SEER*Stat Tools Webinar Series

Overview of the April 2025 SEER Data Release

Mandi Yu, PhD Valentina Petkov, MD MPH Anne-Michelle Noone, PhD

Panelists:







Updates on SEER Race Recode and **Population Data**

- Mandi Yu, PhD

Agenda

New Data Fields

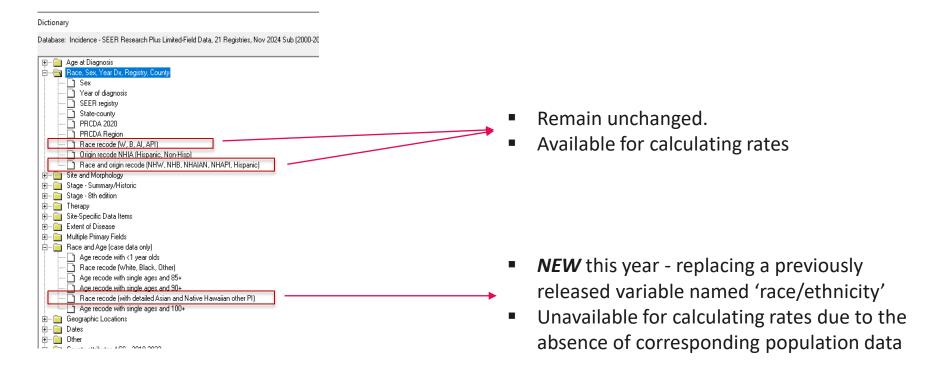
Valentina Petkov, MD, MPH

Overview of data release and new trends

- Anne-Michelle Noone, PhD

SEER Race Recode Update

Race recodes in SEER Research & Research+ databases



More about this new race recode

- It features detailed Asian American (AA) and Native Hawaiian and other Pacific Islander (NHoPI) races
- For AA and NHoPI races, two significant improvements* were made
 - Improved granularity and completeness NOS** cases were reassigned with a specific race using the NAPIIA algorithm
 - Less populous races are grouped with "Other Asian" or "Other PI" to avoid disclosure.
- Non-AA and Non-NHoPI races remain unchanged

- * For details, refer to https://seer.cancer.gov/seerstat/variables/seer/race ethnicity/
- ** NOS=Not Otherwise Specified

Asian American – SEER 21 excl. AK, Nov 2024 sub. 2000-2022

			New "Ra	ce recode	(with detaile	ed Asian and I	Native Ha	waiian other Pl)"	
					Asian				Other	
Durant and IID and feet at the II	Ola i sa a a a		- ::::		Indian,	\ <i>G</i> = t =		V	Asian	T - 4 - 1
Previous "Race/ethnicity"		Japanese	Filipino	Korean	Pakistani	Vietnamese	Laotian	Kampuchean	American	Total
Chinese	126,850									126,850
Japanese		76,667								76,667
Filipino			137,340							137,340
Korean (1988+)				46,186						46,186
Asian Indian or Pakistani, NOS (1988+)					17,446					17,446
Asian Indian (2010+)					37,203					37,203
Pakistani (2010+)					2,897					2,897
Vietnamese (1988+)						48,260				48,260
Laotian (1988+)							4,405			4,405
Kampuchean (1988+)								5,460		5,460
Hmong (1988+)									1,471	1,471
Thai (1994+)									5,223	5,223
Other Asian (1991+) NOS	15,169	5,759	9,769	4,099	4,464	4,767	3	139	31,365	75,534
Total	142,019	82,426	147,109	50,285	62,010	53,027	4,408	5,599	38,059	584,942
% of Reassigned	10.7%	7.0%	6.6%	8.2%	7.2%	9.0%	0.1%	2.5%		

# of cases reassigned:	44,169
# of remaining Asian NOS:	31,365
% of reassigned:	58.5%



Native Hawaiian and other Pacific Islander (NHoPI)—SEER 21 excl. AK, Nov 2024 sub. 2000-2022

	New "Race recode (with detailed Asian and Native Hawaiian other PI)"					
Previous "Race/ethnicity"	Hawaiian	Guamanian/Chamorro	Samoan	Other Pacific Islander	Total	
Hawaiian	33,470				33,470	
Chamorran (1991+)		272			272	
Guamanian, NOS (1991+)		1,372			1,372	
Samoan (1991+)			5,496		5,496	
Micronesian, NOS (1991+)				1,750	1,750	
Polynesian, NOS (1991+)				387	387	
Tahitian (1991+)				96	96	
Tongan (1991+)				1,706	1,706	
Melanesian, NOS (1991+)				114	114	
Fiji Islander (1991+)				993	993	
New Guinean (1991+)				39	39	
Pacific Islander, NOS (1991+)	289	113	141	5,393	5,936	
Total	33,759	1,757	5,637	10,478	51,631	
% of Reassigned	0.9%	6.4%	2.5%			

