# Summary

This document describes the datasets included in project 1718-0026 (approved by the Public Benefit Privacy Panel for Health and Social Care in October 2018). This is a linked dataset of Scottish CRC patients for the period 2006-2018.

At present, storage and access to this data is limited to the named researchers on the project within Scotland and within the agreed time frame for the storage of the data.

In what follows is a description of some of the key CRC Scottish datasets alongside an overview of the variables included in them. All data sets are ready to be linked to one another through a unique patient ID.

# Data Profiles:

## Scottish Cancer Registry (SMR06)

This dataset includes information on all new diagnoses of cancer occurring within Scotland. These data are collected by Public Health Scotland and contain diagnostic, staging and treatment information on on all cancers, including every colorectal and anal cancer diagnosed in Scotland. The SMR06 data is routinely linked with NRS deaths data and hospital admissions data as part of the Information Services Division linked data catalogue. The CORECT-R Scotland database contains the SMR06 records for all patients who had a diagnosis of colorectal cancer between January 2006 and December 2018. All SMR06 records for non-colorectal cancer diagnoses are included if the patient also had a non-colorectal cancer diagnosis during the study period.

Data Summary

| | |
|---|---|
| Data Provider | **Public Health Scotland** |
| Temporality of the data | *January 2006 – December 2018* |
| Geographical extent | Scotland |
| Data tables | Patient |
| Further information and data dictionary | https://www.isdscotland.org/Health-Topics/Cancer/Scottish-Cancer-Registry/Cancer-Metadata/_docs/SMR06-Current-Dataset-from-20190101.pdf |

## National Records of Scotland (NRS) Deaths

The NRS are responsible for the registration of all life events occurring in Scotland including births, deaths, marriages, civil partnerships and adoptions. They are also responsible for Scottish census. The vital events deaths data set contains information on deaths including the leading cause of death, other causes of death, place of death, duration of illness and much more. For the purposes of the CORECT-R Scotland data, deaths data were collected for any CRC patient who died throughout the study period (2006-2019).

Data Summary

| | |
|---|---|
| Data Provider | **National Records of Scotland** |
| Temporality of the data | *January 2006 – August 2019* |
| Geographical extent | Scotland |
| Data tables | NRS Vital Events- Deaths |
| Further information and data dictionary | https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/deaths<br><br>https://www.nrscotland.gov.uk/statistics-and-data/statistics/statistics-by-theme/vital-events/deaths/deaths-background-information |

## Scottish Morbidity Records (SMR)

The Scottish Morbidity Records (SMR) contain healthcare data for individual patients. There are four main SMR series for the general types of healthcare received during an episode and/or the nature or status of the patient. Those are outpatient attendance (SMR00), general/acute inpatient and day case (SMR01), maternity inpatient and day case (SMR02) and mental health inpatient and day case (SMR04).

The largest of the SMR series, the SMR01 database, contains episode level data for all general/acute inpatient or day cases in Scottish NHS hospitals or Scottish NHS beds in non-NHS-institutions. An SMR01 record is generated for an inpatient or day case for the following reasons: when they are admitted to an NHS hospital from a location external to the NHS; when they are admitted to a contracted NHS bed in a non-NHS-institution; when they change speciality; when they transfer from another NHS hospital; when they change consultant but not specialty. Further, an SMR01 record is generated when an inpatient moves into and/or out of one of the valid significant facilities and when they return to hospital after been on pass for more than 5 days.

The Scottish CORECT-R linked data set contains all concurrent and historic (from January 1997) SMR01 records for patients who had a colorectal cancer diagnosis during the study period (January 2006-December 2018) as identified from the SMR06 registry.

Data Summary

| | |
|---|---|
| Data Provider | **Public Health Scotland** |
| Temporality of the data | *January 1997-August 2019* |
| Geographical extent | Scotland |
| Data tables | SMR01- General/acute inpatient and day case |
| Further information | https://www.ndc.scot.nhs.uk/Dictionary-A-Z/Definitions/index.asp?Search=S&ID=460&Title=SMR01%20-%20General/Acute%20Inpatient%20and%20Day%20Case |
| Data dictionaries | https://www.ndc.scot.nhs.uk/Data-Dictionary/SMR-Datasets//SMR01-General-Acute-Inpatient-and-Day-Case/ |

## Chemocare

Chemocare data is held separately by the three regional cancer networks in Scotland. Those are the South East Cancer Network (SCAN), West of Scotland Cancer Network (WoSCAN) and the North of Scotland Cancer Network (NoSCAN). Each network uses the Chemocare system to electronically record chemotherapy prescribing information for all cancer patients treated within their respective cancer network. This includes information on the patient including their height and weight, the drugs and doses prescribed, regimens etc.

The Chemocare information in the Scottish CORECT-R is limited to only those individuals who have had a diagnosis of colorectal or anal cancer between January 2013 and June 2019. These records can be linked to all other datasets held in the Scottish CORECT-R. Given that Chemocare is more up to date than the records held in SMR06, Chemocare may contain patients who are not present in the SMR06 registry. Moreover, there are some differences in the recording of data between the three cancer networks. Where this is the case, this is highlighted in the data table.

Data Summary

| | |
|---|---|
| Data Provider | **WoSCAN; NoSCAN; SCAN** |
| Temporality of the data | *January 2013 - June 2019* |
| Geographical extent | Scotland |
| Data tables | Chemocare |
| Further information | |
| Data dictionary | There is currently no data dictionary for Chemocare |

## Quality Performance Indicators (QPI)

As part of the NHS Scotland Healthcare Quality Strategy in 2010, the National Quality Performance Indicators (QPIs) were developed. These are a set of cancer specific, outcome focussed, evidence based indicators used to drive quality improvement in cancer care. Currently, these indicators are in place for 18 tumour types, including colorectal cancer. The QPI audits are carried out by each of the three cancer networks across Scotland and the data are curated by Public Health Scotland.

The Scottish CORECT-R include the Colorectal QPI indicators for those patients who had a colorectal cancer diagnosis during the period January 2013 to March 2018.

