Feature Mapping

Paper Specification Text ¹	SEER*Stat text ²	match	reason if dropped, modified, or added	variable type	reason if excluded
Patient ID	Patient ID	yes		removed	irrelevancy
SEER registry		drop	confidentiality, data item not available at individual level	categorical	
Marital status at diagnosis	Marital status at diagnosis	yes		categorical	
Race/ethnicity	Race/ethnicity	yes		categorical	
			Originally coded 0-8 for different hispanic countries. Not available		
NHIA Derived Hisp Origin		drop	in 2023 database.	categorical	
Sex	Sex	yes		categorical	
Age at diagnosis	Age recode with single ages and 90+	mod	originally capped at 130, best available fit is "Single ages and 90+"	continuous	
Year of birth		calc	calculate as (age at diagnosis - year of diagnosis)	continuous	
Sequence number	Sequence number	yes		categorical	
Month of diagnosis		drop	confidentiality, data item not available at individual level	categorical	
Year of diagnosis	Year of diagnosis	yes		continuous	
Primary site ICD-O-2 (1973+)	Primary Site	yes		categorical	
Laterality	Laterality	yes		categorical	
			entire "Site and Morphology - Historic" category. Not available in		
Histologic Type ICD-O-2		drop	2023 database.	categorical	
			entire "Site and Morphology - Historic" category. Not available in		
Behavior Code ICD-O-2*/		drop	2023 database.	categorical	
Histologic Type ICD-O-3	Histologic Type ICD-O-3	yes		categorical	
Behavior code ICD-O-3	Behavior code ICD-O-3	yes		categorical	
Grade	Grade Recode (thru 2017)	yes		categorical	
Diagnostic confirmation	Diagnostic Confirmation	yes		removed	post-dx
Type of reporting source	Type of Reporting Source	yes		categorical	
EOD 10 - size (1988+)	EOD 10 - size (1988-2003)	yes	FYI, This field is only coded for cases diagnosed 1988-2003. For 2004-2015, similar type of information was collected in CS Tumor Size (2004-2015) in the collaborative stage variables. This field is only coded for cases diagnosed 1988-2003. For 2004-2015, similar type of information was collected in CS extension (2004-2015) and CS Mets at DX (2004-2015) in the collaborative	categorical	
EOD 10 - extension	EOD 10 - extent (1988-2003)	ves	stage variables.	categorical	
EOD 10 - path extension	EOD 10 - Prostate path ext (1995-2003)	yes	This field is only coded for prostate cases diagnosed 1995-2003. For 2004+, similar type of information was collected in CS SSF 3 in the collaborative stage variables This field is only coded for cases diagnosed 1988-2003. For 2004-	categorical	
			2015, similar type of information was collected in CS Lymph Nodes	s	
EOD 10 - lymph node	EOD 10 - nodes (1988-2003)	yes	(2004-2015) in the collaborative stage variables.	categorical	
EOD 10 - positive lymph nodes examined	Regional nodes positive (1988+)	yes		continuous	
EOD 10 - number of lymph nodes examined	Regional nodes examined (1988+)	yes		continuous	
EODold 13 digit		drop	compound information, only valid 1973-1982	removed	compound info
EODold 2 digit		drop	compound information, only valid 1973-1982	removed	compound info
EODold 4 digit		drop	compound information, only valid 1973-1982	removed	compound info
Coding system for EOD	Coding system-EOD (1973-2003)	yes		categorical	
Tumor marker 1	Tumor marker 1 (1990-2003)	yes		categorical	
Tumor marker 2	Tumor marker 2 (1990-2003)	yes		categorical	
Tumor marker 3	Tumor marker 3 (1998-2003)	yes		categorical	
CS Tumor size	CS tumor size (2004-2015)	yes		continuous	
CS Extension	CS extension (2004-2015)	yes		categorical	
CS Lymph Nodes	CS lymph nodes (2004-2015)	yes		categorical	
CS Mets at DX	CS mets at dx (2004-2015)	yes		categorical	
CS Site-Specific Factor 1	CS site-specific factor 1 (2004-2017 varying by schema)	yes		categorical	

CS Site-Specific Factor 3	00.014.014.01	100 ti			T	
C3 site sperific Factor 4 C3 site sperific Factor 5 C3 site sperific Factor 6 C3 site specific Factor 6	CS Site-Specific Factor 2	CS site-specific factor 2 (2004-2017 varying by schema)	yes		categorical	
CS 18te Specific Factor 5	•	, , , , ,				
Site Specific Factor 6	·				-	
