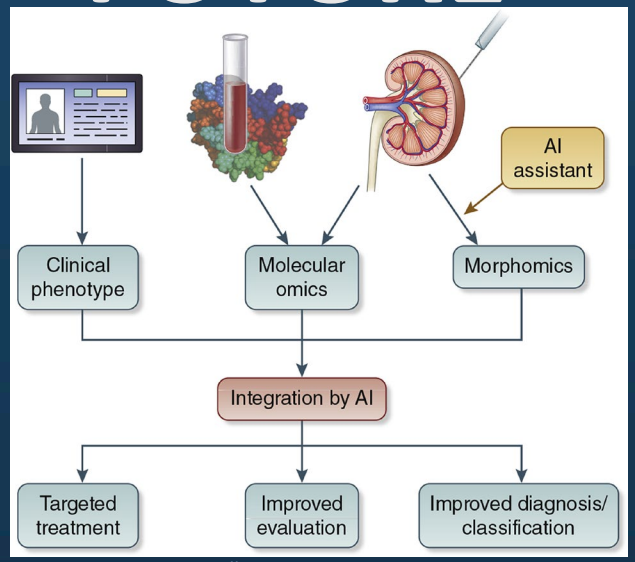
PRESENT



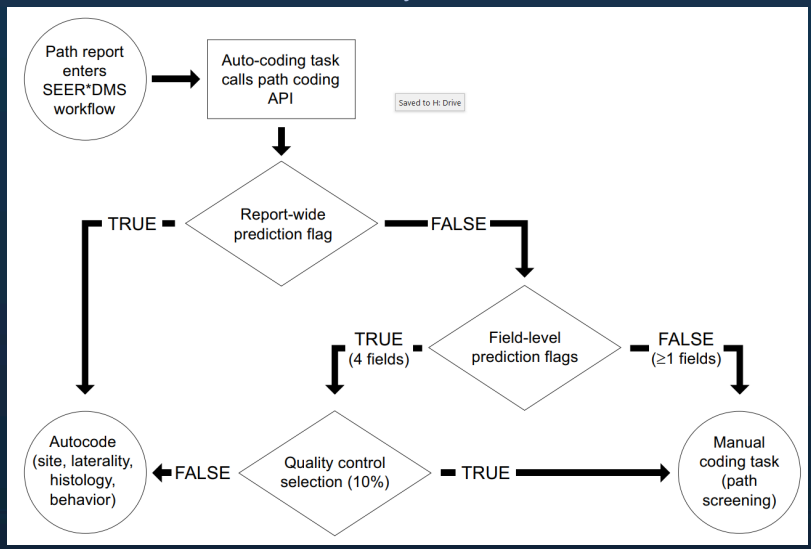
Mahmood, H, *et al.* Artificial Intelligence-based methods in head and neck cancer diagnosis: an overview. *Br J Cancer* **124**, 1934–1940 (2021). https://doi.org/10.1038/s41416-021-01386-x.

National Cancer Registrars Association. n.d. https://www.ncra-usa.org/portals/68/Images/cancer-registry-eduction.jpg.

FUTURE



Huo, Yuankai et al. "AI applications in renal pathology." *Kidney international* vol. 99,6 (2021): 1309-1320. doi:10.1016/j.kint.2021.01.015



Hsu, Elizabeth et al. "Machine learning and deep learning tools for the automated capture of cancer surveillance data." *Journal of the National Cancer Institute. Monographs* vol. 2024,65 (2024): 145-151. doi:10.1093/jncimonographs/lgae018

For More Information

alison.vandyke@nih.gov

Visit the Cancer PathCHART website & CPC*Search tool
https://seer.cancer.gov/cancerpathchart/

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

?s

Submit all Cancer PathCHART questions to Ask a SEER Registrar
Select Cancer PathCHART
https://seer.cancer.gov/registrars/contact.html