Systemic Anti-Cancer Therapy (SACT) Data Dictionary

Version 5.1

Date: 26 March 2021
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.

<table>
<thead>
<tr>
<th>Version</th>
<th>Summary of Change</th>
<th>Prepared By</th>
<th>Reviewed By</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>Initial draft</td>
<td>Helen Strongman</td>
<td>Rachael Williams</td>
</tr>
<tr>
<td>2.0</td>
<td>Modified</td>
<td>Rachael Williams</td>
<td>Eleanor Yelland</td>
</tr>
<tr>
<td>3.0</td>
<td>Modified</td>
<td>Helen Booth</td>
<td>Eleanor Yelland</td>
</tr>
<tr>
<td>4.0</td>
<td>Modified</td>
<td>Sonia Coton</td>
<td>Eleanor Yelland</td>
</tr>
<tr>
<td>5.0</td>
<td>Modified</td>
<td>Eleanor Yelland</td>
<td>Hilary Shepherd</td>
</tr>
<tr>
<td>5.1</td>
<td>Modified</td>
<td>Hilary Shepherd</td>
<td></td>
</tr>
</tbody>
</table>

Version 1.0
- First version (20 June 2017) for set 14
- No changes for set 15

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

Version 5.1
- Updated for set 21
- Updated branding
- Added DOIs
- Added v3 SACT variables: regoutsum_cur_not_com_plan, regoutsum_non_curat, regoutsum_toxic, regoutsum_cur_com_plan, adjunctive therapy
- Dropped v2 SACT variables: stage_at_start, programme_number, regimen_number, chemo_radiation, number_of_cycles_planned, opcs_procurement_code, regimen_modification_time_delay, regimen_modification_stopped_early, regimen_outcome_summary, opcs_delivery_code, date_of_final_treatment, organisation_code_of_provider

DOI
Please cite in any publications using these data:

CPRD GOLD SACT August 2021 - https://doi.org/10.48329/4jwp-4951

CPRD Aurum SACT August 2021 - https://doi.org/10.48329/nt7-ex36
### 1. Demographics and Consultant - One row per patient

<table>
<thead>
<tr>
<th>Column description</th>
<th>Column Name</th>
<th>Details</th>
<th>Field Type</th>
<th>Valid Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPRD patient identifier</td>
<td>e_patid</td>
<td>Unique patient identifier based on CPRD primary care data</td>
<td>ID</td>
<td>Number</td>
</tr>
<tr>
<td>CR patient Identifier</td>
<td>e_cr_patid</td>
<td>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.</td>
<td>ID</td>
<td>Number</td>
</tr>
<tr>
<td>Pseudonymised Tumour Identifier</td>
<td>e_cr_id</td>
<td>Tumour Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Consultant code</td>
<td>consultant_gmc_code</td>
<td>Pseudonymised code of the consultant who initiated the SACT treatment</td>
<td>NUMBER</td>
<td>Number</td>
</tr>
<tr>
<td>Consultant Specialty Code</td>
<td>consultant_speciality_code</td>
<td>Main specialty code of the consultant who initiated the SACT treatment</td>
<td>NUMBER</td>
<td>Max Length 8 characters, Range 0 - 991</td>
</tr>
</tbody>
</table>
2. Clinical Status