Same specific Scare 25 Same 25 Same specific Scare 25 Same specific Scare 25 Same spec		, , , , ,				
Deriver ACCT Deriver ACCT Per Potential (CCC, 10 et al (2004-2015) Yes Categorical	•				_	
Derived ACC N	·					
Derived ACC Mg						
Deniver ACC Stage Group		, , ,				
Derived 551977 Derived 551977 See 2016						
Commerced Stations Commerced Summary Stage 2000 (2004-2017) mode Stationary Stage 2000 (2004-2015) stationa	Derived AJCC Stage Group	Derived AJCC Stage Group, 6th ed (2004-2015)	yes		categorical	
Derived ASS/2000 SERR Combined Summary Stage 2000 (2004-2017) mod						
Served S0000 SER Commined Summary Stage 2000 (2004-2017) mode mane as "SER summary stage 2000" categorical	Derived SS1977		drop		categorical	
Derived ALC: - Flags Servicion input original (20 version input original (2004-2015) ves Servicion input original (2004						
CS Version Input Criginal CS Version Imput Compant (2004-2015) Ves CS Version Imput Common (2004-2015) RS SummStope Reg LN Sur (2003-1) Ves CS Version Imput Common (2004-2015) RS SummStope Reg LN Sur (2003-1) Ves CS Version Imput Common (2004-2015) RS SummStope Reg LN Sur (2003-1) Ves CS Version Imput Common (2004-2015) RS SummStope Reg LN Sur (2003-1) Ves CS Version Imput Common (2004-2015) Ves CREC Summ-Sur (2004-2015) Ves CREC Summ-	Derived SS2000	SEER Combined Summary Stage 2000 (2004-2017)	mod	, , ,	categorical	
CS Version Input Current CS Version Input Curr	Derived AJCC - flag		drop	always 1 after 2004. Not available in 2023 database.	categorical	
SS Version Input Current SS Journal (S. Syresion input Current (2004-2015) SS Journal Supprise (1988 - 2003-7) SS Journal	CS Version Input Original	CS version input original (2004-2015)	yes		categorical	
RX Summ-sug prim site RX Summ-sug Prim Site (1984) Ves removed treatment info RX Summ-sug prim Site (1984) Ves removed treatment info RX Summ-sug oth reg/dis (CS Version Derived	CS version derived (2004-2015)	yes		categorical	
MX Summ-scope Reg LM Sur (2003+) Yes	CS Version Input Current	CS version input current (2004-2015)	yes		categorical	
RX Summ-Surg Oth Reg/Dis (2003+) yes removed treatment info Reson no cancer-directed surgery yes removed treatment info Scope of frught had year (1984-1997) site specific surgery 1987-1997 site specific surgery 1987-1997 yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Record (1988-2002) yes removed treatment info Record (1988-2003) yes removed treatment info Reco	RX Summsurg prim site	RX SummSurg Prim Site (1998+)	yes		removed	treatment info
RX Summ-Surg Oth Reg/Dis (2003+) yes removed treatment info Reson no cancer-directed surgery yes removed treatment info Scope of frught had year (1984-1997) site specific surgery 1987-1997 site specific surgery 1987-1997 yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Scope of frught had year (1988-2002) yes removed treatment info Record (1988-2002) yes removed treatment info Record (1988-2003) yes removed treatment info Reco						_
Number of Imph nodes						
Resson no cancer directed surgery Resson no cancer directed surgery Site specific surgery (1923-1997) Record mynh note surgery 98-02 Scope of reg lymph note surgery 98-02 Scope of reg lymph note surgery 98-02 Surgery to that regists sites (1998-2002) Record mynher Record mynher Record mynher recode Groph of a surgery of the regists sites (1998-2002) Record mynher Record mynher recode Groph of a surgery of the regists sites (1998-2002) Recode (CD-0-2 to 9 Recode (CD-0-2 to 9 Recode (CD-0-2 to 9 Recode (CD-0-2 to 10 CCC site recode (CD-0-3/WHO 2008 ICCC site recode a detion/IARC 2017 More of the surgery of the register of the regist	Number of lymph nodes					_
Site specific surgery (1383-1997) Site specific surgery (1381-1997) Site specific surgery (1381-1997) Stope of Imph not surgery 1382-202 Scope of Imph not surgery 1382-202 Scope of Imph not surgery 1382-202 Surgery to other sites Surgery to other sites Surgery to other sites Surgery to other sites Record number Record number Record number recode Recode ICD-0-3/WHO 2008 Site recode ICD-0-3/WHO 2008 Site recode ICD-0-3/WHO 2008 Site recode ICD-0-3/WHO 2008 Recode ICD-0-10	, ,					