Data Summary

| | |
|---|---|
| Data Provider | **Public health Scotland (collected by WoSCAN; NoSCAN; SCAN)** |
| Temporality of the data | *January 2013 – March 2019* |
| Geographical extent | Scotland |
| Further information | https://www.ndc.scot.nhs.uk/National-Datasets/data.asp?ID=2&SubID=21 |

| | |
|---|---|
| Data dictionary | https://www.isdscotland.org/Health-Topics/Cancer/Cancer-Audit/docs/Colorectal/Colorectal-Cancer-QPI-Dataset-V3-4-Final.pdf |

## Data Summary

| | |
|---|---|
| Data Provider | National Safe Haven Scotland (Project number: 1718-0026) |
| Temporality of the data | *January 1997-June 2019 (dates vary depending on data table)* |
| Geographical extent | Scotland |
| Data tables | SMR06 |
| | NRS Deaths |
| | SMR01 |
| | QPI |
| | Chemocare |

## Content

In what follows, we outline four data tables from the first batch of the CORECT-R Scotland data. Those are for NRS deaths, SMR06, SMR01, QPI and Chemocare. Some variables within the datasets are derived using pre-existing variables in that data set. The main spine of patients comes from the SMR06 registry. From here, all patients can be linked to their SMR01 records via their master ID. Further, if the patient died during the study period their SMR06 record can also be linked to their NRS deaths record. Patient records can also be linked to Chemocare and QPI, though it is possible that some patients who are present in Chemocare, are not present in SMR06, due to Chemocare covering a more recent time period than SMR06.

| SMR06 data set (2006-2018) | | | | |
|---|---|---|---|---|
| **Data Item** | **Variable Name** | **Description of field content** | **Format** | **Further info** |
| **Identifiers** | | | | |
| Master ID | master_index | Pseudononymised person ID | Text | |
| **Patient information** | | | | |
| Date of birth | dob_fmt | Date of Birth: MMYYYY | Date | |
| Gender | sex | Sex 0 = Not Known; 1 = Male; 2 = Female; 9 = (includes not stated by patient, or not recorded) | Number | |
| Age | age_in_years | Age at incidence date | Number | Derived from dob_fmt |
| 10 year age bands | age_bands | "<35" "35-44" "45-54" "55-64" "65-74" "75-84" "85+" | Number | Derived from age_in_years |
| Date of death | dod_fmt | Date of Death: DDMMYYYY | Date | |
| Vital status | died | Vital status of patient. 0 = No, 1 = Yes | Number | Derived |
| Age of death | age_died | Age at death | Number | |
| Cause of death | cause_i | Cause of death i where i = 1,...,8. ICD10 Codes. | Text | |
| Died from crc | crc_death | CRC on death certificate as cause of death. 1 = Yes, 0 = No | Number | |
| Emigration date | embarkation_date | The date of emigration from Scotland: DDMMYYYY | Date | |
| End of follow up | end_of_follow_up | Date last observed (died, left country or censored): DDMMYYYY | Date | Derived from date of death, left country or final patient death. |
| Left Scotland | left_scotland | Patient left Scotland (based on presence of embarkation date) | Number | Derived from embarkation date |
| Survival time | survival | Years and months from date of incidence to date of follow up | Number | Derived from date of incidence and date of follow up. |
| Incidence date | incidence_date_fmt | Date of incidence: DDMMYYYY | Date | |
| Death certificate | death_certificate | The case has FIRST come to light ONLY as a result of a death. 0 = No, 1 = Yes | Number | |
| Death certificate only | death_certificate_only | The case has been registered from the death certificate only, since no other evidence of the tumour can be found. 0 = No, 1 = Yes | Number | |
| SIMD Decile | simdyear_sc_decile | Where year = 2004,2006,2009,2012 and 2016. Scottish Index of Miltiple Deprivation Decile. | Number | |
| SIMD Quintile | simd2016_sc_quintile | Where year = 2004,2006,2009,2012 and 2016. Scottish Index of Miltiple Deprivation Quintile. | Number | |

| | | | | |
|---|---|---|---|---|
| Urban Rural (6 score) | ur6_year | Where year = 2003,…2016. Scottish Government 6 fold urban/rural classification. 1 = "Large urban area" 2 = "Other urban area" 3 = "Accessible small town" 4 = "Remote small town" 5 = "Accessible rural" 6 = "Remote rural" | Number | |
| Urban Rural (8 score) | ur8_year | Where year = 2003,…2016. Scottish Government 8 fold urban/rural classification. | Number | |
| **Tumour information** | | | | |
| ICD10 Code | icd10s_cancer_site | The anatomical site of origin of the primary tumouor, format ICD10 | Text | |
| ICD02 Code | icdo2_icdo2 | The anatomical site of origin of the primary tumouor, format ICD02 | Text | |
| ICD03 Code | type_icdo3 | The histology of the tumour and comprises the first four digits of the ICD03 morphology | Text | |
| Grade classification system | grade_classification | The classification system used for grading the tumour. 1 = "Grading for Breast Cancer" 2 = "ICD0/UICC grading system" 3 = "Gleason Score (prostate)" 5 = "Fuhrman Nuclear Grade" 6 = "WHO grade for brain and CNS tumours" 8 = "Other" 9 = "Not determined/not stated/not appicable" | Number | |
| Differentiation | grade_cell_type | Indicates the degree of differentiation of malignant tumours. 0 - 10, G1, G2, GX. | Text | |
| Side | side | This indicates the side or laterality in the case of paired organs. 0 = "Not applicable" 1 = "Right" 2 ="Left" 3 = "Bilateral" 9 = "Not known" | Number | |
| Detection method | method_1st_detection | Indicates how the tumour was first detected. 1 = "Screening examination" 2 = "Incidental finding" 3 = "Clinical presentation" 4 = "Incidental finding at autopsy" 5 = "Interval Cancer" 8 = "Other" 9 = "Not known" | Number | |
| Most valid base of diagnosis | mvb_diag | Most valid base of diagnosis: indicates the method judged to have provided or validated the diagnosis during the course of the illness. 1 = "Clinical only" 2 = "Clinical investigation (including x-ray, ultrasound, etc." 3 = "Exploratory surgery/endoscopy/autopsy (without concurrent or previous histology)" 4 = "Specific biochemical and/or immunological tests" 5 = "Cytology (including blood film or bone marrow aspirate)" 6 = "Histology of metastasis" 7 = "Histology of primary" 8 = "Autopsy with concurrent or previous histology" 9 = "Not known" 10 = "Death certificate" | Number | |
| Microscopic confirmation | hist_ver | Microscopic confirmation of the histological or cytological diagnosis. 1 = Verified, 2 = Not Verified. | Number | |
| Microinvasive | microinvasive | Degree of invasion which is not associated with any risk of nodal metastasis and is sufficiently small to treat by local or conservative means. 0 = No, 1 = Yes, 9 = Not known | Number | |
| Clinical T Stage | stage_clinical_t | Indicates the extent of the spread of the tumour at diagnosis in terms of clinical findings. Stage is associated with invasive tumours only. T-size/extent of primary tumour based on clinical examination +- imaging | Text | |
| Clinical N Stage | stage_clinical_n | Condition of regional lymph nodes/glands based on clinical examination +- imaging. | Text | |

