



Medicines & Healthcare products
Regulatory Agency



CPRD

Systemic Anti-Cancer Therapy (SACT) Data Dictionary

Version 5.0

Date: 07 October 2020

Documentation Control Sheet

Over time, it may be necessary to issue amendments or clarifications to parts of this document. This form must be updated whenever changes are made.

| Version | Affected Areas Summary of Change | Prepared By | Reviewed By |
|---------|-----------------------------------------------------------------------|------------------|------------------|
| 1.0 | First CPRD-NCRAS-SACT data release for set 14. No changes for set 15. | Helen Strongman | Rachael Williams |
| 2.0 | Updated for set 16 | Rachael Williams | Eleanor Yelland |
| 3.0 | Updated for set 17 | Helen Booth | Eleanor Yelland |
| 4.0 | Updated for set 18 | Sonia Coton | Eleanor Yelland |
| 5.0 | Updated for set 19 | Eleanor Yelland | Hilary Shepherd |

Version 1.0

- First version (20 June 2017)

Version 2.0

- Updated for set 16

Version 3.0

- Updated for set 17

Version 4.0

- Updated for set 18

Version 5.0

- Updated for set 19

Patient - One row per patient

| <i>Column description</i> | <i>Column Name</i> | <i>Details</i> | <i>Field Type</i> | <i>Valid Content</i> |
|---------------------------|--------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------------|
| CPRD patient identifier | e_patid | Unique patient identifier based on CPRD primary care data | ID | Number |
| CR patient Identifier | e_cr_patid | Unique patient identifier based on NCRAS data. In some cases, the same person may have multiple patient IDs. Patient IDs will be retained even after two patient records are found to be the same person. | ID | Number |

1. Tumour

| <i>Column description</i> | <i>Column Name</i> | <i>Details</i> | <i>Field Type</i> | <i>Valid Content</i> |
|---------------------------------|-------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------------------------------|
| CPRD patient identifier | e_patid | Unique patient identifier based on CPRD primary care data | ID | Number |
| CR patient Identifier | e_cr_patid | Unique patient identifier based on NCRAS data. In some cases, the same person may have multiple patient IDs. Patient IDs will be retained even after two patient records are found to be the same person. | ID | Number |
| Pseudonymised Tumour Identifier | e_pseudo_merged_tumour_id | Tumour Identifier (used to link to all subsequent tables) | NUMBER | Max length 6 characters |
| Consultant code | consultant_gmc_code | Pseudonymised code of the consultant who initiated the SACT treatment | NUMBER | Number |
| Consultant Specialty Code | consultant_speciality_code | Main specialty code of the consultant who initiated the SACT treatment | NUMBER | Max Length 8 characters, Range 0 - 991 |
| Organisation code | organisation_code_of_provider | Pseudonymised organisation code of the organisation acting as a Health Care Provider | NUMBER | Number |
| Primary Diagnosis | primary_diagnosis | Primary diagnosis at the initiation of SACT (at the time of decision to treat) (ICD-10) | TEXT | Max Length 6 characters |
| Cleaned Morphology | morphology_clean | Morphology ICD-O at the initiation of SACT (at the time of decision to treat) | TEXT | Max Length 5 characters |
| Stage of Tumour at Start | stage_at_start | Pre-treatment TNM stage. Record the overall clinical TNM stage grouping of the tumour. | TEXT | Max Length 5 characters |

2. Programme and Regimen

| <i>Column description</i> | <i>Column Name</i> | <i>Details</i> | <i>Field Type</i> | <i>Valid Content</i> |
|----------------------------------|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| CR patient Identifier | e_cr_patid | Unique patient identifier based on NCRAS data. In some cases, the same person may have multiple patient IDs. Patient IDs will be retained even after two patient records are found to be the same person. | ID | Number |
| Pseudonymised Tumour Identifier | e_pseudo_merged_tumour_id | Tumour Identifier | NUMBER | Max length 6 characters |
| Pseudonymised Regimen identifier | e_pseudo_merged_regimen_id | Regimen Identifier (used to link to all subsequent tables) | NUMBER | Max length 6 characters |
| Programme Number | programme_number | SACT - version 2 only. Programmes of chemotherapy are numbered according to their chronological order of commencement in the patient's disease management. | NUMBER | Max length 2 characters, Range - 1 to 99 |
| Regimen number | regimen_number | SACT - version 2 only Regimens are numbered according to their chronological order of commencement in the patient's disease management. | NUMBER | Max length 2 characters, Range - 1 to 99 |
| Treatment Intent | intent_of_treatment | Intent of SACT regimen | TEXT | Intent of SACT regimen - SACT version 2 A = Adjuvant N = Neo-Adjuvant C = Curative P = Palliative D = Disease Modification 03 - Disease Modification from 2018 |