# of cases reassigned:	543
# of remaining PI NOS:	10,478
% of assigned:	4.9%



Population Data Update



Census 2020, Vintage 2023 and Updated Intercensal 2010-2019

- Two new annual bridged race county population estimates
 - 1. Postcensal Vintage 2023 (2020-2023) from Census Bureau through an IAA
 - 2. Updated intercensal 2010-2019 from **Woods & Poole (W&P)** through a contract

Note: Census Bureau delayed the production of 2010-2019 to fall 2025

- Both used the newly released 2020 "blended base" estimates as included in Vintage 2023 files
 - As the starting point for the postcensal 2020-2023 time series
 - As the ending point for the intercensal 2010-2019 time series
- It is to ensure trends not biased by changes made to population data

What changes were made in 2020 "Blended base" for V23?

- "Blended base" integrates information from both 2010 and 2020
- V22 from last year (Vintage 2022 and previous W&P's intercensal)
 - Age and Sex: 2020 demographic analysis (DA) estimates* (note: DA is independent of census and is developed from vital, migration & Medicare records)
 - Race and Hispanic origin: projections from 2010 census
- V23 from this year (Vintage 2023 and updated W&P's intercensal)
 - Sources of age, sex and race are unchanged.
 - Hispanic origin: 2020 census

Updates to Age Adjustment for the 85+ Population

- The default age groups used for age standardization have changed from 19 to 20 groups, with ages 85+ divided into 85–89 and 90+
- Standard populations used for age-adjustment remain unchanged
- Rationale for the change:
 - As life expectancy increases, the 85+ population continues to grow, exhibiting significant variations in health risks.
 - Dividing this older population into finer age groups enables more precise agestandardization.
- More information: https://seer.cancer.gov/stdpopulations/

Bridged race and future changes

- Bridged race enables trend analysis dating back to 1990
 - NCHS stopped producing bridged race populations in 2020
 - NCI/SEER has an agreement with Census Bureau to produce these through data year 2025
 - Note that bridged race is not compliant to OMB 1997 recommendations
- Research to calculate trends by alone-race and in-combination race is currently underway at NCI
- Future alone-race or in-combination race features the distinction of Asian American (AA) from Native Hawaiian and other Pacific Islander (NHoPI)

Comparisons of populations with previous estimates

- Time-series graphs for 50 states, DC, SEER registries, and SEER registry groupings
- Estimates compared include (<u>https://seer.cancer.gov/endofdecade-pops/</u>)

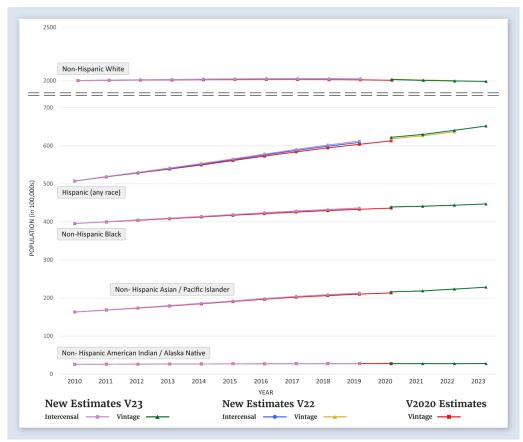
Previous Estimates:

- 1. 2011-2020 populations projected forward from Census 2010 -- Vintage 2020
- V22 estimates
 - 1) Intercensal 2011-2019 anchored to Vintage 2022 -- Intercensal V22
 - 2) 2020-2022 populations projected forward from Census 2020 --Vintage 2022

Current Estimates: V23 estimates

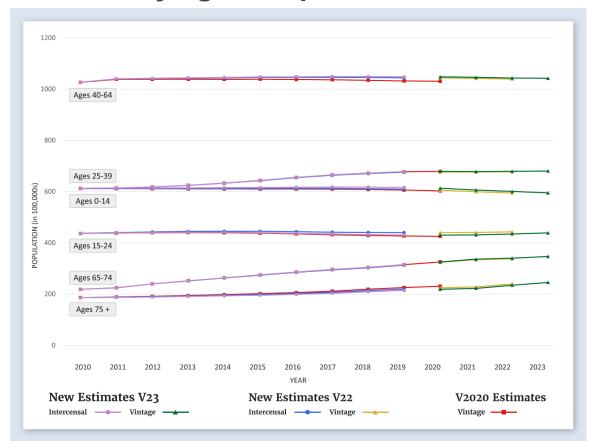
- 1) Intercensal 2011-2019 anchored to Vintage 2023 -- Intercensal V23
- 2) 2020-2023 populations projected forward from Census 2020 -- Vintage 2023