| | | | | |
|---|---|---|---|---|
| Clinical M Stage | stage_clinical_m | Indicates distant metastates | Text | |
| Pathologic T Stage | stage_pathologic_t | Indicates the extent of the spread of the tumour at diagnosis in terms of the pathalogical findings. Stage is associated with invasive tumours only. T-size/extent of primary tumour based on clinical examination +- imaging. Breast and lung | Text | |
| Pathologic N Stage | stage_pathologic_n | Condition of regional lymph nodes/glands based on clinical examination +-imaging | Text | |
| Pathologic M Stage | stage_pathologic_m | Indicates distant metastates | Text | |
| Dukes Stage (full) | stage_colorectal | Indicates the extent of spread of the invasive tumour at diagnosis in terms of the pathological and/or clinical findings for Socrates. Extent of primary tumour for Colorectal Cancer Dukes staging is primarily based on histological findings. Stages: A,B,C,C1,C2,D, Unknown. | Text | |
| Dukes Stage | dukes_stage | Duke's staging of colorectal cancer. 1 = "Duke's Stage A" 2 = "Duke's Stage B" 3 = "Duke's Stage C" 4 = "Duke's Stage D" 9 = "Duke's Stage Unknown" | Number | Derived |
| CRC ICD10 Codes | crc_icd10 | CRC ICD-10 code. 0 = "Other" 1 = "C18:Colon" 2 = "C19: Rectosigmoid junction" 3 = "C20: Rectum" 4 = "C21:Anus/Anal Canal" | Number | |
| CRC Type | crc_type | Type of colorectal cancer. 1 = "Colon/Sigmoid" 2 = "Rectal" | Number | Derived from crc_icd10 |
| CRC Flag | crc_flag | Diagnosed with colon, sigmoid or rectal cancer. 0 = No, 1 = Yes | Number | Derived from crc_type |
| Other cancer flag | other_cancer_flag | Patient had a non-crc diagnois during the study period. 0 = No, 1 = Yes | Number | Derived from icd10s_cancer_site |
| Number of other cancers | count_other_cancer | Number of non-crc diagnoes the patient had during the study period | Number | |
| Pervious cancer | previous_cancer | Patient had a prior cancer diagnosis (pre-2006) | Number | |
| Previous CRC cancer | previous_crc_cancer | Patient had a prior CRC cancer diagnosis (pre-2006) | Number | |
| **Treatment** | | | | |
| Nodes examined | nodes_examined | Pathological nodal status -Indicates what regional lymph nodes were examined. 0 = "No regional lymph nodes removed or aspirated" 1 = "Aspiration or biopsy of regional lymph node" 2 = "Sentinel lymph node biopsy" 3 = "Regional lymph node dissection" 4 = "Not known" | Number | |
| Number of nodes examined | no_of_nodes_examined | Indicates how many of the regional lymph nodes were examined. | Number | |
| Positive nodes | positive_nodes | Indicates if any of the regional lymph nodes were positive. 1 = Yes, 0 = No, 9 = Not known | Number | |
| Number of positive nodes | no_positive_nodes | Pathological nodal status -Indicates how many of the regional lymph nodes were positive.0 onwards, Not Known. | Text | |

| | | | | |
|---|---|---|---|---|
| Diagnosis institution | hosp_gp_diag | Institution code in which the diagnosis was first made. | Text | |
| Health Board of diagnosis | smr06_health_board | Health board in which the diagnosis was first made according to SMR06 | Text | Derived from hosp_gp_diag |
| Referred to radiotherapy | ref_to_rad | Referred to radiotherapy department 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |
| Radiotherapy | treated_with_rad | Treated with radiotherapy. 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |
| Type of radiotherapy | type_radio_type | Type of radiotherapy administered from 01.01.2006 (variable added to the file during June 2018) 1 =Brachytherapy, 2= External beam/Teletherapy, 3 =Proton beam therapy, 4 =Radioisotope therapy, 9 =Not Known | Number | |
| Radiotherapy to primary site | rad_to_primary | Radiotherapy to primary site. 1 = Yes, 0 = No, 9 = Not known | Number | |
| Radiotherapy to metastases | rad_to_mets | Radiotherapy to Metastases. 1 = Yes, 0 = No, 9 = Not known | Number | |
| Other radiotherapy | rad_other | Radiotherapy treatment Other such as organ ablation or prophylaxis. 1 = Yes, 0 = No, 9 = Not known | Number | |
| Date of first radio therapy | dor_fmt | Date of first radiotherapy treatment: YYYY/MM/DD | Date | |
| Institution radiotherapy | hosp_1st_rad | Institution code/practice code of hospital/GP practice of first radiotherapy (if applicable) | Text | |
| Chemotherapy | chemo | Indicates if the patient has had systemic chemotherapy treatment. 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |
| Date of first chemo | doc_fmt | Date of first chemo treatment: YYYY/MM/DD | Date | |
| Institution chemotherapy | hosp_gp_1st_chemo | Institution code/practice code of hospital/GP practice of first chemotherapy (if applicable) | Text | |
| Hormone therapy | horm_therapy | Indicates if the patient has had hormone therapy treatment. 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |
| Date of first hormone therapy | doh_fmt | Date of first hormone treatment: YYYY/MM/DD | Date | |
| Institution hormone therapy | hosp_gp_1st_hormone_therapy | Institution code/practice code of hospital/GP practice that initiated the first hormone therapy (if applicable) | Text | |
| Surgery | surgery | Indicates if the patient has been treated with surgery. 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |
| Date of first surgery | dos_fmt | Date of first surgical treatment: YYYY/MM/DD | Date | |
| Institution surgery | hosp_gp_1st_surgery | Institution code/practice code of hospital/GP practice where a surgical treatment was first carried out (if applicable) | Text | |
| Palliative surgery | palliative_surgery | Treated with palliative surgery (variable added to the file during June 2018). 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |

| | | | | |
|---|---|---|---|---|
| Date of first palliative surgery | dops_fmt | Date of first palliative surgical treatment: YYYY/MM/DD | Date | |
| Institution palliative surgery | hosp_gp_1st_pal_surgery | Institution code/practice code of hospital/GP practice where first palliative surgery was carried out(if applicable) | Text | |
| Immunotherapy | type_immun_type | Indicates if immunitherapy/biotherapy was administered. From 01.01.1997 (variable added to the file during June 2018) 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |
| Date of first immunotherapy | doit_fmt | Date of first immunotherapy type treatment: YYYY/MM/DD | Date | |
| Institution immunotherapy | hosp_gp_1st_immun_type | Institution code/practice code of hospital/GP practice that initiated the first immunotherapy (if applicable) | Text | |
| Other therapy | other_therapy | Patient treated with other therapy. 1 = Yes, 0 = No, 7 = Planned, 9 = Not known | Number | |
| Date of first other therapy | doot_fmt | Date of first other therapy treatment: YYYY/MM/DD | Date | |
| Institution other therapy | hosp_gp_1st_other_therapy | Institution code/practice code of hospital/GP practice that initiated the first other therapy (if applicable) | | |
| Other therapy | type_other_therapy | Indicates the type of other therapy treatment carried out. | Text | |
| Objectives of treatment | therapy_objectives | Objective of treatment. 1 = "Curative intent" 2 = "Non-curative intent (palliative)" 9 = "Not Known" | Number | |