| | | | | |
|-----------------------------------|----------------------------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | | | | <p>Intent of SACT regimen - SACT version 3</p> <p>01 = Curative - aiming to permanently eradicate disease</p> <p>02 = Palliative - Aiming to extend life expectancy</p> <p>03 = Palliative - Aiming to relieve and/or control malignancy related symptoms</p> <p>04 = Palliative - Aiming to achieve remission</p> <p>05 = Palliative - Aiming to delay tumour progression</p> <p>98 = Other</p> <p>99 = Not Known</p> |
| Regimen Analysis Group | analysis_group | <p>This is a cleaned version of variables mapping regimen information into consistent description groups. A SACT regimen identifies a standard for a combination of drugs (or single drug) given in a planned schedule</p> <p>Note: includes 'Not Chemo', 'Not matched'.</p> | TEXT | Max Length 59 characters |
| Regimen Benchmark Group | benchmark_group | <p>Benchmark_Group - Maps regimen information into consistent high-level groups (e.g. 'CHOP R – 21 days' is mapped to CHOP R or DEGARELIX, BICALUTAMIDE + GOSERELIN and TAMOXIFEN are mapped to HORMONES). Note includes 'Not Chemo', 'Not matched'.</p> | TEXT | Max Length 59 characters |
| Person Height at start of Regimen | height_at_start_of_regimen | Height in metres at the start of the SACT Regimen | NUMBER | Range -1 to 9.99 Max 2 d.p |
| Person Weight at | weight_at_start_of_regimen | Weight in kilograms at the start of the SACT | NUMBER | Range -1 to 999 |

| start of Regimen | | Regimen | | Max 3 d.p |
|----------------------------------------|------------------------------|-------------------------------------------------------------|------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Performance Status at Start of Regimen | perf_status_start_of_regimen | Performance Status Indicator of Person at start of regimen. | TEXT | <p>Performance status (Adult) - SACT version 3 0 = Able to carry out all normal activity without restriction 1 = Restricted in physically strenuous activity, but able to walk and do light work 2 = Able to walk and capable of all self care, but unable to carry out any work. Up and about more than 50% of waking hours 3 = Capable of only limited self care, confined to bed or chair more than 50% of waking hours 4 = Completely disabled. Cannot carry on any self care. Totally confined to bed or chair</p> <p>Performance status (adult and young person)- SACT version 2 0 = Able to carry out all normal activity without restriction 1 = Restricted in physically strenuous activity, but able to walk and do light work 2 = Able to walk and capable of all self care, but unable to carry out any work. Up and about more than 50% of waking hours 3 = Capable of only limited self</p> |

| | | | | |
|-------------|------------------------|-----------------------------------------------|------|---------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | | | | <p>care, confined to bed or chair more than 50% of waking hours</p> <p>4 = Completely disabled. Cannot carry on any self care. Totally confined to bed or chair</p> <p>00 = 100% - Fully active, normal</p> <p>01 = 90% - Minor restrictions in physically strenuous activity</p> <p>02 = 80% - Active, but tires more quickly</p> <p>03 = 70% - Both greater restriction of, and less time spent in, play activities</p> <p>04 = 60% - Up and around, but minimal active play; keep busy with quieter activities</p> <p>05 = 50% - Gets dressed but lies around much of the day; no active play; able to participate in all quiet play and activities</p> <p>06 = 40% - Mostly in bed; participates in quiet activities</p> <p>07 = 30% - In bed; needs assistance even for quiet play</p> <p>08 = 20% - Often sleeping; play entirely limited to very passive activities</p> <p>09 = 10% - No play; does not get out of bed</p> <p>10 = 5% - Unresponsive</p> <p>11 = 0% - Dead</p> |
| Comorbidity | comorbidity_adjustment | Indicator of whether or not patient's overall | TEXT | Y = Yes, N = No |

