**Code table for Cancer Staple Dataset**

**Data available from 1968 onwards to latest available year.**

| **SNO** | **Data Label** | **Description** | **Code Table** |
| --- | --- | --- | --- |
| 1 | UIN | Patient unique identifier (anonymised) |  |
| 2 | GENDER | Gender | 1=male, 2=female |
| 3 | ETHNIC\_GROUP | Ethnic | 1=chinese, 2=malay, 3=indian, 4=others |
| 4 | YEAR\_OF\_DIAGNOSIS | Year of diagnosis |  |
| 5 | AGE\_AT\_DIAGNOSIS | Age at diagnosis = diagnosis date - dob (In years) |  |
| 6 | PRI\_SITE | Primary site | ICD-9 [1968 - 1992]  ICD-O-2 [1993 - 2002]  IDC-O-3 [2003 onward] |
| 7 | LATERALITY | Laterality (from 2003) | 1=right side |
|  | 2=left side |
|  | 3=one side only involved, but right or left not specified |
|  | 4=both sides involved |
|  | blank=missing / NA |
| 8 | HISTO\_DIAGNOSIS | Histological diagnosis | refer to ICD-O-2 manual (2000-2002), ICD-O-3 manual (2003-2012)  /0=benign, /1=borderline, /2=in-situ, /3=invasive, 9999/9=unknown histology |
|  |  |
| 9 | GRADE\_DIFF | Grade differentiation (from 2003) | 1=well, 2=moderate, 3=poor, 4=undifferentiated, 5=T-cell, 6=B-cell, 7=Null cell, 8=NK cell, 9=NOS blank: missing - before 2003 |
| 10 | CLINICAL\_STAGE\_T | T category (AJCC staging system) (from 2003) | refer to AJCC manual (6th edn 2003-2009, 7th edn 2010-2017, 8th edn 2018 onwards) |
| 11 | CLINICAL\_STAGE\_N | N category (AJCC staging system) (from 2003) |
| 12 | CLINICAL\_STAGE\_M | M category (AJCC staging system) (from 2003) |
| 13 | STAGE\_GROUPING | Stage grouping (AJCC staging system) (from 2003) |
| 14 | PRI\_TUMOUR | Primary tumour (before 2003) | 0=carcinoma-in-situ |
|  | 1=cancer restricted to primary organ or tissue of origin |
|  | 2=local extension beyond primary organ or tissue of origin |
|  | 8=carcinoma-in-situ which subsequently becomes invasive  9=unknown or not applicable |
|  | blank=missing - i) after 2003, ii) missing |
| 15 | LYMPH\_NODE | Lymph node involvement (before 2003) | 1=yes, 2=no, 9=unknown or not applicable  blank=missing - i) after 2003, ii) missing |
| 16 | REMOTE\_META | Remote metastases (before 2003) | 1=yes, 2=no, 9=unknown or not applicable  blank=missing - i) after 2003, ii) missing |
| 17 | NO\_TREATMENT | No treatment (from 2008) | 0=no/NA/missing, 1=yes |
| 18 | SURGERY | Surgery | 0=no/NA/missing, 1=yes |
| 19 | AGE\_AT\_SURGERY | Age at surgery = date of surgery initiation - dob (In years) |  |
| 20 | DAYS\_DIAGNOSIS\_SURGERY | Days\_diagnosis\_surgery = date of surgery initiation - date of diagnosis |  |
| 21 | RADIO | Radiotherapy | 0=no/NA/missing, 1=yes |
| 22 | AGE\_AT\_RADIOTHERAPY | Age at radiotherapy = date of radiotherapy - dob (In years) |  |
| 23 | DAYS\_DIAGNOSIS\_RADIOTHERAPY | Days\_diagnosis\_radiotherapy = date of radiotherapy - date of diagnosis |  |
| 24 | CHEMO | Chemotherapy | 0=no/NA/missing, 1=yes |
| 25 | AGE\_AT\_CHEMOTHERAPY | Age at chemotherapy = date of chemotherapy initiation - dob (In years) |  |
| 26 | DAYS\_DIAGNOSIS\_CHEMOTHERAPY | Days\_diagnosis\_chemotherapy = date of chemotherapy initiation - date of diagnosis |  |
| 27 | HORMONES | Hormones therapy | 0=no/NA/missing, 1=yes |
| 28 | AGE\_AT\_HORMONE\_THERAPY | Age at hormone therapy = date of hormone therapy initiation - dob (In years) |  |
| 29 | DAYS\_DIAGNOSIS\_HORMONE\_THERAPY | Days\_diagnosis\_hormone therapy = date of hormone therapy initiation - date of diagnosis |  |
| 30 | BIOLOG | Biological therapy | 0=no/NA/missing, 1=yes |
| 31 | AGE\_AT\_BIOLOGICAL\_THERAPY | Age at Biological therapy = date of Biological therapy initiation - dob (In years) |  |
| 32 | DAYS\_DIAGNOSIS\_BIOLOGICAL\_THERAPY | Days\_diagnosis\_Biological therapy = date of Biological therapy initiation - date of diagnosis |  |
| 33 | OTHER\_THERAPY | Other therapy | 0=no/NA/missing, 1=yes |
| 34 | AGE\_AT\_OTHER\_THERAPY | Age at Other therapy = date of Other therapy initiation - dob (In years) |  |
| 35 | DAYS\_DIAGNOSIS\_OTHER\_THERAPY | Days\_diagnosis\_Other therapy = date of Other therapy initiation - date of diagnosis |  |
| 36 | DEATH | Death indicator | 0= alive, 1=died |
| 37 | YEAR\_OF\_DEATH | Year of death |  |
| 38 | AGE\_AT\_DEATH | Age at Death = date of death - dob (In years) |  |
| 39 | SURVIVAL\_TIME | Survival\_time = date of death - date of diagnosis. Censored at 31 Dec of the following data year (In years). For example, If the latest data is till 2014, it is matched till 31 Dec 2015. |  |
| 40 | DEATH\_CODE | Death cause (in ICD codes) | refer to ICD-9 (before 2012) and ICD-10 manual (2012 & after). |
| 41 | DEATH\_CAUSE | Death cause (text description) |  |

**Note 1:** Treatment data of invasive cancers only is limited to 6 months from the date of diagnosis and only the first treatment of each type. Data not available from private clinics

**Note 2**: Carcinoma in-situ cancers are included

**Note 3:** Item 4-5, 19-20, 22-23, 25-26, 28-29, 31-32, 34-35, 38-39 are derived variables