Total U.S. by Race/Ethnicity



- Revised Upward
 - Hispanic
- This reflected the change in 2020 'blended base', where Hispanic data is updated using census 2020 in Vintage 2023

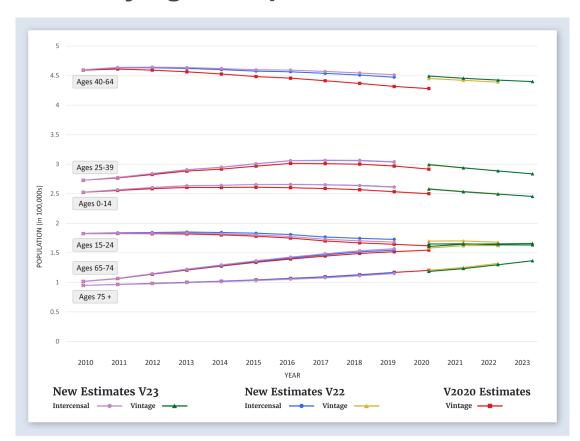
Total U.S. by Age Group



Key observations:

- Revised Downward (by ~3%)
 - o Ages 15-24
 - o Ages 75+
- Revised Upward
 - Ages 0-14 (more salient)
 - Remaining groups with smaller magnitudes
- County and census tract revisions could be greater.

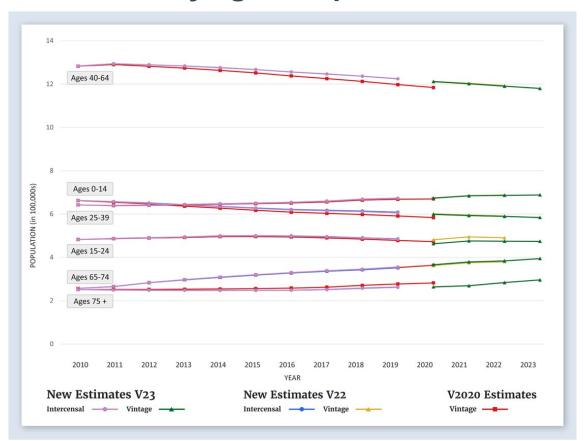
Hawaii by Age Group



Revised Downward

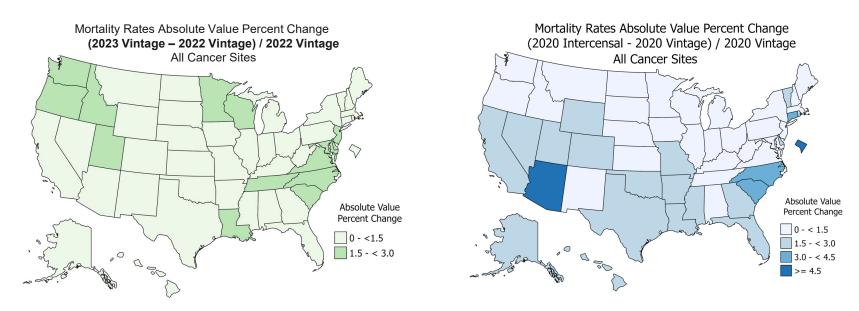
- o Ages 15-24
- o Ages 75+
- Revised Upward
 - Ages 40-64
 - Ages 65-74

Connecticut by Age Group



- Revised Downward
 - Ages 15-24

Impacts of population revisions on 2020 mortality rates by State



Maps also available for rates of All causes of death, Breast, Colon & rectum, Lung & bronchus, and Prostate https://seer.cancer.gov/endofdecade-pops/

SEER*Stat Tools Webinar Series

New Data Fields

Valentina Petkov, MD, MPH Surveillance Research Program

New Data Fields



Multiple Myeloma Terminology Recode (2012+)

- Multiple myeloma (MM) and smoldering/ asymptomatic myeloma are coded with the same ICD-O-3.2 code of 9732/3
- The new data field allows to distinguish MM from smoldering or asymptomatic myeloma which have different prognosis and implications for the Revised International Staging System (RISS) for MM
- Combines two data fields
 - Site Specific Factor, SSF, (2012-2017) 3 for Collaborative Stage (CS) data collection system, schema Plasma Cell Disorders including MM
 - Site Specific Data Item (SSDI) Schema Discriminator 1: Multiple Myeloma Terminology (NAACCR #3926) in Plasma Cell Myeloma, Schema ID 00821
 - Available in Research and Research Plus except limited fields DBs

