

# Cancer

## PROGRAMS

---

AMERICAN COLLEGE OF SURGEONS

● ● ● NATIONAL  
● ● ● CANCER  
● ● ● DATABASE

## **Quick Start Guide to PUF Data Structure**

### **2017 National Cancer Database Participant Use File (NCDB PUF)**

The data included in the zipped file provided in the PUF download, are provided in a flat text file format, and should be read with software such as SAS, SPSS (PASW), STATA, or any other statistical software of your choosing. Also included in this zipped file are two scripts that will allow you to read the data in SAS (NCDB\_PUF\_Labels\_2017.sas), or SPSS (NCDB\_PUF\_Labels\_2017.sps). You must have SAS, or SPSS/PASW, software installed on your computer to be able to use these files. These two scripts include 1) input statements that will read the provided flat text input file and assign variable names, 2) variable labels and 3) either format statements (SAS) or value labels (SPSS/PASW). If you are using other software, such as STATA, you can use the file layout below to read in the PUF data.

The table appearing on the following pages provides a list of each of the item made available in PUF. In addition, “standard” or “analytic” variable names are provided, the diagnosis years for which individual items are available, and finally the item length and column location of each item is specified. The item and variable names in this table are the same as those documented in the online 2017 PUF Data Dictionary. The data dictionary includes the item names, descriptions and value definitions of the variables included in your PUF file. Please refer to this data dictionary before beginning your analysis to ensure your familiarity with the provided data. Note that the items included, and diagnosis year availability are only applicable to the 2017 PUF, which includes cases diagnosed in 2004-2017.

Please note per the Data Use Agreement that any abstracts or publications using these data should include the following acknowledgement in all presentations/publications:

“The data used in the study are derived from a de-identified NCDB file. The American College of Surgeons and the Commission on Cancer have not verified and are not responsible for the analytic or statistical methodology employed, or the conclusions drawn from these data by the investigator.”

The abstracts and publications also need to indicate:

“The NCDB is a joint project of the Commission on Cancer of the American College of Surgeons and the American Cancer Society.” If you have any questions regarding these files, e-mail your questions to [NCDB\\_PUF@facs.org](mailto:NCDB_PUF@facs.org).

