Standard Column Reference Table

Complete Reference: Standard Column Names and Expected Values

STANDARD_COLUMN_NAME	Data Type	Expected Values / Format	Description	Example
COORDINATE_ID	Character	Unique sample identifier	Unique ID for each RNA-seq sample	TCGA-D3-A1Q5-01A
SAMPLE_ID	Character	Any unique sample ID	Sample identifier	Sample_001
PATIENT_ID	Character	Any unique patient ID	Patient identifier	Patient1, Pt001
FILE_NAME	Character	Filename string	Original filename or data source	data_file_001.bam
DATASET	Character	Dataset name	Source dataset (auto-generated)	TCGA_SKCM, Riaz, Liu
AGE_YEARS	Numeric	0-120	Patient age in years	65, 58.5
AGE_CATEGORY	Character	"0-18 years" "18-30 years" "31-40 years" "41-50 years" "51-60 years" "61-70 years" "71-80 years" "81-100 years"	Age group categories	"51-60 years"
GENDER	Character	"male" "female"	Patient sex/gender	male
RACE	Character	"White" "Black or African American" "Asian" Other descriptive values	Patient race	White
ETHNICITY	Character	"Hispanic or Latino" "Not Hispanic or Latino"	Patient ethnicity	Not Hispanic or Latino

STANDARD_COLUMN_NAME	Data Type	Expected Values / Format	Description	Example
VITAL_STATUS	Character	"Alive" "Dead"	Patient vital status	Alive
TIME_TO_DEATH_DAYS	Numeric	Positive number	Days from diagnosis/treatment to death	245.5
TIME_TO_LAST_FOLLOWUP_DAYS	Numeric	Positive number	Days from diagnosis/treatment to last follow-up	730.0
OS_MONTHS	Numeric	Positive number	Overall survival in months	24.5
OS_STATUS	Character	"0" (Alive) "1" (Dead)	Overall survival status (binary)	1
PFS	Numeric	Positive number	Progression-free survival in months	12.3
PROGRESSION	Character	"Yes" "No"	Disease progression status	Yes
RECURRENCE	Character	"Yes" "No"	Disease recurrence status	No
WHO_GRADE	Character	"WHO 1" "WHO 2" "WHO 3" "WHO 4"	WHO tumor grade	WHO 3
AJCC_STAGE	Character	"Stage 0" "Stage I" "Stage II" "Stage III" "Stage IV"	AJCC cancer stage	Stage III
T_STAGE	Character	"T0", "T1", "T1a", "T1b", "T1c" "T2", "T2a", "T2b", "T2c" "T3", "T3a", "T3b", "T3c"	Primary tumor stage	T2a

STANDARD_COLUMN_NAME	Data Type	Expected Values / Format	Description	Example
		"T4", "T4a", "T4b", "T4c"		
N_STAGE	Character	"N0" "N1", "N1a", "N1b", "N1c" "N2", "N2a", "N2b", "N2c" "N3", "N3a", "N3b", "N3c"	Regional lymph node stage	N1
M_STAGE	Character	"M0" "M1" "M1A" "M1B" "M1C" "IIIC"	Distant metastasis stage	M1A
PRIMARY_MET	Character	"Primary" "Metastatic"	Primary tumor vs metastatic	Metastatic
DIAGNOSIS	Character	"acral" "mucosal" "cutaneous" "nodular" "occult" "ocular/uveal" "superficial spreading" "other"	Melanoma subtype/diagnosis	cutaneous
TREATMENT_IMMUNOTHERAPY	Character	"Nivolumab" "Pembrolizumab" "Ipilimumab" "Ipilimumab+Nivolumab" "Ipilimumab+Pembrolizumab" Other combinations	Immunotherapy treatment	Pembrolizumab
TREATMENT_CHEMOTHERAPY	Character	Drug names or "Yes"/"No"	Chemotherapy treatment	Dacarbazine
TREATMENT_TARGETED	Character	Drug names or "Yes"/"No"	Targeted therapy treatment	Vemurafenib
TREATMENT_RADIATION	Character	"Yes" "No" Descriptive text	Radiation therapy	No

STANDARD_COLUMN_NAME	Data Type	Expected Values / Format	Description	Example
PRIOR_TREATMENT	Character	"Yes" "No"	Any prior cancer treatment	Yes
PRIOR_MALIGNANCY	Character	"Yes" "No"	Prior cancer diagnosis	No
SAMPLE_TIMEPOINT	Character	"Pre-treatment" "On-treatment" "Post-treatment"	When sample was collected	Pre-treatment
SITE_OF_RESECTION	Character	"Lymph node" "Skin/Soft tissue" "Brain" "Lung" "Liver" "Bone" "Abdomen/GI" "Thorax" "Primary tumor" "Other organ" "Pelvis" "Mass/Nodule"	Anatomical site of biopsy	Lymph node
RECIST_RESPONSE	Character	"CR" (Complete Response) "PR" (Partial Response) "PRCR" (PR+CR) "SD" (Stable Disease) "PD" (Progressive Disease) "NE" (Not Evaluable) "MR" (Minor Response)	RECIST tumor response	CR
RESPONDER_NONRESPONDER	Character	"Responder" "Non-responder"	Binary response classification	