Code Label

00 Multiple myeloma

- Myeloma, NOS
- Non-secretory myeloma
- Plasma cell myeloma (PCM)
- Ultra-High-Risk Smoldering MM (SMM)
- Smoldering plasma cell myeloma (SPCM)
 - Asymptomatic plasma cell myeloma
 - Early myeloma
 - Evolving myeloma
- Other terminology describing myeloma
 - Unknown terminology used

Heritable Trait Recode (2018+) for Retinoblastoma

- SSDI (NAACCR #3856) in Schema ID 00680
- Captures both
 - Clinical features associated with heritable mutation in RB1 gene (without testing)
 - RB1 germline mutation
- Clinical relevance:
 - Included in staging
 - Prognostic
 - Implications for management
- Recoded to indicate "Not available" for IL and TX
- Available in all SEER Research Plus DBs including limited fields DBs

Code	Label
00	H0: Normal RB1 alleles No clinical evidence of mutation
01	H1: RB1 gene mutation OR Clinical evidence of mutation
07	Test ordered, results not in chart
09	HX: Not documented; No/unknown test; Insufficient evidence
12	Not available (IL 2018+)
13	Not available (TX 2018-2021)

PSA (Prostatic Specific Antigen) Lab Value (2018+)

- SSDI (NAACCR #3920) in Schema ID 00580
- PSA: protein produced by the cells of prostate gland; used in staging and treatment monitoring
 - Captures detailed values up to 999.9 ng/ml and >/=1000 ng/ml
 - Available in all SEER Research and Research Plus DBs including limited fields DBs
- PSA available in the PSA Lab Value Recode (2010+)
 - Detailed values up to 97.9 ng/ml and >/=98.0 ng/ml

Code	Label
00001	0.1 or less nanograms/milliliter (ng/ml)
00002- 09999	0.2 - 999.9 ng/ml
10001	1,000 ng/ml or greater
10002	Lab value not available, physician states PSA is negative/normal
10003	Lab value not available, physician states PSA is positive/elevated/high
10007	Test ordered, results not in chart
10009	Not documented; not assessed; unknown

Bilirubin Pretreatment Total Lab Value Recode (2018+) in Liver Cancer

- SSDI (NAACCR #3813) for Schema ID 00220
- Prognostic serologic marker
- Recoded to indicate "Not available" for IL and TX
- Available in all SEER Research Plus DBs including limited data fields DBs
- Values are associated with different units of measurement
 - Use with Bilirubin Pretreatment Unit of Measure Recode (2018+) to convert

Code	Label
0000	0.0 milligram/deciliter (mg/dL) 0.0 micromole/liter (umol/L)
00000- 09999	0.1-999.9 milligram/deciliter (mg/dL) 0.1-999.9 micromole/liter (umol/L)
10000	1000 milligram/deciliter (mg/dL) or greater 1000 micromole/liter (umol/L) or greater
10001	Test ordered, results not in chart
10002	Not applicable: Information not collected for this case
10003	Not documented; Not assessed or unknown if assessed
10013	Not available (IL 2018+)
10014	Not available (TX 2018-2021)

Bilirubin Pretreatment Unit of Measurement Recode (2018+) in Liver Cancer

- SSDI (NAACCR #3814) in Schema ID 00220
- Specifies unit of measurement for bilirubin
- Use to convert Bilirubin Pretreatment Total Lab Value
 - E.g. to convert umol/L in mg/dL divide by a conversion factor of 17.1
- Recoded to indicate "Not available" for IL and TX registries
- Available in all SEER Research Plus DB including limited fields DBs

Code	Label
01	Milligrams per deciliter (mg/dL)
02	Micromoles/liter (umol/L)
07	Test ordered, results not in chart
80	Not applicable: Information not collected for this case
09	Not documented; Unit of measure not assessed or unknown if assessed
12	Not available (IL 2018+)
13	Not available (TX 2018-2021)