NCDB 2017 PUF Item List		Adult Only/All*	PUF Item Name	Diagnosis Years Available**	File Layout			Data Type
					Characters	Start	End	
Case Key	Case Key	All	PUF_CASE_ID	All	37	1	37	String
Facility	Facility Key	All	PUF_FACILITY_ID	All	10	38	47	String
	Facility Type	Adult*	FACILITY_TYPE_CD	All	1	48	48	Numeric
	Facility Location	Adult*	FACILITY_LOCATION_CD	All	1	49	49	Numeric
Patient Demographics	Patient Age at Diagnosis	All	AGE	All	3	50	52	Numeric
	Sex	All	SEX	All	1	53	53	Numeric
	Race	All	RACE	All	2	54	55	Numeric
	Spanish Hispanic Origin	All	SPANISH_HISPANIC_ORIGIN	All	1	56	56	Numeric
	Primary Payor at Diagnosis	All	INSURANCE_STATUS	All	2	57	58	Numeric
	Income 2000	All	MED_INC_QUAR_00	All	1	59	59	Numeric
	Education 2000	All	NO_HSD_QUAR_00	All	1	60	60	Numeric
	Urban/Rural 2003	All	UR_CD_03	All	1	61	61	Numeric
	Income 2008-2012	All	MED_INC_QUAR_12	All	1	62	62	Numeric
	Education 2008-2012	All	NO_HSD_QUAR_12	All	1	63	63	Numeric
	Urban/Rural 2013	All	UR_CD_13	All	1	64	64	Numeric
	Great Circle Distance	All	CROWFLY	All	8	65	72	Numeric
	Charlson-Deyo Score	All	CDCC_TOTAL_BEST	All	2	73	74	Numeric
Cancer Identification	Sequence Number	All	SEQUENCE_NUMBER	All	2	75	76	String
	Class of Case	All	CLASS_OF_CASE	All	2	77	78	Numeric
	Year of Diagnosis	All	YEAR_OF_DIAGNOSIS	All	4	79	82	Numeric
	Primary Site	All	PRIMARY_SITE	All	4	83	86	String
	Laterality	All	LATERALITY	All	1	87	87	Numeric
	Histology	All	HISTOLOGY	All	4	88	91	Numeric
	Behavior	All	BEHAVIOR	All	1	92	92	Numeric
	Grade	All	GRADE	All	1	93	93	Numeric
	Diagnostic Confirmation	All	DIAGNOSTIC_CONFIRMATION	All	1	94	94	Numeric
	Size of Tumor	All	TUMOR_SIZE	All	3	95	97	Numeric
	Regional Lymph Nodes Positive	All	REGIONAL_NODES_POSITIVE	All	2	98	99	Numeric
	Regional Lymph Nodes Examined	All	REGIONAL_NODES_EXAMINED	All	2	100	101	Numeric
Stage of Disease	Diagnostic and Staging Procedure, Days from Dx	All	DX_STAGING_PROC_DAYS	All	8	102	109	Numeric
	Diagnostic and Staging Procedure	All	RX_SUMM_DXSTG_PROC	All	2	110	111	Numeric
	AJCC Clinical T	All	TNM_CLIN_T	All	5	112	116	String
	AJCC Clinical N	All	TNM_CLIN_N	All	5	117	121	String
	AJCC Clinical M	All	TNM_CLIN_M	All	5	122	126	String
	AJCC Clinical Stage Group	All	TNM_CLIN_STAGE_GROUP	All	4	127	130	String
	AJCC Pathologic T	All	TNM_PATH_T	All	5	131	135	String
	AJCC Pathologic N	All	TNM_PATH_N	All	5	136	140	String
	AJCC Pathologic M	All	TNM_PATH_M	All	5	141	145	String