| | | | | |
|---------------------------|--------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------------------------------------------------------------------------------------------------------------------------|
| Adjustment Indicator | | physical state (other diseases and conditions) was a significant factor in deciding on regimen, or in varying the dose or treatment interval. Y = Yes; N= No | | |
| Decision to Treat Date | date_decision_to_treat | This is the date that the consultation between the patient and the clinician took place and a Planned Cancer Treatment was agreed | DATE | ddmmyyyy |
| Start Date of Regimen | start_date_of_regimen | This is the first administration date of the first cycle of a regimen. | DATE | ddmmyyyy |
| Clinical Trial | clinical_trial | Indicates whether an individual episode of care was delivered as part of a clinical trial. | TEXT | 01 = PATIENT is taking part in a CLINICAL TRIAL 02 = PATIENT is not taking part in a CLINICAL TRIAL 99 = Not Known |
| Chemo-Radiation Indicator | chemo_radiation | This field identifies regimens which are given as part of a combined treatment with radiation | TEXT | SACT version 2 only. Y = Yes, N = No |
| Number of Planned Cycles | number_of_cycles_planned | The number of cycles specified in the prescription. This may be the number of cycles in the standard regimen or be modified by the prescriber. | NUMBER | Max length 3 characters |

3. Cycle

| <i>Column description</i> | <i>Column Name</i> | <i>Details</i> | <i>Field Type</i> | <i>Valid Content</i> |
|----------------------------------|----------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|------------------------------------------|
| CR patient Identifier | e_cr_patid | Unique patient identifier based on NCRAS data. In some cases, the same person may have multiple patient IDs. Patient IDs will be retained even after two patient records are found to be the same person. | ID | Number |
| Pseudonymised Tumour Identifier | e_pseudo_merged_tumour_id | Tumour Identifier | NUMBER | Max length 6 characters |
| Pseudonymised Regimen identifier | e_pseudo_merged_regimen_id | Regimen Identifier | NUMBER | Max length 6 characters |
| Pseudonymised Cycle Identifier | e_pseudo_merged_cycle_id | Unique identifier for a Cycle within a Regimen (used to link to drug detail table). | NUMBER | Max length 7 characters |
| Cycle Number | cycle_number | Number of cycle | NUMBER | |
| Start Date (Cycle) | start_date_of_cycle | Date of the first drug administration in each Cycle. | DATE | ddmmyyyy |
| Person Weight | weight_at_start_of_cycle | Weight in kilograms at the start of the Cycle. Note that the completion of this variable is optional as required for local purposes. | NUMBER | Range: 1 - 985 Recorded to 3 d.p. |
| Performance Status | perf_status_start_of_cycle | Performance status at start of cycle | TEXT | See previous performance status variable |
| Primary Procedure | opcs_procurement_code | OPCS code which is used to identify the primary procedure carried out | TEXT | Max length 4 characters |

4. Drug detail

| <i>Column description</i> | <i>Column Name</i> | <i>Details</i> | <i>Field Type</i> | <i>Valid Content</i> |
|-------------------------------------------------------------|--------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|----------------------------------------------------------------------------------------------------------------------|
| CR patient Identifier | e_cr_patid | Unique patient identifier based on NCRAS data. In some cases, the same person may have multiple patient IDs. Patient IDs will be retained even after two patient records are found to be the same person. | ID | Number |
| Pseudonymised Tumour Identifier | e_pseudo_merged_tumour_id | Tumour Identifier | NUMBER | Max length 6 characters |
| Pseudonymised Regimen identifier | e_pseudo_merged_regimen_id | Regimen Identifier | NUMBER | Max length 6 characters |
| Pseudonymised Systemic Anti-Cancer Therapy Cycle Identifier | e_pseudo_merged_cycle_id | Cycle Identifier is a unique identifier for an Anti-Cancer Drug Cycle within an Anti-Cancer Drug Regimen. | NUMBER | Max length 7 characters |
| Drug Detail identifier | e_pseudo_merged_drug_detail_id | Drug Detail Identifier | NUMBER | Max length 7 characters |
| Drug Type Grouping | drug_group | This is a cleaned version of the drug name. Note includes 'Not Chemo', 'Not matched'. | TEXT | Max Length 30 characters |
| Actual Dose | actual_dose_per_administration | The actual chemotherapy dose can be given in milligrams (mg) or other applicable units, for example, 400 milligrams, 200 units, 1.5 grams etc. | NUMBER | Range: Min -1, Max 6575 |
| SACT drug route Of Administration | administration_route | Route of administration for SACT drug | NUMBER | 01 =Intravenous 02 = Oral 03 = Intrathecal 04 = Intramuscular 05 = Subcutaneous 06 = Intraarterial |