New Fields Released in October 2024

Data field name in SEER*Stat	Schema ID	Clinical utility
Microsatellite Instability (MSI) (2018+)	Colon and Rectum (200)	Prognostic and predictive biomarker (BM); Lynch syndrome
KRAS Recode (2018+)	Colon and Rectum (200)	Prognostic and predictive BM
	Oropharynx HPV-Mediated (p16+) (100);	
Schema discriminator 2: Oropharyngeal p16 (2018+)	Oropharynx (p16-) (111)	Staging, prognostic BM
Methylation of O6-Methylguanine-Methyltransferase Recode (MGMT) (2018+)	Brain (721); CNS Other (722)	Prognostic and predictive BM
KIT Gene Immunohistochemistry Recode (2018+)	GIST (430)	Diagnostic
JAK2 Recode (2018+)	HemeRetic (830)	Diagnostic and predictive BM; monitor response to treatment
High Risk Cytogenetics Recode (2018+)	Plasma Cell Myeloma (821)	Staging, prognostic BM
Sentinel Lymph Nodes Examined Recode (2018+)	Breast (480); Melanoma Skin (470)	Prognosis, treatment
Sentinel Lymph Nodes Positive Recode (2018+)	Breast (480); Melanoma Skin (470)	Prognosis, treatment

KRAS Recode (2018+) in Colorectal Cancer

- SSDI (NAACCR #3866) in Schema ID 00200
- KRAS is a signaling protein in growth receptor pathway – cell proliferation and survival
- Mutations in K-ras gene indicate unlikely response to anti-EGFR therapy and are associated with worse prognosis
- Recoded to indicate that IL and TX registries did not collect it
- Available in all Research Plus DB including limited fields DB

Code	Label
00	Normal (wild type); Negative for mutations,
01	Abnormal (mutated) in codon(s) 12, 13 and/or 61
02	Abnormal (mutated) in codon 146 only
03	Abnormal (mutated), but not in codon(s) 12, 13, 61, or 146
04	Abnormal (mutated), NOS, codon(s) not specified
07	Test ordered, results not in chart
80	Not applicable: Information not collected for this case
09	Not documented; KRAS not assessed or unknown if assessed
12	Not available (IL 2018+)
13	Not available (TX 2018-2021)

Microsatellite Instability (MSI) (2018+) in Colorectal Cancer

- SSDI (NAACCR #3890) for Schema ID 00200
- MSI is a molecular alteration in which short tandem repeats (microsatellites) are prone to accumulating mutations in DNA
- Due to DNA mismatch-repair (MMR) system deficiency because of germline/somatic mutations in the MMR genes
- MSI-High or MMR-Deficient are associated with better prognosis and have treatment implications
- Available in all SEER Research Plus DB including limited field DBs

Code	Label
00	Microsatellite instability (MSI) - stable; microsatellite stable (MSS); Negative and/or Mismatch repair intact/proficient (MMR-P)
01	MSI unstable low (MSI-L)
02	MSI unstable high (MSI-H) and/or Mismatch repair deficient (MMR-D)
08	Not applicable: Information not collected for this case
09	Not documented in medical record MSI-indeterminate; Not assessed or unknown if assessed

Schema Discriminator 2: Oropharyngeal p16

- SSDI (NAACCR #3927) in Schema IDs 00110 and 00111
- Captures p16 testing of oropharyngeal tumors
- Used in staging
- HPV-associated oropharyngeal cancers have better prognosis
- Available in all Research Plus DBs including limited fields DBs

Code	Label	Schema ID #/Description
01	p16 Negative	00111: Oropharynx (p16-)
02	p16 Positive;	00100: Oropharynx HPV- Mediated (p16+)
09	p16 Unknown	00111: Oropharynx (p16-)

Methylation of O6-Methylguanine-Methyltransferase Recode (2018+) for Selected Brain Cancers

- SSDI (NAACCR #3889) in Schema ID 00721 and 00722
- MGMT is an enzyme that repairs DNA
 - Could decrease cytotoxic effect of chemo
- Methylation of the MGMT gene leads to epigenetic silencing of the gene and decreased production of the enzyme
 - Associated with better response to alkylating agents and nitrosoureas
- Usually assessed in high grade glioma and astrocytoma
- Available in all SEER Research Plus DBs including limited field DBs

Code	Label	
00	MGMT methylation absent/not present, unmethylated MGMT	
01	MGMT methylation present, low level Hypomethylated; Partial methylated	
02	MGMT methylation present, high level Hypermethylated	
03	MGMT methylation present, level unspecified	
06	Benign or borderline tumor	
07	Test ordered, result not in chart	
08	Not applicable: Information not collected for this case	
09	Not documented; Pathologist can't determine Not assessed or unknown if assessed	
12	Not available (IL 2018+)	
13	Not available (TX 2018-2021)	

KIT Gene Immunohistochemistry (IHC) Recode (2018+) in Gastrointestinal Stromal Tumors (GIST)