Scope of Imply node surgery 8-02 Scope of Teg Imply not surg (1998-2002) yes removed treatment info Surgery to other sydfs sites (1998-2002) yes removed treatment info Surgery to other sydfs sites (1998-2002) yes removed treatment info Surgery to other sydfs sites (1998-2002) yes removed treatment info Surgery of other sydfs sites (1998-2002) yes removed treatment info Surgery of other sydfs sites (1998-2002) yes removed treatment info Surgery of other sydfs sites (1998-2002) yes removed treatment info Surgery of other sydfs sites (1998-2002) yes removed treatment info Surgery of other sydfs sites (1998-2002) yes removed treatment info Surgery of other sydfs sites (1998-2002) yes removed treatment info Surgery of other sydfs sites (1998-2002) yes removed treatment info Surgery of Surgery sydfs sydfs sites (1998-2002) yes removed treatment info Surgery of the Surgery of the Surgery of the Surgery of Surgery sydfs sites (1998-2003) yes removed treatment info Surgery sydfs syd					1	
Surgery to other sites Surgery of ath reg/dis sites (1998-2002) Wes Record number Record number recode Recode Record number recode Record number recode Record number rec						_
Record number Record number recode mod some small changes to order of sequential number removed irrelevancy type of following expected drop Mot available in 2023 database removed irrelevancy mode (CP - 3/WHO 2008 Ster recode ICD - 0-3/WHO 2008 ICC ster recode a detail or ster ster and Morphology - Historic" category dropped. Not available in 2023 database. Categorical enter "Site and Morphology - Historic" category dropped. Not available in 2023 database. Categorical enter "Site and Morphology - Historic" category dropped. Not drop available in 2023 database. Categorical enter site and Morphology - Historic" category dropped. Not drop available in 2023 database. Categorical selection of the control of the co						
Type of followup expected Age recode https://doi.org/10.1008/ Age recode https://doi.org/10.1008/https://doi.org/10.1008/<				some small changes to order of sequential number		_
Age recode <a (cd-o-2="" (cd-o-3="" 2008="" 9<="" <a="" age="" age)="" href="Age recode" recode="" recode"="" td="" to="" who="" with="" =""><td></td><td>Record Hamber recode</td><td>-</td><td>,</td><td></td><td></td>		Record Hamber recode	-	,		
Site recode ICD-O-3/WHO 2008 Sterecode ICD-O-3/WHO 2008 Recode ICD-O-3/WHO 2008 Recode ICD-O-2 to 9 drop available in 2023 database. entire "Site and Morphology - Historic" category dropped. Not available in 2023 database. categorical entire "Site and Morphology - Historic" category dropped. Not available in 2023 database. categorical categorical categorical categorical switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values categorical switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values categorical switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values categorical switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categ		Aga ragada with <1 year alds		Not available III 2023 database		irrelevancy
Recode ICD-O-2 to 9 drop available in 2023 database. categorical	-					
Recode ICD-O-2 to 10 drop available in 2023 database. categorical entire "Site and Morphology - Historic" category dropped. Not available in 2023 database. categorical savilable in 2023 database. categorical validable in 2023 database. categorica	Site recode ICD-O-3/WHO 2008	Site recode ICD-O-3/WHO 2008	yes	autica IICita and Manufalana IIII-taniall actaona durana di Nat	Categorical	
Recode ICD-O-2 to 10 Recode ICD-O-3 twithout the first mail morphology - Historic" category dropped. Not adaptive from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values and the properties of the properties	D I . I CD . O . 2 O					
Recode ICD-O-2 to 10 CCC site recode ICD-O-3/WHO 2008 ICCC site recode and edition/IARC 2017 mod switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2017 definitions, categorical switched from WHO 2008 definitions to IARC 2	Recode ICD-U-2 to 9		arop		categoricai	
Switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values CCC site recode ICD-O-3/WHO 2008 CCC site recode extended 3rd edition/IARC 2017 mod same type of information but different values categorical						
ICCC site recode ICD-O-3/WHO 2008 ICCC site recode and edition/IARC 2017 mod same type of information but different values switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values categorical switched from WHO 2008 definitions to IARC 2017 definitions, same type of information but different values categorical same type of information but different values categorical same type of information but different values categorical sets prain and CNS Recode Brain recode SEER Brain and CNS Recode SEER Brain and CNS Recode Mod Same type of information but different values categorical sets provided CS Schema v0204*/ TMM 7/CS v0204+ Schema (thru 2017) yes Race recode A Race recode (W, B, AI, API) yes Race recode Y Race recode (W, B, AI, API) Negro Congin Recode NHIA (Hispanic, Non-Hisp) yes SEER historic stage A SEER historic stage A (1973-2015) yes SEER summary Stage and titling (1988+) ALCC stage 3rd edition (1988-2003) yes SEER Summary Stage 1977 (1995-2000) ACC stage 3rd edition (1988-2003) SEER Summary Stage 1977 (1995-2000) ACC stage 1977 (1995-2000) ACC stage 1977 (1995-2000) ACC stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator ACC stage 2000 2000 (2001-2003) ACC stage 3rd edition (1988-2003) ACC stage 3rd edition (1988-2003) Yes SEER Summary Stage 1977 (1995-2000) ACC stage 3rd edition (1988-2003) ACC stage 3rd edition (1988-2003) Yes SEER Summary Stage 1977 (1995-2000) ACC stage 3rd edition (1988-2003) ACC stage 3rd editi	Recode ICD-O-2 to 10		drop		categorical	
Switched from WHO 2008 definitions to IARC 2017 definitions, categorical sehavior recode for analysis Behavior recode for analysis Benavior recode for analysis Benavior recode for analysis Benavior recode for analysis Broan Histology recode Histology recode - Histology recode - Benavior recode for analysis Broan recode SEER Brain and CNS Recode Mod Same type of information but different values categorical Categorical Categorical Categorical TNM 7/CS v02044 Schema (thru 2017) yes Categorical Race recode (White, Black, Other) Race recode A Race recode (White, Black, Other) Race recode Y Race recode (White, Black, Other) Origin Recode NHIA Origin recode NHIA (Hispanic, Non-Hisp) SEER historic stage A SEER historic stage A SEER historic stage A SEER historic stage A AUCC stage 3rd edition (1988+) SEER modified AUCC stage 3rd edition (1988-2003) SEER modified AUCC stage 3rd edition (1988-2003) SEER modified AUCC stage 3rd edition (1988-2003) SEER summary Stage 1977 (1995-2000) drop same as 2016's "Derived S51977". Not available in 2023 database "Combined Summary Stage field to facilitate stage analyses over time." categorical				1		
ICCC site rec extended ICD-0-3/ WHO 2008 ICCC site recode extended 3rd edition/IARC 2017 ICCC site recode for analysis Initiation but different values Icategorical ICCC site recode for snalysis ICCC site recode for snalysis Initiation but different values Icategorical ICCC site recode for snalysis ICCC site recode for sality in finition but different values Icategorical ICCC site recode for sality in finition but different values Icategorical ICCC site recode for sanalysis ICCC site recode for sanal transfer to sate prical ICCC stage of information but different values Icategorical ICCC stage recode for sanal transfer to sate prical ICCC stage recode for sanal transfer to sate prical	ICCC site recode ICD-O-3/WHO 2008	ICCC site recode 3rd edition/IARC 2017	mod	77	categorical	
Behavior recode for analysis Behavior recode for analysis Broad Histology recode Histology recode SEER Brain and CNS Recode CS Schema v0204*/ TNM 7/CS v0204* Schema (thru 2017) Race recode A Race recode (White, Black, Other) Race recode Y Race recode (White, Black, Other) Resource (Wh						
Broad Histology recode Histology recode - broad groupings Press	ICCC site rec extended ICD-O-3/ WHO 2008	ICCC site recode extended 3rd edition/IARC 2017	mod	same type of information but different values	categorical	