	AJCC Pathologic Stage Group	All	TNM_PATH_STAGE_GROUP	All	4	146	149	String
	TNM Edition Number	All	TNM_EDITION_NUMBER	All	2	150	151	Numeric
	NCDB Analytic Stage Group	All	ANALYTIC_STAGE_GROUP	All	1	152	152	Numeric
	CS Mets at DX	All	CS_METS_AT_DX	2004-2015	2	153	154	String
	CS Mets Eval	All	CS_METS_EVAL	2004-2015	1	155	155	String
	CS Extension	All	CS_EXTENSION	2004-2015	3	156	158	String
	CS Tumor Size/Ext Eval	All	CS_TUMOR_SIZEEXT_EVAL	2004-2015	1	159	159	String
	CS Mets at Dx-Bone	All	CS_METS_DX_BONE	2010-2015	1	160	160	Numeric
	CS Mets at Dx-Brain	All	CS_METS_DX_BRAIN	2010-2015	1	161	161	Numeric
	CS Mets at Dx-Liver	All	CS_METS_DX_LIVER	2010-2015	1	162	162	Numeric
	CS Mets at Dx-Lung	All	CS_METS_DX_LUNG	2010-2015	1	163	163	Numeric
	Lymph Vascular Invasion	All	LYMPH_VASCULAR_INVASION	2010+	1	164	164	Numeric
	CS Site Specific Factor 1	All	CS_SITESPECIFIC_FACTOR_1	***	3	165	167	Numeric
	CS Site Specific Factor 2	All	CS_SITESPECIFIC_FACTOR_2	***	3	168	170	Numeric
	CS Site Specific Factor 3	All	CS_SITESPECIFIC_FACTOR_3	***	3	171	173	Numeric
	CS Site Specific Factor 4	All	CS_SITESPECIFIC_FACTOR_4	***	3	174	176	Numeric
	CS Site Specific Factor 5	All	CS_SITESPECIFIC_FACTOR_5	***	3	177	179	Numeric
	CS Site Specific Factor 6	All	CS_SITESPECIFIC_FACTOR_6	***	3	180	182	Numeric
	CS Site Specific Factor 7	All	CS_SITESPECIFIC_FACTOR_7	***	3	183	185	Numeric
	CS Site Specific Factor 8	All	CS_SITESPECIFIC_FACTOR_8	***	3	186	188	Numeric
	CS Site Specific Factor 9	All	CS_SITESPECIFIC_FACTOR_9	***	3	189	191	Numeric
	CS Site Specific Factor 10	All	CS_SITESPECIFIC_FACTOR_10	***	3	192	194	Numeric
	CS Site Specific Factor 11	All	CS_SITESPECIFIC_FACTOR_11	***	3	195	197	Numeric
	CS Site Specific Factor 12	All	CS_SITESPECIFIC_FACTOR_12	***	3	198	200	Numeric
	CS Site Specific Factor 13	All	CS_SITESPECIFIC_FACTOR_13	***	3	201	203	Numeric
	CS Site Specific Factor 14	All	CS_SITESPECIFIC_FACTOR_14	***	3	204	206	Numeric
	CS Site Specific Factor 15	All	CS_SITESPECIFIC_FACTOR_15	***	3	207	209	Numeric
	CS Site Specific Factor 16	All	CS_SITESPECIFIC_FACTOR_16	***	3	210	212	Numeric
	CS Site Specific Factor 17	All	CS_SITESPECIFIC_FACTOR_17	***	3	213	215	Numeric
	CS Site Specific Factor 18	All	CS_SITESPECIFIC_FACTOR_18	***	3	216	218	Numeric
	CS Site Specific Factor 19	All	CS_SITESPECIFIC_FACTOR_19	***	3	219	221	Numeric
	CS Site Specific Factor 20	All	CS_SITESPECIFIC_FACTOR_20	***	3	222	224	Numeric
	CS Site Specific Factor 21	All	CS_SITESPECIFIC_FACTOR_21	***	3	225	227	Numeric
	CS Site Specific Factor 22	All	CS_SITESPECIFIC_FACTOR_22	***	3	228	230	Numeric
	CS Site Specific Factor 23	All	CS_SITESPECIFIC_FACTOR_23	***	3	231	233	Numeric
	CS Site Specific Factor 24	All	CS_SITESPECIFIC_FACTOR_24	***	3	234	236	Numeric
	CS Site Specific Factor 25	All	CS_SITESPECIFIC_FACTOR_25	***	3	237	239	Numeric
	CS Version Number	All	CS_VERSION_LATEST	All	6	240	245	Numeric
Treatment	Treatment Started, Days from Dx	All	DX_RX_STARTED_DAYS	All	8	246	253	Numeric
	First Surgical Procedure, Days from Dx	All	DX_SURG_STARTED_DAYS	All	8	254	261	Numeric