- SSDI (NAACCR #3865) for Schema ID 00430
- KIT gene regulates cell growth and differentiation
- Gene mutation causes GIST to escape cellular control signals
- Diagnostic (85-90% of GIST have oncogenic mutations)
- Predict response to therapy with Gleevec or Sutent
- Available in all SEER Research Plus DBs including limited field DBs

	,
Code	Label
00	KIT negative/normal; within normal limits
01	KIT positive
07	Test ordered, results not in chart
80	Not applicable: Information not collected for this case
09	Not documented; Cannot be determined by pathologist KIT not assessed or unknown if assessed
12	Not available (IL 2018+)
13	Not available (TX 2018-2021)

JAK2 Recode (2018+) in Myeloproliferative Neoplasms (MPN)

- SSDI (NAACCR #3862) for Schema ID 830
- Janus Kinase (JAK)2 gene is involved in the development of blood cells
- JAK2 gene mutation is found in MPN (such as polycythemia vera, essential thrombocythemia, and primary myelofibrosis)
- Used in diagnosis, treatment choices and response to treatment monitoring
- Recoded to indicate IL and TX registries did not collect it
- Available in all Research Plus DB including limited fields DB

Code	Label
00	JAK2 result stated as negative
01	JAK2 positive for mutation V617F WITH or WITHOUT other mutations
02	JAK2 positive for exon 12 mutation
03	JAK2 positive for other specified mutation
04	JAK2 positive for more than one mutation other than V617F
05	JAK2 positive NOS Specific mutation(s) not stated
07	Test ordered, results not in chart
80	Not applicable: Information not collected for this case
09	Not documented; Not assessed or unknown if assessed
12	Not available (IL 2018+)
13	Not available (TX 2018-2021)

High Risk Cytogenetics Recode (2018+) in Plasma Cell Myeloma

- SSDI (NAACCR #3857) for Schema ID 00821
- Defined as one or more of several chromosomal alterations (translocation t(4;14), t(14;16) or deletion 17p)
- Used in RISS staging along with Serum beta-2 macroglobulin, serum albumin, and serum LDH
- The presence of high-risk cytogenetics is associated with worse prognosis
- Available in all SEER Research Plus DBs including limited fields DBs

Code	Label
00	High-risk cytogenetics not identified/not present
01	High-risk cytogenetics present
05	Schema Discriminator 1: Plasma Cell Myeloma Terminology coded to 1 or 9
07	Test ordered, results not in chart
09	Not documented; Not assessed or unknown if assessed
12	Not available (IL 2018+)
13	Not available (TX 2018-2021)

New data items for Research Plus 21 registries limited fields database

- These fields have been available in Research and Research Plus SEER 8, 12 and 17 registries
- Some registries did not collect all of these data items for all years
 - Recoded to note "Not available" for the registries and years not collected
 - Allows researchers to exclude those cases from analyses

Data field name in SEER*Stat	Primary Site/ Schema	
Lymph nodes examined (1988+)	All sites	
Lymph nodes positive (1988+)	All sites	
Tumor Size Summary (2016+) #	All sites	
Derived EOD 2018 T (2018+)	All sites	
Derived EOD 2018 N (2018+)	All sites	
Derived EOD 2018 M (2018+)	All sites	
Derived EOD 2018 Stage Group (2018+)	All sites	
EOD Primary Tumor (2018+)	All sites	
EOD Regional Nodes (2018+)	All sites	
EOD Mets (2018+)	All sites	
CEA Pretx Interpretation Recode (2010+)	Appendix, Colon and Rectum	
Chromosome 1p status recode (2010+)	Brain and CNS Other	
Chromosome 19q status recode (2010+)	Brain and CNS Other	
Methylation of O6-Methylguanine- Methyltransferase (2018+)	Brain and CNS Other	
Sentinel Lymph Nodes Examined (2018+)	Breast and Melanoma Skin	
Sentinel Lymph Nodes Positive (2018+)	Breast and Melanoma Skin	
KRAS (2018+)	Colon and Rectum	
Perineural Invasion recode (2010+)	Colon and Rectum	
CA-125 Pretreatment Interpreation Recode (2010+)	Fallopian tube, Ovary, PPC	

