



Medicines & Healthcare products
Regulatory Agency



CPRD

Systemic Anti-Cancer Therapy (SACT) Data Dictionary

Version 4.0

Date: 06 December 2019

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

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

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 chemotherapy	NUMBER	Number
Consultant Specialty Code	consultant_speciality_code	Main specialty code of the consultant who initiated the chemotherapy	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 (ICD-10)	TEXT	Max Length 6 characters
Tumour Morphology	morphology	Morphology at the initiation of SACT (ICD-O)	TEXT	Max Length 5 characters
Cleaned Morphology	morphology_clean	Cleaned version of Tumour Morphology	TEXT	Max Length 5 characters
Stage of Tumour at Start	stage_at_start	Pre-treatment TNM stage.	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	Programme Number	NUMBER	Max length 2 characters, Range -1 to 99
Regimen number	regimen_number	Number of the regimen within the programme. Added chronologically	NUMBER	Max length 2 characters, Range -1 to 99
Treatment Intent	intent_of_treatment	Intent of treatment with SACT	TEXT	A = Adjuvant, N = Neoadjuvant, C = Curative, P = Palliative, D = Disease Modification
Regimen Analysis Group	analysis_group	This is a cleaned version of variables mapping regimen information into consistent description groups	TEXT	Max Length 59 characters
Regimen Benchmark Group	benchmark_group	This is a grouped version of the raw description of the regimen	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 start of Regimen	weight_at_start_of_regimen	Weight in kilograms at the start of the SACT Regiment	NUMBER	Range -1 to 999 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	Adult: 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. Young person: Lansky Play - Performance Scale (00 to 11)
Comorbidity Adjustment Indicator	comorbidity_adjustment	Indication of whether a patient's overall physical state was a significant factor in deciding on the type of anti-cancer drug regimen, or influencing whether a dose was reduced, or the time between cycles increased at the commencement of the regimen.	TEXT	Y = Yes, N = No
Decision to Treat Date	date_decision_to_treat	The first date of a definitive decision to treat with chemotherapy or hormone therapy	DATE	ddmmyyyy
Start Date of Regimen	start_date_of_regimen	Planned start date of the chemotherapy or hormone therapy regimen.	DATE	ddmmyyyy
Clinical Trial	clinical_trial	Indicates whether an individual episode of care was delivered as part of a clinical trial.	TEXT	Max length 2 characters. Y = Yes, N = No, NK = Not Known, Null = Not recorded
Chemo-Radiation Indicator	chemo_radiation	Indicates whether regimen was given as part of a combined treatment with	TEXT	Max length 2 characters. Y = Yes, N = No, NK = Not Known, Null = Not recorded

		radiation		
Number of Planned Cycles	number_of_cycles_planned	Number of planned cycles of SACT within regimen	NUMBER	Max length 3 characters, Range -1 to 617

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
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. Recorded to 3d.p.
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 = Intra-vesicular intravesical, 10 = Intratumour intralesional, 11 = Topical cutaneous, 12 = Intradermal, 99 = Other
Administration Date	administration_date	Date on which the chemotherapy was administered, an infusion commenced, or an oral drug was initially dispensed.	DATE	ddmmyyyy
OPCS Delivery Code	opcs_delivery_code	OPCS code for delivery of Systemic Anti-Cancer Therapy drug.	TEXT	Free text Max length 4

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 Systemic Anti-Cancer Therapy)	date_of_final_treatment	Start date of the final cycle of chemotherapy within a Regimen, which is the end date of the chemotherapy treatment	DATE	ddmmyyyy
Modified Regimen by Dose Reduction	regimen_mod_dose_reduction	An indication of whether a Regimen was modified by reducing the dose administered.	TEXT	Y = Yes, N = No
Modified Regimen by Time Delay	regimen_mod_time_delay	An indication of whether a Regimen was modified by extending the time between administration dates.	TEXT	Y = Yes, N = No
Modified Regimen by Stopping Early	regimen_mod_stopped_early	An indication of whether a Regimen was modified by reducing the administration days below the number planned.	TEXT	Y = Yes, N = No

Outcome Summary	regimen_outcome_summary	Whether the treatment was completed as planned, and if not, why.	TEXT	0 = Treatment completed as prescribed, 1= Patient died, 2 = Progressive disease during chemotherapy, 3 = Acute chemotherapy toxicity, 4 = Technical or organisational problems, 5 = Patient choice (stopped or interrupted treatment)
-----------------	-------------------------	--	------	---