| | | | | |
|---------------------------------|------------------------------------|--------------------------------------------------------------------------------------------------------------------------------------------------|--------|--------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| | | | | 07 = Intraperitoneal 08 = Other intracavity Intracavernous 09 = Intravesical (Intra-Vesicular) 10 = Intratumour / Intralesional 11 = Cutaneous (Topical) 12 = Intradermal 13 = Intratumour 14 = Intralesional 98 = Other |
| SACT Administration Date | administration_date | Date on which the anti-cancer drug was administered to a patient, an infusion commenced, or an oral drug was initially dispensed to the patient. | DATE | ddmmyyyy |
| Pseudonymised organisation code | Organisation_code_of_drug_provider | Pseudonymised version of Organisation Code of Drug Provider | NUMBER | |
| Primary procedure (OPCS) | opcs_delivery_code | OPCS code for delivery of Systemic Anti-Cancer Therapy drug. SACT Version 2 only | TEXT | Free text |

5. Outcome

| <i>Column description</i> | <i>Column Name</i> | <i>Details</i> | <i>Field Type</i> | <i>Valid Content</i> |
|-------------------------------------------------|-------------------------------------|-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|-------------------|-------------------------|
| CR patient Identifier | e_cr_patid | Unique patient identifier based on NCRAS data. In some cases, the same person may have multiple patient IDs. Patient IDs will be retained even after two patient records are found to be the same person. | ID | Number |
| Pseudonymised Tumour Identifier | e_pseudo_merged_tumour_id | Tumour Identifier | NUMBER | Max length 6 characters |
| Pseudonymised Regimen identifier | e_pseudo_merged_regimen_id | Regimen Identifier | NUMBER | Max length 6 characters |
| Pseudonymised Outcome Identifier | e_pseudo_merged_outcome_id | Outcome Identifier | NUMBER | Max length 6 characters |
| Start Date (Final Therapy) | date_of_final_treatment | SACT Version 2 only Start date of the final cycle of chemotherapy within a Regimen, which is the end date of the chemotherapy treatment | DATE | ddmmyyyy |
| Regimen modification indicator (dose reduction) | regimen_modification_dose_reduction | Identifies if a regimen was modified by reducing the dose of any anti-cancer drug administered at any point in the regimen after commencement of the regimen. | TEXT | Y = Yes, N = No |
| Regimen modification indicator (Time Delay) | regimen_modification_time_delay | SACT Version 2 only. An indication of whether a Regimen was modified by reducing the administration days below the number planned. | TEXT | Y = Yes, N = No |
| Regimen modification indicator (days reduced) | regimen_mod_stopped_early | SACT Version 2 only. An indication of whether a Regimen was modified by reducing the administration days below the number planned. | TEXT | Y = Yes, N = No |

| | | | | |
|---------------------------------|-------------------------|------------------------------------------------------------------------------------------|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|
| Planned treatment change reason | regimen_outcome_summary | SACT Version 2 only. Whether the treatment was completed as planned, and if not, why. | TEXT | 1= Patient died, 2 = Progressive disease during chemotherapy, 3 = Acute chemotherapy toxicity, 4 = Technical or organisational problems, 5 = Patient choice (stopped or interrupted treatment) |
|---------------------------------|-------------------------|------------------------------------------------------------------------------------------|------|------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------|