(2018+)	GIST	
JAK2 (2018+)	Heme	
Invasion Beyond Capsule Recode (2010+)	Kidney Parenchyma	
AFP PreTX Interpretation Recode (2010+)	Liver	
Separate Tumor Nodules Ipsilateral Recode (2010+)	Lung	
Visceral and Parietal Pleural Invasion Recode (2010+)	Lung	
Measured basal diameter recode (2010+)	Melanoma choroid, ciliary body, Iris	
Measured thickness recode (2010+)	Melanoma choroid, ciliary body, Iris	
Breslow Thickness Recode (2010+)	de (2010+) Melanoma Skin	
Ulceration recode (2010+)	Melanoma Skin	
Mitotic rate recode (2010+)	Melanoma Skin	
High Risk Cytogenetics (2018+)	Plasma Cell Myeloma	
PSA value recode (2010+) #	Prostate	
Gleason pattern clinical recode (2010+) Prostate		
Gleason pattern pathological recode (2010+)	Prostate	
Gleason score clinical recode (2010+)	Prostate	
Gleason score pathological recode (2010+)	Prostate	
Number of cores examined recode (2010+)	Prostate	
Number of cores positive recode (2010+)	Prostate	

KIT Gene Immunohistochemistry

Primary Site/ Schema

Available in Research SEER 21 registers

Overview of data release and new trends

Annie Noone, PhD

2025 SEER data release

- Released April 16, 2025
 - Includes diagnoses 1975 to 2022

- 2020 data point
 - Excluded from Joinpoint trends and risks of developing cancer (DevCan)
 - Included in all other statistics and SEER databases

https://seer.cancer.gov/data/covid-impact.html

How to access SEER data

Cancer Stat Facts

https://seer.cancer.gov/statfacts/

Did you know? Video series

https://seer.cancer.gov/statistics/videos/

How to access SEER data

SEER*Explorer

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

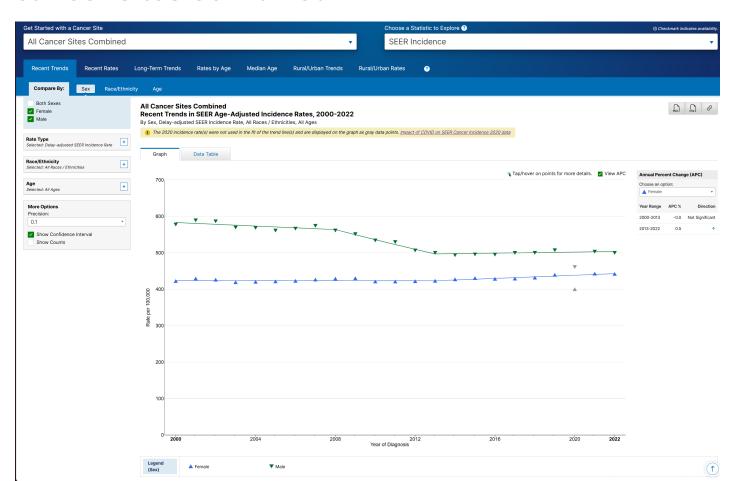
New feature: Rates and trends by rural/urban status

SEER*Stat

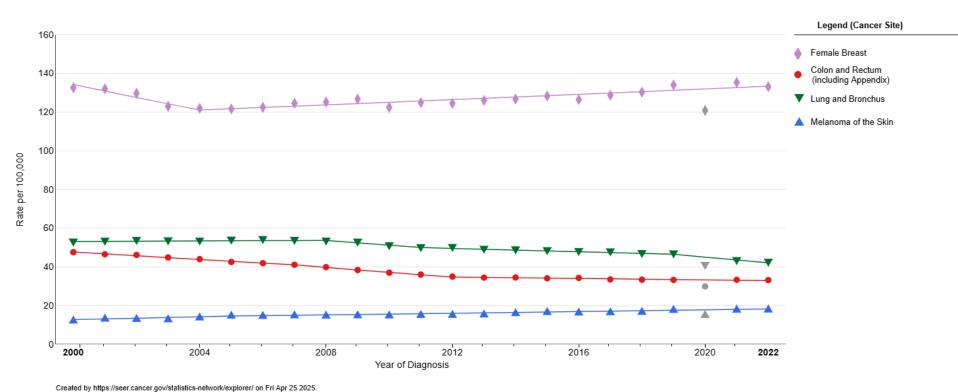
https://seer.cancer.gov/seerstat/

- Request data: https://seer.cancer.gov/data/access.html
- Data policy: https://seer.cancer.gov/data/access-policy.html