Brain recode SEER Brain and CNS Recode TMM 7/CS v0204+ Schema (thru 2017) Race recode A Race recode (White, Black, Other) Race recode Y Race recode (White, Black, Other) Race recode NHIA Origin recode NHIA (Hispanic, Non-Hisp) SEER historic stage A SEER historic stage A (1973-2015) SEER historic stage A (1973-2015) SEER modified AICC stage 3rd ed (1988+) SEER modified AICC stage 3rd ed (1988+) SEER Summary Stage 1977 (1995-2000) SEER Summary Stage 1977 (1995-2000) SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator SEER summary First malignant primary indicator SEER state-county recode Couse of death to SEER site recode COD to site recode COD to site recode TMM 7/CS v0204+ Schema (thru 2017) yes SEER by service and set of information but different values categorical categorical same type of information but different values categorical categorical categorical categorical categorical categorical categorical categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical categorical categorical categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical c	Behavior recode for analysis	Behavior recode for analysis	yes		categorical	
CS Schema v0204*/ Race recode (White, Black, Other) Race recode (Wh. Al, API) Origin Recode NHIA Origin recode NHIA (Hispanic, Non-Hisp) SEER historic stage A SEER historic stage A SEER historic stage A (1973-2015) SEER modified AICC stage 3rd edition (1988+) SEER modified AICC stage 3rd edition (1988-2003) SEER modified AICC stage 3rd edition (1988-2003) SEER summary Stage 1977 (1995-2000) SEER Summary Stage 1977 (1995-2000) SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator SEER summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator SEER summary Stage summary Stage (2004+) First malignant primary indicator SEER summary Stage summary Stage summary Stage (2004+) First malignant primary indicator SEER summary Stage summary Stage summary Stage (2004+) First malignant primary indicator SEER summary Stage summary Stage summary Stage (2004+) SEER summary Stage su	Broad Histology recode	Histology recode - broad groupings	yes		categorical	
Race recode A Race recode (White, Black, Other) Race recode Y Race recode (W, B, AI, API) Recode NHIA Origin Recode NHIA (Hispanic, Non-Hisp) Rese Ristoric stage A SEER historic stage A (1973-2015) SEER nistoric stage A (1973-2015) SEER modified AJCC stage 3rd edition (1988+) SEER modified AJCC stage 3rd edition (1988+) SEER modified AJCC stage 3rd edition (1988+) SEER modified AJCC stage 3rd (1988-2003) SEER Summary Stage 1977 (1995-2000) SEER Summary Stage 1977 (1995-2000) SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) Mod Tirst malignant primary indicator First malignant primary indicator First malignant primary indicator SEER stage and the first maligna	Brain recode	SEER Brain and CNS Recode	mod	same type of information but different values	categorical	
Race recode Y Race recode (W, B, AI, API) Prigin Recode NHIA Origin recode NHIA (Hispanic, Non-Hisp) Prigin Recode NHIA Origin recode NHIA Origin recode NHIA (Hispanic, Non-Hisp) Prigin Recode NHIA Origin rec	CS Schema v0204*/	TNM 7/CS v0204+ Schema (thru 2017)	yes		categorical	
Origin Recode NHIA Origin recode NHIA (Hispanic, Non-Hisp) Yes Categorical SEER historic stage A SEER historic stage A (1973-2015) SEER modified AICC stage 3rd edition (1988-2003) SEER modified AICC stage 3rd edition (1988-2003) SEER modified AICC stage 3rd (1988-2003) SEER summary Stage 1977 (1995-2000) SEER Summary Stage 1977 (1995-2000) SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator First malignant primary indicator SEER summary Stage S	Race recode A	Race recode (White, Black, Other)	yes		categorical	