| **NRS deaths data set (2006-2019)** | | | | |
|---|---|---|---|---|
| **Data Item** | **Variable Name** | **Description of field content** | **Format** | **Further info** |
| **Identifiers** | | | | |
| Master ID | master_index | Pseudononymised person ID | Text | |
| **Patient information** | | | | |
| Date of birth | dob_fmt | Date of Birth: MMYYYY | Date | |
| Date of death | dod_fmt | Date of death: DDMMYYYY | Date | |
| Gender | sex | Sex 1 = Male; 2 = Female | Number | |
| Age died | age_died | Age at death. | Number | Derived from dob_fmt and dod_fmt |
| SIMD Decile | simdyear_sc_decile | Where year = 2004,2006,2009,2012 and 2016. Scottish Index of Miltiple Deprivation Decile. | Number | |
| SIMD Quintile | simd2016_sc_quintile | Where year = 2004,2006,2009,2012 and 2016. Scottish Index of Miltiple Deprivation Quintile. | Number | |

| | | | | |
|---|---|---|---|---|
| Urban Rural (6 score) | ur6_year | Where year = 2003,…2016. Scottish Government 6 fold urban/rural classification. 1 = "Large urban area" 2 = "Other urban area" 3 = "Accessible small town" 4 = "Remote small town" 5 = "Accessible rural" 6 = "Remote rural" | Number | |
| Urban Rural (8 score) | ur8_year | Where year = 2003,…2016. Scottish Government 8 fold urban/rural classification. | Number | |
| **Details of death** | | | | |
| Underlying cause of death | underlying_cause_of_death | The disease or injury which initiated the chain of morbid events leading directly to death, or the accident/act which produced the fatal injury. ICD10 codes. | Text | |
| Other causes of death | cause_of_death_code_i | Other causes of death mentioned on the death certificate, after the primary cause of death. ICD10 codes. Where i = 0,…,9. | Text | |
| Duration of illness in years | duration_of_illness_years_1i | Approximate interval between onset of illness and death in years. Where i= a,b,c,d. | Number | |
| Duration of illness in months | duration_of_illness_months_1i | Approximate interval between onset of illness and death in months. Where i= a,b,c,d. | Number | |
| Duration of illness in days | duration_of_illness_days_1i | Approximate interval between onset of illness and death in days. Where i= a,b,c,d. | Number | |
| Place of death | place_of_death | Place where death occurred. 0 = "Institution invalid or irrelevant" 1 = "NHS Hospital" 2 = "Home/Private Address" 3 = "Hospice" 4 "Private care homes and care homes" 5 "Homes for the elderly" 6 "Private hospital" 7 "Other" | Number | |
| Institution of death | institution | The institution code for where the death occurred. | Text | |
| Primary household occupation | occupation | Primary household occupation code. | Text | |
| Primary household occupation group | Major_occ_group | Primary household major occupation group. 0 = "Large employers and higher managerial occupations" 1 = "Higher professional occupations" 2 = "Lower managerial and professional occupations" 3 = "Intermediate occupations" 4 = "Small employers and own account workers" 5 = "Lower supervisory and technical occupations" 6 = "Semi-routine occupations" 7 = "Routine occupations" 8 = "Never worked and long term unemployed" 9 = "Students, not stated or not classifiable" | Number | Derived from occupation |
| Occupation | deceased_occupation_code | Classification of the deceased persons occupation code. | Text | |
| Major occupation group | dec_major_occ_group | Deceased persons major occupation group. 0 = "Large employers and higher managerial occupations" 1 = "Higher professional occupations" 2 = "Lower managerial and professional occupations" 3 = "Intermediate occupations" 4 = "Small employers and own account workers" 5 = "Lower supervisory and technical occupations" 6 = "Semi-routine occupations" 7 = "Routine occupations" 8 = "Never worked and long term unemployed" 9 = "Students, not stated or not classifiable" | Number | Derived from deceased_occupation_code |

| | | | | |
|---|---|---|---|---|
| Employment status | deceased_employment_status | The employment status of the deceased. 0 = “Other- student, unemployed, not available, etc” 1 = “Employee, apprentice, armed forces- other rank etc” 2 = “Manager, superintendent, armed forces- officer etc” 3 = “Supervisor, foreman, charge hand etc” 4 “Self-employed- with employees” 5 “Self-employed- without employees” | Number | |
| Colorectal cancer related death | crc_death | An indicator that shows if any of the causes of death mentioned on the death certificate were due to colorectal cancer. 0 = Non-CRC death, 1 = CRC death | Number | |
| Country of residence | country_of_residence | Country of residence code for the deceased person. IS03166 codes. | Text | |
| Scottish resident | scottish_resident | Indicates if the deceased was a Scottish resident. 0 = No, 1 = Yes | Number | Derived from country_of_residence |
| Health Board | deaths_health_board | Health board where the death occurred. A= "Ayrshire and Arran" B = "Borders" C = "Argyll and Clyde" F = "Fife" G = "Greater Glasgow" H = "Highland" L ="Lanarkshire" N = "Grampian" R = "Orkney" S = "Lothian" T = "Tayside" V = "Forth Valley" W = "Western Isles" Y = "Dumfries and Galloway" Z = "Shetland" | Text | Derived from institution. |

| **SMR01 data set (1997-2018)** | | | | |
|---|---|---|---|---|
| **Data Item** | **Variable Name** | **Description of field content** | **Format** | **Further info** |
| **Identifiers** | | | | |
| Master ID | master_index | Pseudononymised person ID | Text | |
| **Patient information** | | | | |
| Date of birth | dob_fmt | Date of Birth: MMYYYY | Date | |
| Gender | sex | Sex 0 = Not Known; 1 = Male; 2 = Female; 9 = (includes not stated by patient, or not recorded) | Number | |
| SIMD Decile | simdyear_sc_decile | Where year = 2004,2006,2009,2012 and 2016. Scottish Index of Miltiple Deprivation Decile. | Number | |
| SIMD Quintile | simd2016_sc_quintile | Where year = 2004,2006,2009,2012 and 2016. Scottish Index of Miltiple Deprivation Quintile. | Number | |
| Urban Rural (6 score) | ur6_year | Where year = 2003,…2016. Scottish Government 6 fold urban/rural classification. 1 = "Large urban area" 2 = "Other urban area" 3 = "Accessible small town" 4 = "Remote small town" 5 = "Accessible rural" 6 = "Remote rural" | Number | |
| Urban Rural (8 score) | ur8_year | Where year = 2003,…2016. Scottish Government 8 fold urban/rural classification. | Number | |
| **Episode level information** | | | | |