All cancer sites combined

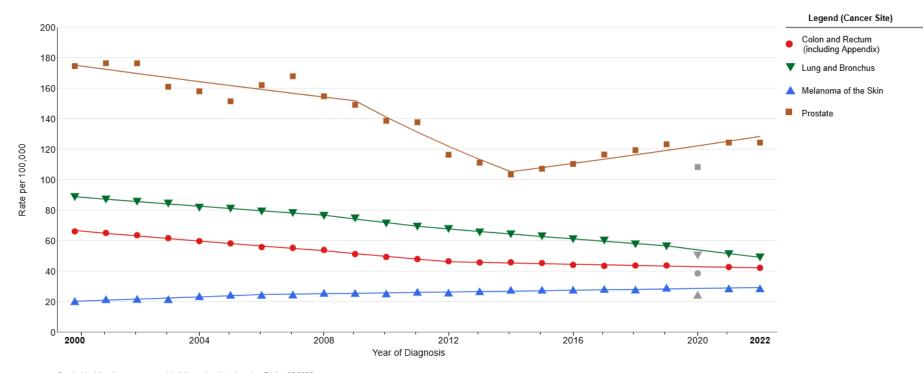


Incidence trends for women, 2000-2022



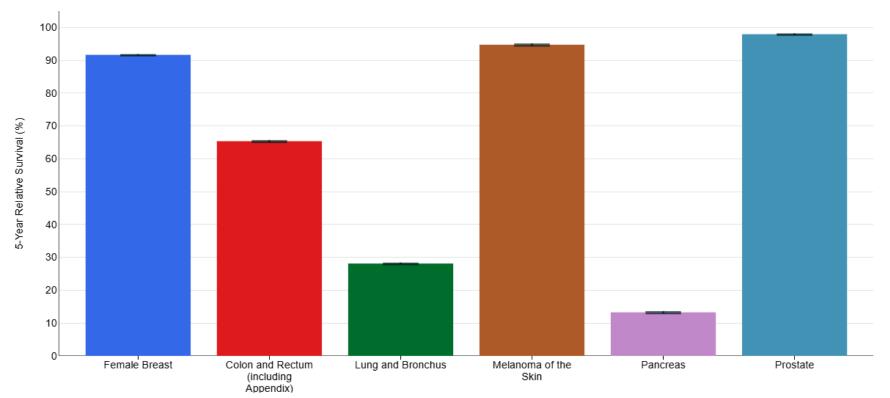
41

Incidence trends for men, 2000-2022



Created by https://seer.cancer.gov/statistics-network/explorer/ on Fri Apr 25 2025.

5-year relative survival rates, 2015-2021



Created by https://seer.cancer.gov/statistics-network/explorer/ on Wed May 07 2025.

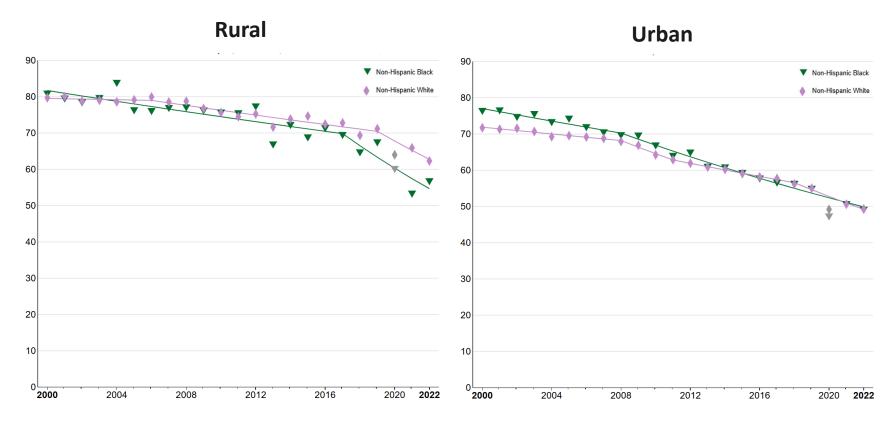


People alive with cancer on January 1, 2022

Cancer Site	Number of people alive with cancer	Percent of U.S. cohort population
All Cancer Sites Combined	18,000,110	5.4
Female Breast	4,091,181	2.4
Colon and Rectum (including Appendix)	1,416,499	0.4
Lung and Bronchus	635,547	0.2
Melanoma of the Skin	1,504,676	0.5
Pancreas	107,988	<0.1
Prostate	3,518,978	2.1

Lung cancer incidence rates by rural/urban status and race

Men and women combined



Data release June 2025

- Mortality estimates with rural/urban status
- Preliminary estimates
 https://seer.cancer.gov/statistics/preliminary-estimates/
- Recorded webinar December 18, 2024
 https://seer.cancer.gov/news/seerstat-webinars.html

Thank You!

https://seer.cancer.gov/news/seerstat-webinars.html



cancer.gov/espanol