Origin Recode NHIA Origin recode NHIA (Hispanic, Non-Hisp) Yes Categorical SEER historic stage A SEER historic stage A (1973-2015) SEER modified AICC stage 3rd edition (1988-2003) SEER modified AICC stage 3rd edition (1988-2003) Yes Categorical SEER modified AICC stage 3rd ed (1988+) SEER modified AICC stage 3rd (1988-2003) SEER modified AICC stage 3rd (1988-2003) SEER summary Stage 1977 (1995-2000) SEER Summary Stage 1977 (1995-2000) Combined Summary Stage (2004+) SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) Mod Time." Categorical "Combined Summary Stage field to facilitate stage analyses over time." Categorical categorical "Combined Summary Stage field to facilitate stage analyses over time." Categorical categorical categorical categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features Cause of death to SEER site recode COD to site recode COD to site recode COD to site recode COD to site recode SEER historic stage A Yes Categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical	Race recode Y	Race recode (W, B, AI, API)	yes			
SEER historic stage A SEER historic stage A (1973-2015) yes categorical AJCC stage 3rd edition (1988+) AJCC stage 3rd edition (1988-2003) yes categorical SEER modified AJCC stage 3rd ed (1988+) SEER modified AJCC stage 3rd (1988-2003) yes categorical SEER Summary Stage 1977 (1995-2000) drop same as 2016's "Derived SS1977". Not available in 2023 database categorical "Combined Summary Stage field to facilitate stage analyses over time." categorical stage. SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) mod time." categorical stage. First malignant primary indicator yes confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical cate	Origin Recode NHIA				_	
AJCC stage 3rd edition (1988+) AJCC stage 3rd edition (1988-2003) AJCC stage 3rd ed (1988-2003) AJCC	SEER historic stage A					
SEER modified AJCC stage 3rd ed (1988+) SEER modified AJCC stage 3rd (1988-2003) SEER summary Stage 1977 (1995-2000) Combined Summary Stage (2004+) First malignant primary indicator Combined Summary Stage field to facilitate stage analyses over time." categorical "Combined Summary Stage field to facilitate stage analyses over time." categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical categorical categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical categorical						
SEER Summary Stage 1977 (1995-2000) Combined Summary Stage (2004+) First malignant primary indicator State-county recode Cause of death to SEER site recode drop same as 2016's "Derived SS1977". Not available in 2023 database categorical categorical categorical categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features removed post-dx						
SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator First malignant primary indicator State-county recode Cause of death to SEER site recode "Combined Summary Stage field to facilitate stage analyses over time." categorical categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features removed post-dx	, , , , , , , , , , , , , , , , , , , ,		1,			
SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator First malignant primary indicator State-county recode Cause of death to SEER site recode "Combined Summary Stage field to facilitate stage analyses over time." categorical categorical confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features removed post-dx	SEER Summary Stage 1977 (1995-2000)		drop	same as 2016's "Derived SS1977" Not available in 2023 database	categorical	
SEER Summary Stage 2000 2000 (2001-2003) Combined Summary Stage (2004+) First malignant primary indicator yes confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical categorical categorical categorical removed post-dx	522 Sammary Stuge 1577 (1555 2000)		Патор		- Categoricai	
First malignant primary indicator First malignant primary indicator First malignant primary indicator First malignant primary indicator yes confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical with "County Attributes - Time Dependent" features removed post-dx	SEER Summary Stage 2000 2000 (2001-2003)	Combined Summary Stage (2004+)	mod	, , ,	categorical	
State-county recode confidentiality, data item not available at individual level, replaced with "County Attributes - Time Dependent" features categorical removed post-dx				unic.		