| | | | | |
|---|---|---|---|---|
| Date of admission | doa_fmt | Date of admission: YYYYMMDD | Date | |
| Date of discharge | dodis_fmt | Date of discharge: YYYYMMDD | Date | |
| Year of admission | yoa | Year of admission | Number | Derived from doa_fmt |
| Health Board | smr01_health_board | Health Board. A= "Ayrshire and Arran" B = "Borders" C = "Argyll and Clyde" F = "Fife" G = "Greater Glasgow" H = "Highland" L ="Lanarkshire" N = "Grampian" R = "Orkney" S = "Lothian" T = "Tayside" V = "Forth Valley" W = "Western Isles" Y = "Dumfries and Galloway" Z = "Shetland" | Text | |
| Institution type | smr01_institution_type | Institution type. C = "Clinic Premises" H = "NHS Hospital" J = "Joint user hospital" K = "Contractural hospital" V = "Private nursing home, private hospital" | Text | |
| Type of admission | admission_type | Admission type. 10 ="Routine Admission - no additional detail added" 11= "Routine elective (i.e. from waiting list as planned, excludes planned transfers)" 12 ="Patient admitted on day of decision to admit, or following day, not for medical reasons, but because suitable resources are available" 18 ="Planned transfers" 19 ="Routine Admission - type not known" 20 ="Urgent Admission - no additional detail added" 21 ="Patient delay (for domestic, legal or other practical reasons)" 22= "Hospital delay (for administrative or clinical reasons e.g. arranging appropriate facilities, for tests to be carried out, specialist equipment, etc.)" 30 ="Emergency Admission - no additional detail added" 31= "Patient Injury - Self Inflicted (Injury or Poisoning)" 32 ="Patient Injury - Road Traffic Accident (RTA)" 33 ="Patient Injury - Home Incident (incl. assault or accidental poisoning)" 34 ="Patient Injury - Incident at Work (incl. assault or accidental poisoning)" 35= "Patient Injury - Other Injury (including assault or accidenta poisoning other than in the home or at work)" 36 ="Patient Non-Injury (e.g. stroke, MI, ruptured appendix)" 38 ="Other Emergency Admission (including emergency transfers)" 39= "Emergency Admission - type not known" | Number | |
| Type of admission group | admission_type_group | Type of admission- higher level grouping. 1 = "Routine admission" 2 ="Urgent admission" 3 ="Emergency admission" 9 = "Admission type unknown" | Number | |
| Patient category | patient_category | Patient category. 1 = "Amenity" 2= "Paying" 3 = "NHS" 4 = "Overseas visitor - liable to pay for treatment" 5 = "Overseas visitor - not liable to pay" 8 = "Other (including Hospice)" | Number | |

| | | | | |
|---|---|---|---|---|
| Continuous inpatient stay | cis_marker | Continuous inpatient stay marker- CIS is the unbroken period of care that a patient spends as an inpatient | Number | |
| Specialty | specialty | The division of medicine or dentistry covering a specific area of clinical activity | Text | |
| Significant facility | significant_facility | Type of clinical facility which is identified for clinical and/or costing purposes | Text | |
| Discharge/transfer to | discharge_transfer_to | Type of location a patient is discharged/transferred to following an episode of care | Text | |
| Discharge/transfer to group | dis_trans_to_group | Main grouping for discahrge/transfer to group. 1 = "Patient died" 10 = "Private residence" 20 = "Institution" 30 = "Temporary place of residence" 40 = "Transferred within the same health board/health care provider" 50 = "Transferred to another health board/health care provider" 60 = "Other type of location" | Number | Derived from discharge_transfer_to |
| Discharge type group | dis_type_group | Main grouping for discharge type. 10 = "Regular discharge" 20 = "Irregular discharge" 40 = "Death" | Number | Derived from discharge_type |
| Discharge type | discharge_type | Indicates whether a discharge from an inpatient or day case epsiode is regular,irregular or due to patient death. 10= "Regular Discharge - no additional detail added" 11 = "Discharge from NHS inpatient/day case care" 12 ="Transfer within the same Health Board/ Health Care Provider" 13 = "Transfer to another Health Board/ Health Care Provider" 18 = "Other type of regular discharge" 19= "Regular discharge - type not known" 20 = "Irregular Discharge - no additional detail added" 21= " Patient discharged himself/herself against medical advice" 22= "Patient discharged by relative" 28 Other type of irregular discharge" 29= Irregular discharge - type not known" 40 ="Death - no additional detail added" 41 =Death - Post Mortem" | Number | |
| HRG | hrg | Healthcare resource group. | Text | |
| Length of stay | length_of_stay | Length of stay in hospital days (at episode level) | Number | |
| **Medical conditions and treatment** | | | | |
| Main medical condition | main_condition | Main medical (or social) condition managed/investigated during the patient's stay | Text | |
| Other medical condition | other_condititon_i | Where i= 1,…,5. The i'th other medical condition managed/investigated during the patient's stay after the main condition | Text | |

| | | | | |
|---|---|---|---|---|
| Main condition colorectal cancer | main_condition_crc | Indicates the type of CRC if the main condition was CRC. 0 = "Non-CRC" 1 = "C18:Colon" 2 = "C19: Rectosigmoid junction" 3 = "C20: Rectum" | Number | |
| Any condition colorectal cancer | any_condition_crc | Indicates the type of CRC if the any condition was CRC. 0 = "Non-CRC" 1 = "C18:Colon" 2 = "C19: Rectosigmoid junction" 3 = "C20: Rectum" | Number | |
| Main operation | main_operation | Main operation carried out during patient stay (OPCS4) | Text | |
| Other operations | other_operation_i | Where i= 1,…,3. Other operation carried out during patient stay (i'th after main) | Text | |
| OPCS4 Code | opcode2 | First OPCS4 code of main_operation | Text | Derived from main_operation |
| OPCS4 Code 2B | opcode2b | Second OPCS4 code of main_operation | Text | Derived from main_operation |
| OPCS4 Code Group | opcode_derived | Grouping variable for colorectal operation. 1 = "Major resection" 2 = "Minor resection" 3 = "Bypass/stoma" 4 = "Stent" 5 = "Appendix" 6 = "Liver" 7 = "Other/currently unknown" | float | |
| OPCS4 Code Group 2B | opcode2b_derived | Grouping variable for colorectal operation where additional info is given for the same operation. 1 = "Major resection" 2 = "Minor resection" 3 = "Bypass/stoma" 4 = "Stent" 5 = "Appendix" 6 = "Liver" 7 = "Other/currently unknown" | float | |
| Acute myocardial infarction | ami | Acute myocardial infarction. 0 = No, 1 = Yes | Number | |
| Congestive heart failure | chf | Congestive heart failure. 0 = No, 1 = Yes | Number | |
| Peripheral vascular disease | pvd | Peripheral vascular disease. 0 = No, 1 = Yes | Number | |
| Cerebrovascular disease | cevd | Cerebrovascular disease. 0 = No, 1 = Yes | Number | |
| Dementia | dementia | Dementia. 0 = No, 1 = Yes | Number | |