	Definitive Surgical Procedure, Days from Dx	All	DX_DEFSURG_STARTED_DAYS	All	8	262	269	Numeric
	Surgical Procedure of the Primary Site	All	RX_SUMM_SURG_PRIM_SITE	All	2	270	271	Numeric
	Surgical Approach	All	RX_HOSP_SURG_APPR_2010	2010+	1	272	272	Numeric
	Surgical Margins	All	RX_SUMM_SURGICAL_MARGINS	All	1	273	273	Numeric
	Scope of Regional LN Surgery	All	RX_SUMM_SCOPE_REG_LN_SUR	All	1	274	274	Numeric
	Surgery Other Site	All	RX_SUMM_SURG_OTH_REGDIS	All	1	275	275	Numeric
	Surgical Inpatient Stay, Days from Surgery	All	SURG_DISCHARGE_DAYS	All	8	276	283	Numeric
	Readmission Within 30 Days of Surgical Discharge	All	READM_HOSP_30_DAYS	All	1	284	284	Numeric
	Reason for No Surgery	All	REASON_FOR_NO_SURGERY	All	1	285	285	Numeric
	Radiation, Days from Dx	All	DX_RAD_STARTED_DAYS	All	8	286	293	Numeric
	Location of Radiation Therapy	All	RAD_LOCATION_OF_RX	All	1	294	294	Numeric
	Radiation Surgery Sequence	All	RX_SUMM_SURGRAD_SEQ	All	1	295	295	Numeric
	Radiation Ended, Days from Start of Radiation	All	RAD_ELAPSED_RX_DAYS	All	3	296	298	Numeric
	Reason for No Radiation	All	REASON_FOR_NO_RADIATION	All	1	299	299	Numeric
	Systemic, Days from Dx	All	DX_SYSTEMIC_STARTED_DAYS	All	8	300	307	Numeric
	Chemotherapy, Days from Dx	All	DX_CHEMO_STARTED_DAYS	All	8	308	315	Numeric
	Chemotherapy	All	RX_SUMM_CHEMO	All	2	316	317	Numeric
	Hormone Therapy, Days from Dx	All	DX_HORMONE_STARTED_DAYS	All	8	318	325	Numeric
	Hormone Therapy	All	RX_SUMM_HORMONE	All	2	326	327	Numeric
	Immunotherapy, Days from Dx	All	DX_IMMUNO_STARTED_DAYS	All	8	328	335	Numeric
	Immunotherapy	All	RX_SUMM_IMMUNOTHERAPY	All	2	336	337	Numeric
	Hematologic Transplant and Endocrine Procedures	All	RX_SUMM_TRNSPLNT_ENDO	All	2	338	339	Numeric
	Systemic Surgery Sequence	All	RX_SUMM_SYSTEMIC_SUR_SEQ	2006+	1	340	340	Numeric
	Other Treatment, Days from Dx	All	DX_OTHER_STARTED_DAYS	All	8	341	348	Numeric
	Other Treatment	All	RX_SUMM_OTHER	All	1	349	349	Numeric
	Palliative Care	All	PALLIATIVE_CARE	All	1	350	350	Numeric
	RX Summ--Treatment Status	All	RX_SUMM_TREATMENT_STATUS	All	1	351	351	Numeric
Outcomes and Additional Treatment, Case, and Facility Items	Thirty Day Mortality	All	PUF_30_DAY_MORT_CD	2004-2016	1	352	352	Numeric
	Ninety Day Mortality	All	PUF_90_DAY_MORT_CD	2004-2016	1	353	353	Numeric
	Last Contact or Death, Months from Dx	All	DX_LASTCONTACT_DEATH_MONTHS	2004-2016	8	354	361	Numeric
	PUF Vital Status	All	PUF_VITAL_STATUS	2004-2016	1	362	362	Numeric
	Surgery at This Facility	All	RX_HOSP_SURG_PRIM_SITE	All	2	363	364	Numeric
	Chemotherapy at This Facility	All	RX_HOSP_CHEMO	All	2	365	366	Numeric
	Immunotherapy at This Facility	All	RX_HOSP_IMMUNOTHERAPY	All	2	367	368	Numeric
	Hormone Therapy at This Facility	All	RX_HOSP_HORMONE	All	2	369	370	Numeric
	Other Treatment at This Facility	All	RX_HOSP_OTHER	All	2	371	372	Numeric
	PUF Multiple Source	All	PUF_MULT_SOURCE	All	1	373	374	Numeric
	Reference Date Flag	All	PUF_REFERENCE_DATE_FLAG	All	1	374	374	Numeric
	Scope of Regional LN Surgery, Dx Years 2012+	All	RX_SUMM_SCOPE_REG_LN_2012	2012+	1	375	375	Numeric