State-county recode drop with "County Attributes - Time Dependent" features categorical Post-dx Cause of death to SEER site recode COD to site recode yes removed post-dx	Thist manghant primary multator	That manghant primary mulcator	yes		categorical	
State-county recode drop with "County Attributes - Time Dependent" features categorical Post-dx Cause of death to SEER site recode COD to site recode yes removed post-dx				confidentiality data item not available at individual level	.	
Cause of death to SEER site recode COD to site recode yes removed post-dx	State assumb users de					
		COD to the constitution		with County Attributes - Time Dependent" reatures		
LOD to site rec KM COD to site rec KM yes removed post-dx					1	
	COD to site rec KM	COD to site rec KM	yes		removed	post-dx

Vital status recode (study cutoff used)	Vital status recode (study cutoff used)	ves		removed	post-dx
IHS link	IHS Link	yes		categorical	1
Historic SSG 2000 Stage	Summary stage 2000 (1998-2017)	ves		categorical	
-	, , , ,	<u> </u>			
AYA site recode/WHO 2008	AYA site recode 2020 Revision	mod	updated revision, same type of information but different values	categorical	
Lymphoma subtype recode/WHO 2008	Lymphoid neoplasm recode 2021 Revision	mod	updated revision, similar except doesn't incorporate grade	categorical	
SEER cause of death classification	SEER cause-specific death classification	yes		target	
SEER other cause of death classification	SEER other cause of death classification	yes		removed	post-dx
CS EXT/Size Eval	CS Tumor Size/Ext Eval (2004-2015)	yes		removed	treatment
CS Nodes Eval	CS Reg Node Eval (2004-2015)	yes		removed	treatment
CS Mets Eval	CS Mets Eval (2004-2015)	yes		removed	treatment
Primary by International Rules	Primary by international rules	yes		categorical	
ER Status Recode Breast Cancer (1990+)*/	ER Status Recode Breast Cancer (1990+)	yes		categorical	
PR Status Recode Breast Cancer (1990+)*/	PR Status Recode Breast Cancer (1990+)	ves		categorical	
CS Schema - AJCC 6th Edition	CS Schema - AJCC 6th Edition	yes		categorical	
Cs Site-specific Factor 8	CS site-specific factor 8 (2004-2017 varying by schema)	ves		categorical	
CS Site-Specific Factor 10*/	CS site-specific factor 10 (2004-2017 varying by schema)	ves		categorical	
CS Site-Specific Factor 11*/	CS site-specific factor 11 (2004-2017 varying by schema)	yes		categorical	
CS Site-Specific Factor 13*/	CS site-specific factor 13 (2004-2017 varying by schema)	ves		categorical	
CS Site-Specific Factor 15*/	CS site-specific factor 15 (2004-2017 varying by schema)	ves		categorical	
CS Site-Specific Factor 16*/	CS site-specific factor 16 (2004-2017 varying by schema)	yes		categorical	
Lymph-vascular Invasion (2004+)*/	Lymph-vascular Invasion (2004+ varying by schema)	ves		categorical	
Survival months	Survival months	ves		target	
Survival months flag	Survival months flag	ves		removed	irrelevancy
			Caution is urged when using this variable due to Medicare		,