| | | | | |
|---|---|---|---|---|
| COPD | copd | Chronic obstructive pulmonary disease. 0 = No, 1 = Yes | Number | |
| Rheumatoid disease | rheumd | Rheumatoid disease - Connective tissue disease. 0 = No, 1 = Yes | Number | |
| Peptic ulcer disease | pud | Peptic ulcer disease. 0 = No, 1 = Yes | Number | |
| Mild liver disease | mld | Mild liver disease. 0 = No, 1 = Yes | Number | |
| Diabetes no complications | diab | Diabetes no complications. 0 = No, 1 = Yes | Number | |
| Diabetes w complications | diabwc | Diabetes w complications. 0 = No, 1 = Yes | Number | |
| Hemiplegia or paraplegia | hp | Hemiplegia or paraplegia. 0 = No, 1 = Yes | Number | |
| Renal disease | renal | Renal disease. 0 = No, 1 = Yes | Number | |
| Cancer (any malignancy) | cancer | Cancer (any malignancy). 0 = No, 1 = Yes | Number | |
| Moderate or severe liver disease | msld | Moderate or severe liver disease. 0 = No, 1 = Yes | Number | |
| Metastatic solid tumour | metacancer | Metastatic solid tumour. 0 = No, 1 = Yes | Number | |
| HIV | hiv | HIV. 0 = No, 1 = Yes | Number | |

| **QPI data set (2013-2018)** | | | | |
|---|---|---|---|---|
| **Data Item** | **Variable Name** | **Description of field content** | **Format** | **Further info** |
| **Identifiers** | | | | |
| Master ID | master_index | Pseudononymised person ID | Text | |
| **Patient information** | | | | |
| Date of birth | dob_fmt | Date of Birth: MMYYYY | Date | |
| Sex | sex | Sex. 1 = Male; 2 = Female; | Number | |

| | | | | |
|---|---|---|---|---|
| Date of death | dod_fmt | Date of Death: DDMMYYYY | Date | |
| Diagnosis | | | | |
| Source of cancer referral | mrefer | Source of Cancer Referral. 1 "Primary care" 2 "Screening" 3 "Incidental" 4 "Review clinic" 5 "Cancer genetic clinic" 6 "Self-referral A&E" 7 "GP direct to hospital" 8 "Previous GP but subsequent to hospital" 11 "Primary care clinician (dental)" 12 "Private healthcare" 13 "Other" 99 "Not recorded" | Number | |
| Location of diagnosis | hosp | Location of diagnosis | Text | |
| Health board of diagnosis(QPI) | qpi_health_board | Health Board according to QPI hosp of diagnosis. A= "Ayrshire and Arran" B = "Borders" C = "Argyll and Clyde" F = "Fife" G = "Greater Glasgow" H = "Highland" L ="Lanarkshire" N = "Grampian" R = "Orkney" S = "Lothian" T = "Tayside" V = "Forth Valley" W = "Western Isles" Y = "Dumfries and Galloway" Z = "Shetland" | Derived from hosp | |
| Date of diagnosis | diagdate_fmt | Date of Diagnosis : DDMMYYYY | float | |
| Date of histological diagnosis | hdiag | Date of Histological Diagnosis | long | |
| Most valid basis of diagnosis | valid | Most valid basis of diagnosis. 1 ="Clinical only" 2 ="Clinical Investigation" 3 ="Exploratory surgery/endoscopy/autopsy" 4= "Tumour specific markers" 5= "Cytology" 6= "Histology of metastasis" 7= "Histology of primary" 99= "Not known" | Number | |
| **Staging/Imaging** | | | | |
| Staging investigations complete | sinvest | Staging investigations complete. 1 "Complete CTCAP" 2 "Complete CTCAP and MRI" 3 "Incomplete" 4 "Incomplete - Contraindications" 95 "Patient refused" 96 "Not applicable" 99 "Not recorded" | byte | |
| Data staging investigations complete | sinvestdate_fmt | Date Staging Inv completed: DDMMYYYY | Date | |
| Large bowel imaging | lbtype | Large bowel imaging. 1 "Yes Colonoscopy or CT Colonography" 2 "Incomplete" 3 "Not performed" 4 "Incomplete due to obstructing tumour" 94 "Patient died before treatment" 95 "Patient refused" 96 "Not applicable" 99 "Not recorded" | Number | |
| Date of large bowel imaging | lbdate_fmt | Date of Large Bowel Imaging : DDMMYYYY | Date | |
| **Treatment** | | | | |
| | | | | |
| Date of discussion by MDT | mdtdate_fmt | Date discussed by multidisciplinary team (MDT): DDMMYYYY | Date | |

| | | | | |
|---|---|---|---|---|
| Seen by stoma nurse | stomanurse | Seen by Stoma Nurse. 1= "Yes" 2 ="No" 95= "Patient refused" 96 ="Not applicable" 99 ="Not recorded" | Number | |
| Date seen by stoma nurse | stomandate_fmt | Date of Stoma nurse : DDMMYYYY | Date | |
| Stoma site marked | stomamark | Stoma Site Marked Pre-op. 1= "Yes" 2 ="No" 95= "Patient refused" 96 ="Not applicable" 99 ="Not recorded" | Number | |
| Date of first treatment | firsttreatdate_fmt | Date of first treatment: DDMMYYYY | Date | |
| Type of first cancer treatment | firsttreatmode | Type of First Cancer Treatment. 1= "Surgery" 2 ="Radiotherapy" 3 ="Chemotherapy" 4= "Chemoradiotherapy" 5= "Endoscopic" 7= "Supportive Care Only" 11= "Other therapy" 94= "Patient died before treatment" 95 ="Patient refused treatment" 99= "Not recorded" | Number | |
| Date of definitive treatment | deftreatdate_fmt | Date of definitive treatment : DDMMYYYY | Date | |
| ASA status | asa | American Society of Anaesthesiologists (ASA) status. The ASA PS classification globally assesses the degree of "sickness" or "physical state" prior to selecting the anaesthetic or prior to performing surgery 1= "Normal healthy patient" 2 ="Mild systemic disease" 3 ="Severe systemic disease" 4 ="Severe systemic disease constant threat to life" 5 ="Moribund" 6 ="Brain-dead" 96 ="Not applicable" 99 ="Not recorded" | Number | |
| Location of surgery | hospsurg | Location of Surgery | Text | |
| OPCS4 Code | opcode2 | Final Definitive Surgery Performed CRC | Text | |
| OPCS4 Code 2B | opcode2b | Final Definitive Surgery Performed CRC 2b | Text | |
| OPCS4 Code Group | opcode_derived | Grouping variable for colorectal operation. 1 = "Major resection" 2 = "Minor resection" 3 = "Bypass/stoma" 4 = "Stent" 5 = "Appendix" 6 = "Liver" 7 = "Other/currently unknown" | Number | |
| OPCS4 Code Group 2B | opcode2b_derived | Grouping variable for colorectal operation where additional information is given for the same operation. 1 = "Major resection" 2 = "Minor resection" 3 = "Bypass/stoma" 4 = "Stent" 5 = "Appendix" 6 = "Liver" 7 = "Other/currently unknown" | Number | |
| Total Mesorectal Excision | texcision | Total Mesorectal Excision. 1 ="Yes" 2 ="No" 96 ="Not applicable" 99= "Not recorded" | Number | |
| Type of surgical approach | surgappr | Surgical Approach. 1 ="Open" 2 ="Laparoscopic completed" 3 ="Laparoscopic converted" 4 ="TEM" 5 "TART" 96 ="Not applicable" 99= "Not known" | Number | |
| Date of definitive surgery | finsurgdate | Date of Definitive Surgery CRC: DDMMYYYY | Date | |
| Presentation type | present | Presentation type. 1 ="Elective routine" 2= "Emergency" 96= "Not applicable" 99= "Not recorded" | Number | |

