

# Get Data Out technical documentation: Treatment

GDO 0023, 2021-05-10

## Background

The fact of treatment with resection, chemotherapy, and radiotherapy (in the initial treatment period) is defined as a “yes/no” binary indicator for each malignant cancer registered. Three variables indicate whether any resection, chemotherapy, or radiotherapy was received. There are also 8 mutually exclusive variables that indicate the combination of resection, chemotherapy, and radiotherapy received.

### 1. Datasets used

The AV2018.AV\_TREATMENT\_TABLE\_1318\_4P6 dataset in CASREF01 and the end of year AV2018.AV\_TUMOUR dataset is used to produce these figures. The methodology for creating AV2018.AV\_TREATMENT\_TABLE\_1318\_4P6 is detailed in the [CAS-SOP #4.6\\_treatments1318](#).

- AV2018.AV\_TUMOUR was used to select the cohort, tumour and demographic details as well as provide links to other datasets via the ID fields.
- AV2018.AV\_TREATMENT\_TABLE\_1318\_4P6 was used to identify treatments for the tumours in the cohort captured by the cancer registry, Inpatient Hospital Episodes Statistics, the Radiotherapy Datasets collected by NATCANSAT, and the Systematic Anti-Cancer Therapy dataset (see [SOP](#) for details).

### 2. Cohort

The cohort was defined to align with the cohort specified in [GDO\\_0021\\_tech\\_doc\\_Incidence](#), which in turn is generated from the methodology and standard restrictions in the [CAS SOP #1 – Counting Cancer Cases](#). This SOP was followed, with the additional exclusions of testicular tumours in female patients and ovarian tumours in male patients for tumours whose ICD code starts with D. (The SOP handles C-coded tumours but does not yet perform any exclusions on D-coded tumours: sex-exclusions are performed on all tumours for our dataset). The cohort was created from the end of year

AV2018.AV\_TUMOUR dataset. Cases not meeting the SOP criteria are excluded from the cohort. Once the restrictions have been applied the cohort is refined using diagnosis year and site codes to select the relevant cohort of interest

The original treatment flags were only defined for ICD-10 C-coded tumours. In this release, the cohort has been widened to run for some D-coded tumours in line with [CAS-SOP #4.6\\_treatments1318](#), including: brain, bladder, urethra, renal pelvis, and ureter and kidney. However, the flags are not produced for all D-coded tumours covered by GDO – D-coded sarcomas still have no treatment flags. Where data cannot be calculated because the group includes D-coded tumours, this is denoted by this value '.d'.

In some cases, full treatment data could not be calculated because surgery codes have not yet been defined by the therapeutics team. This means that while chemotherapy and radiotherapy data can be provided, both individual and combinatory treatments involving surgical resection cannot be reported on. This issue affects bone cancer, soft tissue and peripheral nerve cancer, most sarcomas, and D-coded tumours of bladder, urethra, renal pelvis and ureter and kidney. Where data cannot be calculated because the group lacks definition of the relevant surgery codes, this is denoted by this value '.m'.

### 3. Definitions of treatments

Treatment with resection, chemotherapy, and radiotherapy was defined in the AV2018.AV\_TREATMENT\_TABLE\_1318\_4P6 dataset. The timeframe after diagnosis to include as part of a patient's first course of treatment varies by cancer site and treatment. Resections were also defined using a list of OPCS4 codes, agreed with clinicians. These are detailed in [CAS-SOP #4.6\\_treatments1318](#). The 8 mutually exclusive variables were calculated from the figures in AV\_TREATMENT\_TABLE\_1318\_4P6.

### 4. Variation from published figures

GDO aims to produce clinically meaningful groupings of cancer sites. Care should be taken when comparing these figures with earlier publications, as the definition of cancer sites may vary. For more information, see the supporting information for incidence figures for each cancer site on the [Get Data Out website](#).

## 5. Other caveats

Where a group size is very small, data are not available as a measure to protect patient confidentiality. This is indicated with '.k'.

“Other care” represents the group of patients who had no record of chemotherapy, tumour resection, or radiotherapy in the time frame assessed. This may include patients who received other treatments (such as hormonal therapy or management of symptoms), treatment outside of the time frame assessed, treatment in a private setting, or there may be data missing from the datasets used.

## 6. QA

The QA was done internally by the Get Data Out team.

## 7. Alterations to technical specification

Column names for the treatment data have been changed since our previous publication. We know this may make difficulties for people loading the data into standard templates, and we aim to minimise this. The rational for these changes are

- Merging of 'cohort' columns – the underlying cohort of 'all tumours' is the same for all types of treatment and publishing the same count of tumours 11 times bloats the file without adding information. Rather than having the same cohort repeated for each treatment type, there is now one column representing the treatment cohort
- Removal of '(95%)' from column names – this is standardising with the Routes file. The characters '(', ')', and '%' in column names cause difficulties when loading data into some software such as R

Changes to the column names from the previous release are documented below:

Previous column name	New column name
CT cohort	Treatment cohort
RT cohort	Treatment cohort
Surgery cohort	Treatment cohort
CT only cohort	Treatment cohort
RT only cohort	Treatment cohort

Surgery only cohort	Treatment cohort
CT and RT cohort	Treatment cohort
Surgery and CT cohort	Treatment cohort
Surgery and RT cohort	Treatment cohort
Surgery and CT and RT cohort	Treatment cohort
Other care cohort	Treatment cohort
CT percentage LCI (95%)	CT percentage LCI
CT percentage UCI (95%)	CT percentage UCI
RT percentage LCI (95%)	RT percentage LCI
RT percentage UCI (95%)	RT percentage UCI
Surgery percentage LCI (95%)	Surgery percentage LCI
Surgery percentage UCI (95%)	Surgery percentage UCI
CT only percentage LCI (95%)	CT only percentage LCI
CT only percentage UCI (95%)	CT only percentage UCI
RT only percentage LCI (95%)	RT only percentage LCI
RT only percentage UCI (95%)	RT only percentage UCI
Surgery only percentage LCI (95%)	Surgery only percentage LCI
Surgery only percentage UCI (95%)	Surgery only percentage UCI
CT and RT percentage LCI (95%)	CT and RT percentage LCI
CT and RT percentage UCI (95%)	CT and RT percentage UCI
Surgery and CT percentage LCI (95%)	Surgery and CT percentage LCI
Surgery and CT percentage UCI (95%)	Surgery and CT percentage UCI
Surgery and RT percentage LCI (95%)	Surgery and RT percentage LCI
Surgery and RT percentage UCI (95%)	Surgery and RT percentage UCI
Surgery and CT and RT percentage LCI (95%)	Surgery and CT and RT percentage LCI
Surgery and CT and RT percentage UCI (95%)	Surgery and CT and RT percentage UCI
Other care percentage LCI (95%)	Other care percentage LCI
Other care percentage UCI (95%)	Other care percentage UCI

Any questions, please contact [charlotte.eversfield@phe.gov.uk](mailto:charlotte.eversfield@phe.gov.uk),  
[sally.vernon@phe.gov.uk](mailto:sally.vernon@phe.gov.uk), [anna.cooper@phe.gov.uk](mailto:anna.cooper@phe.gov.uk), [sean.mcphail@phe.gov.uk](mailto:sean.mcphail@phe.gov.uk)