Diagnostic and Staging Procedure at This Facility	All	RX_HOSP_DXSTG_PROC	All	2	376	377	Numeric
Palliative Care at This Facility	All	PALLIATIVE_CARE_HOSP	All	1	378	378	Numeric
Tumor Size Summary 2016	All	TUMOR_SIZE_SUMMARY_2016	2016+	3	379	381	Numeric
Mets at Diagnosis Other	All	METS_AT_DX_OTHER	2016+	1	382	382	Numeric
Mets at Diagnosis Distant Lymph Nodes	All	METS_AT_DX_DISTANT_LN	2016+	1	383	383	Numeric
Mets at Diagnosis Bone	All	METS_AT_DX_BONE	2016+	1	384	384	Numeric
Mets at Diagnosis Brain	All	METS_AT_DX_BRAIN	2016+	1	385	385	Numeric
Mets at Diagnosis Liver	All	METS_AT_DX_LIVER	2016+	1	386	386	Numeric
Mets at Diagnosis Lung	All	METS_AT_DX_LUNG	2016+	1	387	387	Numeric
No High School Degree Quartiles 2016	All	NO_HSD_QUAR_2016	All	1	388	388	Numeric
Median Income Quartiles 2016	All	MED_INC_QUAR_2016	All	1	389	389	Numeric
*PUF Medicaid Expansion Code	Adult*	PUF_MEDICAID_EXPN_CODE	All	1	390	390	Numeric
Phase I Radiation Primary Treatment Volume	All	PHASE_I_RT_VOLUME	All	2	391	392	Numeric
Phase I Radiation to Draining Lymph Nodes	All	PHASE_I_RT_TO_LN	All	2	393	394	Numeric
Phase I Dose Per Fraction	All	PHASE_I_DOSE_FRACT	All	5	395	399	Numeric
Phase I Number of Fractions	All	PHASE_I_NUM_FRACT	All	3	400	402	Numeric
Phase I External Beam Planning Technique	All	PHASE_I_BEAM_TECH	All	2	403	404	Numeric
Phase I Total Dose	All	PHASE_I_TOTAL_DOSE	All	6	405	410	Numeric
Phase I Radiation Treatment Modality	All	PHASE_I_RT_MODALITY	All	2	411	412	Numeric
Phase II Radiation Primary Treatment Volume	All	PHASE_II_RT_VOLUME	All	2	413	414	Numeric
Phase II Radiation to Draining Lymph Nodes	All	PHASE_II_RT_TO_LN	All	2	415	416	Numeric
Phase II Dose per Fraction	All	PHASE_II_DOSE_FRACT	All	5	417	421	Numeric
Phase II Number of Fractions	All	PHASE_II_NUM_FRACT	All	3	422	424	Numeric
Phase II External Beam Planning Technique	All	PHASE_II_BEAM_TECH	All	2	425	426	Numeric
Phase II Total Dose	All	PHASE_II_TOTAL_DOSE	All	6	427	432	Numeric
Phase II Radiation Treatment Modality	All	PHASE_II_RT_MODALITY	All	2	433	434	Numeric
Phase III Radiation Primary Treatment Volume	All	PHASE_III_RT_VOLUME	2017+	2	435	436	Numeric
Phase III Radiation to Draining Lymph Nodes	All	PHASE_III_RT_TO_LN	2017+	2	437	438	Numeric
Phase III Dose Per Fraction	All	PHASE_III_DOSE_FRACT	2017+	5	439	443	Numeric
Phase III Number of Fractions	All	PHASE_III_NUM_FRACT	2017+	3	444	446	Numeric
Phase III External Beam Planning Technique	All	PHASE_III_BEAM_TECH	2017+	2	447	448	Numeric
Phase III Total Dose	All	PHASE_III_TOTAL_DOSE	2017+	6	449	454	Numeric
Phase III Radiation Treatment Modality	All	PHASE_III_RT_MODALITY	2017+	2	455	456	Numeric
Number of Phases of RT Treatment	All	NUMBER_PHASES_RAD_RX	2017+	2	457	458	Numeric
Radiation Discontinued Early	All	RAD_RX_DISC_EARLY	2017+	2	459	460	Numeric
Total Dose	All	TOTAL_DOSE	2017+	6	461	466	Numeric

\*Facility type, Facility Location, and Medicaid Expansion Codes are suppressed for ages 0-39.

\*\*Based on 2017 PUF including cases diagnosed in 2004-2017 only.

\*\*\*Availability differs by site and factor. Please visit the Collaborative Stage website for more information: <https://cancerstaging.org/cstage/schema/Pages/version0205.aspx>