| | | | | |
|---|---|---|---|---|
| Anastomotic leak | anasleak | Anastomotic Leak. 1= "Yes" 2 ="No anastomotic leak" 96 ="Not applicable" 99 ="Not known" | Number | |
| Intent of surgery | opintent | Intent of Surgery. 1= "Curative" 2= "Palliative" 96 ="Not applicable" 99 ="Not recorded" | Number | |
| Re-operation | reoper | Re-operation. 1 ="Yes" 2= "No" 96= "Not applicable" 99 ="Not recorded" | Number | |
| Extramural venous invasion | extra | Extramural venous invasion. 1 ="Present" 2 ="Not present" 96 ="Not applicable" 99= "Not recorded" | Number | |
| Circumferential margin involved | circmargin | Circumferential margin involved. 1 ="Involved" 2 ="Not involved" 96= "Not applicable" 99 ="Not recorded" | Number | |
| Location of SACT | hospsact | Location of SACT | Text | |
| Location of radiotherapy | hospradio | Location of radiotherapy | Text | |
| Neo-Adjuvant Oncology Treatment Type. | neoonc | Neo-Adjuvant Oncology Treatment Type. 1 ="Short Course RT" 2= "Long Course RT with chemo" 3 ="Long Course RT alone" 4 ="Chemotherapy" 80 ="Patient died before radiotherapy" 81 ="Patient died before SACT" 82 ="Patient died before chemoRT" 83 ="Patient refused RT" 84 "Patient refused SACT" 85 ="Patient refused chemoRT" 86 "RT contraindicated" 87 ="Chemotherapy contraindicate" 88 ="ChemoRT contraindicated" 96 ="Not applicable" 99 ="Not recorded" | Number | |
| Date Neo-adjuvant oncology treatment started | neoadjdate_fmt | Date Neo-adjuvant oncology treatment started : DDMMYYYY | Date | |
| Date Neo-adjuvant oncology treatment completed | neoadjcom_fmt | Date Neo-adjuvant oncology treatment completed : DDMMYYYY | Date | |
| Primary/Palliative /Adjuvant Oncology Treatment Type | adjonc | Primary/Palliative/Adjuvant Oncology Treatment Type. 1 ="Adjuvant Long Course RT with chemotherapy" 2= "Adjuvant chemotherapy" 3= "Adjuvant RT" 4= "Primary Chemotherapy" 5= "Primary/Radical RT" 6 ="Palliative RT" 7 ="Palliative Chemotherapy" 8= "Biological Therapy" 9 ="Palliative Chemoradiotherapy" 80 ="Patient died before radiotherapy" 81 ="Patient died before SACT" 82 ="Patient died before chemoRT" 83 ="Patient refused RT" 84 ="Patient refused SACT" 85 ="Patient refused chemoRT" 96 ="Not applicable" 99= "Not recorded" | Number | |
| Date Primary/Palliative /Adjuvant Oncology Treatment started | adjoncdate_fmt | Date Primary/Palliative/Adjuvant Oncology Treatment started : DDMMYYYY | Date | |

| | | | | |
|---|---|---|---|---|
| Date Primary/Palliative /Adjuvant Oncology Treatment completed | adjcom_fmt | Date Primary/Palliative/Adjuvant Oncology Treatment completed : DDMMYYYY | Date | |
| Health Board | hbyear | Health Board code using various year codes. Where year = 2006, 2014, 2018, 2019. | Text | |
| **Tumour characteristics** | | | | |
| Site of origin of primary tumour | site | Site of Origin of Primary Tumour. 1 ="Caecum" 2 ="Currently unknown await update" 3= "Ascending colon" 4= "Hepatic flexure" 5 ="Transverse colon" 6 ="Splenic flexure" 7= "Descending colon" 8 ="Sigmoid colon" 9 ="Overlapping lesion of colon" 10 ="Colon, unspecified" 11= "Rectum" 12 ="Not recorded" | Number | |
| Circumferential margin involved | circmargin | Circumferential margin involved. 1 ="Involved" 2 ="Not involved" 96= "Not applicable" 99 ="Not recorded" | Number | |
| Grade of differentiation | different | Grade of differentiation. 1= "Well/moderate" 2 ="Poor" 3= "Not assessable" 96 ="Not applicable" 99= "Not known" | Number | |
| Number of lymph nodes examined | lnexamine | Final total number of lymph nodes examined microscopically | Number | |
| Number of lymph nodes involved | lninvolve | Number of lymph nodes involved | Number | |
| TNM Tumour | finalt | TNM tumour | Number | |
| TNM Nodes | finaln | TNM nodes | Number | |
| TNM Metastasis | finalm | TNM Metastasis | Number | |
| Dukes staging | dukes | Dukes stage. 1 ="Dukes A" 2 ="Dukes B" 3= "Dukes C1" 4 ="Dukes C2" 5= "Dukes D" 96= "Not applicable" 99 ="Not recorded" | Number | |
| Dukes staging derived from TNM | dukes_derived_qpi | Dukes stage derived from TNM staging. 1 ="Dukes A" 2 ="Dukes B" =3 "Dukes C" 4 ="Dukes D" | Number | Derived from finalt, finaln, finalm |
| **Cancer Network specific** | | | | |
| Cancer Network | location | Indicates regional cancer network. 1 = SCAN, 2 = WoScan, 3 = NoSCAN | Number | Derived |
| SCAN | SCAN_flag | Indicates if the record is from SCAN. 0 = No, 1 = Yes | Number | |
| WoSCAN | Wos_flag | Indicates if the record is from Wo-SCAN. 0 = No, 1 = Yes | Number | |
| Distance from anal verge (SCAN only) | analverge_SCAN | Distance from Anal Verge | Number | |
| CT Chest Result (SCAN only) | cxr_SCAN | CT Chest Result. 1 ="No metastases" 2 ="Metastases" 3 ="Equivocal" 4 ="Not performed" 95 ="Patient refuses investigation" 96 ="Not applicable" 99 ="Not recorded" | Number | |