Insurance Recode (2007+)		drop	eligibility. Not available in 2023 database.	categorical	
Derived AJCC T 7th ed	Derived AJCC T, 7th ed (2010-2015)	ves		categorical	
Derived AJCC N 7th ed	Derived AJCC N, 7th ed (2010-2015)	yes		categorical	
Derived AJCC M 7th ed	Derived AJCC M, 7th ed (2010-2015)	yes		categorical	
Derived AJCC 7 Stage Group	Derived AJCC Stage Group, 7th ed (2010-2015)	ves		categorical	
Adjusted AJCC 6th T (1988+)	Breast - Adjusted AJCC 6th T (1988-2015)	ves	only differ 126/Blank added for data after 2015	categorical	
Adjusted AJCC 6th N (1988+)	Breast - Adjusted AJCC 6th N (1988-2015)	ves	only differ 126/Blank added for data after 2015	categorical	
Adjusted AJCC 6th M (1988+)	Breast - Adjusted AJCC 6th M (1988-2015)	ves	only differ 126/Blank added for data after 2015	categorical	
Adjusted AJCC 6th Stage (1988+)	Breast - Adjusted AJCC 6th Stage (1988-2015)	ves	only differ 126/Blank added for data after 2015	categorical	
CS Site-Specific Factor 7	Adjusted CS site-specific factor 7 (2004-2017 varying by schema)	yes	no immediately apparent differences	categorical	
CS Site-specific Factor 9	CS site-specific factor 9 (2004-2017 varying by schema)	ves	, , , , , , , , , , , , , , , , , , ,	categorical	
CS Site-Specific Factor 12	CS site-specific factor 12 (2004-2017 varying by schema)	ves		categorical	
Derived HER2 Recode (2010+)	Derived HER2 Recode (2010+)	yes		categorical	
Breast Subtype (2010+)	Breast Subtype (2010+)	ves		categorical	
Lymphoma - Ann Arbor Stage (1983+)	Lymphoma - Ann Arbor Stage (1983-2015)	ves		categorical	
CS mets at DX-bone (2010+)	SEER Combined Mets at DX-bone (2010+)	yes	only differ 14/Blank added	categorical	
CS mets at DX-brain (2010+)	SEER Combined Mets at DX-brain (2010+)	ves	only differ 14/Blank added	categorical	
CS mets at DX-liver (2010+)	SEER Combined Mets at DX-liver (2010+)	ves	only differ 14/Blank added	categorical	
CS mets at DX-lung (2010+)	SEER Combined Mets at DX-lung (2010+)	ves	only differ 14/Blank added	categorical	
T value - based on AJCC 3rd (1988-2003)	T value - based on AJCC 3rd (1988-2003)	yes	, , ,	categorical	
N value - based on AJCC 3rd (1988-2003)	N value - based on AJCC 3rd (1988-2003)	yes		categorical	
M value - based on AJCC 3rd (1988-2003)	M value - based on AJCC 3rd (1988-2003)	ves		categorical	
Total number of in situ/malignant tumors for patient	Total number of in situ/malignant tumors for patient	ves		removed	post-dx
Total number of benign/borderline tumors for patient	Total number of benign/borderline tumors for patient	ves		removed	post-dx
5 - 1 - 2 - 5 , 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1 - 1		1,		1	
	Median household income inflation adj to 2022	add	added for relevant information to replace 'State-county recode'	categorical	
			The state of the s	32128311001	

 $^{^1} https://seer.cancer.gov/data-software/documentation/seerstat/nov2016/TextData. FileDescription.pdf$

 $^{^2} https://seer.cancer.gov/data-software/documentation/seerstat/nov2023/TextData. File Description-nov2023.pdf$