**Suppl 2.** Variables in analysis. Categorization reflects final variable composition.

|  |  |  |
| --- | --- | --- |
| **Variable (SEER Variable)** | **Variable Name/Description per SEER** | **Categorization** |
| **Exposure** | | |
| Histology - ICD-0-3, malignant | 814x/3 (Adenocarcinoma)  807x/3 (Squamous Cell Carcinoma)  856x/3 (Adenosquamous Carcinoma) (Nasopharyngeal) | Adenocarcinoma – 814x  Squamous Cell Carcinoma – 807x  Adenosquamous Carcinoma – 856x |
| **Outcome** | | |
| Mortality      Survival Months | SEER cause-specific death classification  0 (Alive or dead of other cause), 1 (Dead), 8 (Dead – missing/unknown cause of death), 9 (N/A not first tumor)    Survival months  0-9998 (in months), 9999 (Unknown) | Alive/N/A – 0,9  Dead – 1      Months (0-9998) |
| **Co-variables** | | |
| Age  (Age recode single ages and 90+) | 00-89 (Actual age of diagnosis in years),  90+ (90), 999 (Unknown) | 0-89 (age in years) |
| Sex | 1 (Male), 2 (Female) | Male – 1  Female – 2 |
| Race  (race recode (white black other)) | 1 (white), 2 (black) 3 (other), 9 (unknown) | White – 1  Black – 2  Other – 3 |
| Detection Stage  (SEER historic stage A (1975-2015) &  Combined Stage Grade 2004+ (2016-2020)) | 0 (In situ —A noninvasive neoplasm; a tumor which has not penetrated the basement membrane nor extended beyond the epithelial tissue)  1 (Localized —An invasive neoplasm confined entirely to the organ of origin. It may include intraluminal extension where specified)  2 (Regional —A neoplasm that has extended 1) beyond the limits of the organ of origin directly into surrounding organs or tissues; 2) into regional lymph nodes by way of the lymphatic system; or 3) by a combination of extension and regional lymph nodes)  4/7 (Distant —A neoplasm that has spread to parts of the body remote from the primary tumor either by direct extension or by discontinuous metastasis to distant organs, issues, or via the lymphatic system to distant lymph nodes)  9 (Unstaged —Information is not sufficient to assign a stage) | In situ – 0  Localized – 1  Regional – 2  Distant – 4  Unknown – 9 |
| Grade  Differentiation  (Grade recode thru 2017 & Grade pathological 2018+) | 1,A (Grade I; grade i; grade 1; well differentiated; differentiated, NOS)  2,B (Grade II; grade ii; grade 2; moderately differentiated; moderately differentiated; intermediate differentiation)  3,C (Grade III; grade iii; grade 3; poorly differentiated; differentiated),  4,D (Grade IV; grade iv; grade 4; undifferentiated; anaplastic),  5 (T-cell; T-precursor), 6 (B-cell; Pre-B; B-Precursor), 7 (Null cell; Non T-non B), 8 (NK cell (natural killer cell))  9 (cell type not determined, not stated or not applicable) | Well – 1,A  Moderate – 2.B  Poor – 3,C,4,D  Unknown – 9 |
| Surgery  (Reason no cancer-directed surgery) | 0=Surgery performed  1=Not recommended  2=Not recommended, contraindicated due to other cond; autopsy only (1973-2002)  5=Not performed, patient died prior to recommended surgery  6=Recommended but not performed, unknown reason  7=Recommended but not performed, patient refused  8=Recommended, unknown if performed  9=Unknown; death certificate; or autopsy only (2003+) | Yes – 0  No – 1-7 |
| Chemotherapy | 0=None/Unknown  1=Yes | Yes – 1  No – 0 |
| Radiotherapy | 0=None/Unknown  1=Beam radiation  2=Radioactive implants (includes brachytherapy) (1988+)  3=Radioisotopes (1988+)  4=Combination of beam with implants or isotopes  5=Radiation, NOS method or source not specified  7=Refused (1988+)  8=Recommended, unknown if administered | Yes – 1,2,3,4,5  No – 0,7 |