| | | | | |
|---|---|---|---|---|
| Date of CT Chest Result (SCAN only) | xdate_SCAN | Date of CT Chest Result | Date | |
| Liver imaging Result (SCAN only) | liver_SCAN | Liver imaging Result. 1 ="No metastases" 2 ="Metastases" 3 ="Equivocal" 4= "Not performed" 95 ="Patient refused investigation" 96 ="Not applicable" 99 ="Not recorded" | Number | |
| Date of liver imaging completed (SCAN only) | liverdate_SCAN_fmt | Date of liver imaging SCAN completed : DDMMYYYY | Date | |
| Circumferential Resection Margin Predicted (SCAN only) | cmarginpredict_SCAN | Circumferential Resection Margin Predicted. 1 ="Clear" 2 ="Threatened" 3 ="Involved" 96 ="Not applicable" 98= "Not assessed" 99= "Not recorded" | Number | |

| **Chemocare data set (2013-2018)** | | | | |
|---|---|---|---|---|
| **Data Item** | **Variable Name** | **Description of field content** | **Format** | **Further info** |
| **Identifiers** | | | | |
| Master ID | master_index | Pseudononymised person ID | Text | |
| **Patient information** | | | | |
| Gender | sex | Sex of patient. 1 = Male; 2 = Female | Number | |
| Patient height | height | Patient height in meters | Float | |
| Patient weight | weight | Patient weight in kg | Float | |
| Body surface area | body_surface_area | Calculation of body surface area based on height and weight (generated by Chemocare) | Float | |
| Intention of treatment | intention | Intention of treatment (as in Chemocare) | String | |
| Intention of treatment grouping | intention_group | Intention of treatment. 1 ="Adjuvant" 2 ="Curative" 3 ="Radical" 4 ="Neo-Adjuvant" 5 ="Palliative" 6 ="Peri-op" 7= "Durable Remission" 8 ="Day Case" 9 ="MAI" 10 ="INT" 11= "CON" 99= "Unknown" | Number | Derived from intention |
| Appointment date | appointment_date | Appointment date chemotherapy drug was administered: DDMMYYYY | Date | |
| Cycle number | cycle_number | Cycle of chemotherapy (generated by Chemocare) | Number | |
| Day number | day_number | Day of the cycle of chemotherapy (generated by Chemocare) | Number | |
| Drug name | drug_name | Drug name for any drug administered during chemotherapy appointment. | Text | |
| Drug dose | drug_dose | Standard drug dose per meter squared for specific drug or dose band. | Number | |
| Required dose | required_dose | Actual dose delivered to patient. | Number | |

| | | | | |
|---|---|---|---|---|
| Diagnosis | diagnosis | Diagnosis (original text from Chemocare) | Text | |
| Diagnosis grouping | Diagnosis_group | Diagnosis to differentiate colon and rectum, and may indicate the intention of treatment e.g. palliative, curative and line of treatment in metastatic setting e.g. 1st, 2nd, 3rd. . 1 ="Colon" 2 ="Rectal" 3= "Adjuvant Colon" 4 ="Adjuvant Rectum" 5 ="Adj rectum after downstaging" 6 ="Adj Rectal after primary" 7= "Adj rectum post resection" 8 "Pal Colon 1st" 9 = "Pal Colon 2nd" 10= "Pall Colon 3rd" 11= "Pall Colon 4th" 12= "Pall Rectum 1st" 13= "Pall Rectum 2nd" 14 ="Pall Rectum 3rd" 15 ="Pall Rectum 4th" 16= "Neoadj rectum" 17 ="Neoadj colon liver" 18 "Neoadj rectum liver" 19= "Periop colon" 20= "Periop rectum" 21= "Pseudomyxoma 1st" 22 ="Pseudomyxoma 2nd" 23= "Anal" 24 ="Misc" 25= "LowerGI" | Number | Derived from diagnosis |
| Regime | regime | Description of drug regimen given to patient. | Text | |
| **Cancer Network specific** | | | | |
| Cancer Network | location | Location. 1 ="SCAN" 2 ="WoSCAN" 3= "Grampian" 4 ="Highland" 5 ="Tayside" | Number | |
| Protocol (SCAN only) | protocol_SCAN | Further explanation about drug regimen. Available for SCAN only. | Text | |
| Drug type (SCAN only) | drug_type_SCAN | Type of drug e.g. chemotherapy, anti-sickness, anti-diahoerea etc. | Text | |
| Hospital (SCAN only) | hospital_SCAN | Hospital. BGH = Borders General Hospital, DRI = Dumfries and Galloway Royal Infirmary, SJH = St John's Hospital, VHF = Victoria Hospital, WGH = Western General Hospital | Text | |
| Method of dose calculation (WoSCAN. NoSCAN only) | calculation | Method of dose calculation. 1 ="Flat" 2 ="Glomerular Filtration Rate" 3 ="Surface Area" 4 ="Weight" 99= "Unknown" | Number | |
| Unit of drug (NoSCAN only) | unit | Unit of drug. 1 ="Application" 2= "Capsule" 3 ="Tablet" 4 ="Sachet" 5 ="Drop" 6 ="Gram" 7 ="Mg" 8 ="Mcg" 9 ="IU" 10 ="MU" 11= "Ml" 12= "Mmol" 13 ="Unit" 99= "Unknown" | Number | |
| Date BSA recorded (WoSCAN and Grampian only) | date_bsa_recorded | Date body surface area calculated and updated. | Date | |
| Capped (WoSCAN only) | capped | Capped dose for body surface area. 0 "No" 1 "Yes" 99 "Unknown" | Number | |
| Duration (WoSCAN only) | duration | Duration of drug prescription e.g. in days or weeks. | Text | |
| Frequency (WoSCAN only) | freq | Frequency of drug administration e.g. per-day. | Text | |

| | | | | |
|---|---|---|---|---|
| Performance status (WoSCAN, NoSCAN only) | performance_status | Performance status is an overall assessment of the functional/physical performance of the patient. This is a five point scale in which '0' denotes normal activity and '4' a patient who is 100% bed ridden. 0 =Fully active, able to carry on all pre-disease performance without restriction<br>1 =Restricted in physically strenuous activity but ambulatory and able to carry out work of a light or sedentary nature, e.g. light housework, office work<br>2 =Ambulatory and capable of self care but unable to carry out any work activities: up and about more than 50% of waking hours<br>3 =Capable of only limited self care, confined to bed or chair more than 50% of waking hours<br>4 =Completely disabled, cannot carry on any self care, totally confined to bed or chair<br>9 = Not recorded | Number | |
| Drug route (WoSCAN and NoSCAN only) | drug_route | The route the drug was administered (e.g oral, IV) | Text | |
| Drug route grouping (WoSCAN and NoSCAN only) | Drug_route_group | 1 = "Intravenous (IV)" 2 = "Oral" 3 = "Intramusculor" 4 = "Sub Cutaneous" 5 = "Topical-skin" 6 = "Eye- topical, drops, intravetrial" 7 = "Mouth wash" 8 = "Buccal" 9 = "Per rectal" 10 = "Other" 99 = "Missing" | Number | Derived from drug_route |