<table>
<thead>
<tr>
<th>Column description</th>
<th>Column Name</th>
<th>Details</th>
<th>Field Type</th>
<th>Valid Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPRD patient identifier</td>
<td>_patid</td>
<td>Unique patient identifier based on CPRD primary care data</td>
<td>ID</td>
<td>Number</td>
</tr>
<tr>
<td>CR patient Identifier</td>
<td>_cr_patid</td>
<td>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.</td>
<td>ID</td>
<td>Number</td>
</tr>
<tr>
<td>Primary Diagnosis</td>
<td>primary_diagnosis</td>
<td>Primary diagnosis at the initiation of SACT (at the time of decision to treat) (ICD-10)</td>
<td>TEXT</td>
<td>Max Length 6 characters</td>
</tr>
<tr>
<td>Morphology</td>
<td>morphology_clean</td>
<td>Morphology ICD-O at the initiation of SACT (at the time of decision to treat)</td>
<td>TEXT</td>
<td>Max Length 5 characters</td>
</tr>
<tr>
<td>Column Name</td>
<td>Details</td>
<td>Field Type</td>
<td>Valid Content</td>
<td></td>
</tr>
<tr>
<td>-------------------------------------</td>
<td>---------------------------------------------------------------------------------------------------</td>
<td>------------</td>
<td>-------------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>CR patient Identifier</td>
<td>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.</td>
<td>ID</td>
<td>Number</td>
<td></td>
</tr>
<tr>
<td>Pseudonymised Tumour Identifier</td>
<td>Tumour Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
<td></td>
</tr>
<tr>
<td>Pseudonymised Regimen identifier</td>
<td>Regimen Identifier (used to link to all subsequent tables)</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
<td></td>
</tr>
<tr>
<td>Drug treatment intent</td>
<td>Intent of SACT regimen</td>
<td>TEXT</td>
<td>01 = Curative - aiming to permanently eradicate disease</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>02 = Palliative - Aiming to extend life expectancy</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>03 = Palliative - Aiming to relieve and/or control malignancy related symptoms</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>04 = Palliative - Aiming to achieve remission</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>05 = Palliative - Aiming to delay tumour progression</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>98 = Other</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>99 = Not Known</td>
<td></td>
</tr>
<tr>
<td>Adjunctive therapy</td>
<td>Adjunctive therapy</td>
<td>TEXT</td>
<td>1= Adjuvant</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2= Neoadjuvant</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>3= Not Applicable (Primary Treatment)</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>4= Not Known</td>
<td></td>
</tr>
<tr>
<td>Regimen Analysis Group</td>
<td>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</td>
<td>TEXT</td>
<td>Max Length 59 characters</td>
<td></td>
</tr>
<tr>
<td>Regimen Benchmark Group</td>
<td>benchmark_group</td>
<td>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’.</td>
<td>TEXT</td>
<td>Max Length 59 characters</td>
</tr>
<tr>
<td>-------------------------</td>
<td>----------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>-------</td>
<td>------------------------</td>
</tr>
<tr>
<td>Person Height at start of Regimen</td>
<td>height_at_start_of_regimen</td>
<td>Height in metres at the start of the SACT Regimen</td>
<td>NUMBER</td>
<td>Range -1 to 9.99 Max 2 d.p</td>
</tr>
<tr>
<td>Person Weight at start of Regimen</td>
<td>weight_at_start_of_regimen</td>
<td>Weight in kilograms at the start of the SACT Regimen</td>
<td>NUMBER</td>
<td>Range -1 to 999 Max 3 d.p</td>
</tr>
<tr>
<td>Performance Status at Start of Regimen</td>
<td>perf_status_start_of_regimen_clean</td>
<td>Performance Status Indicator of Person at start of regimen (adult)</td>
<td>TEXT</td>
<td>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 Performance status (adult and young person)- SACT</td>
</tr>
<tr>
<td>Version 2</td>
<td>Abgle to carry out all normal activity without restriction</td>
<td>Abgle to carry out all normal activity without restriction</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>----------------------------------------------------------</td>
<td>----------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>Able to carry out all normal activity without restriction</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1</td>
<td>Restricted in physically strenuous activity, but able to walk and do light work</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>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</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Capable of only limited self care, confined to bed or chair more than 50% of waking hours</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Completely disabled. Cannot carry on any self care. Totally confined to bed or chair</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>00</td>
<td>100% - Fully active, normal</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>01</td>
<td>90% - Minor restrictions in physically strenuous activity</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>80% - Active, but tires more quickly</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>03</td>
<td>70% - Both greater restriction of, and less time spent in, play activities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>04</td>
<td>60% - Up and around, but minimal active play; keep busy with quieter activities</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>05</td>
<td>50% - Gets dressed but lies around much of</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Comorbidity Adjustment Indicator</td>
<td>comorbidity_adjustment</td>
<td>Indicator of whether or not patient's overall 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</td>
<td>TEXT</td>
<td>Y = Yes, N = No</td>
</tr>
<tr>
<td>---------------------------------</td>
<td>------------------------</td>
<td>-------------------------------------------------------------------------------------------------</td>
<td>------</td>
<td>----------------</td>
</tr>
<tr>
<td>Decision to Treat Date (Drug Regimen)</td>
<td>date_decision_to_treat</td>
<td>This is the date that the consultation between the patient and the clinician took place and a Planned Cancer Treatment was agreed</td>
<td>DATE</td>
<td>ddmmmm/yyyy</td>
</tr>
<tr>
<td>Month of decision to treat</td>
<td>month_of_decision_to_treat_mm</td>
<td>Month</td>
<td>DATE</td>
<td>mm</td>
</tr>
<tr>
<td>Year of decision to treat</td>
<td>year_of_decision_to_treat_yyyy</td>
<td>Year</td>
<td>DATE</td>
<td>yyyy</td>
</tr>
<tr>
<td>Start Date of Regimen</td>
<td>start_date_of_regimen</td>
<td>This is the first administration date of the first cycle of a regimen.</td>
<td>DATE</td>
<td>ddmmmm/yyyy</td>
</tr>
<tr>
<td>Month start for drug regimen</td>
<td>month_start_for_drug_regimen</td>
<td>month</td>
<td>DATE</td>
<td>mm</td>
</tr>
<tr>
<td>Year of start for drug regimen</td>
<td>year_start_for_drug_regimen</td>
<td>Year</td>
<td>DATE</td>
<td>yyyy</td>
</tr>
<tr>
<td>Clinical Trial</td>
<td>clinical_trial</td>
<td>Indicates whether an individual episode of care</td>
<td>TEXT</td>
<td>01 = PATIENT is taking</td>
</tr>
<tr>
<td>Indicator</td>
<td>was delivered as part of a clinical trial.</td>
<td>part in a CLINICAL TRIAL 02 = PATIENT is not taking part in a CLINICAL TRIAL 99 = Not Known</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------</td>
<td>------------------------------------------</td>
<td>-----------------------------------------------------------------------------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Column description</td>
<td>Column Name</td>
<td>Details</td>
<td>Field Type</td>
<td>Valid Content</td>
</tr>
<tr>
<td>-------------------------------------------</td>
<td>-------------------------------------------------</td>
<td>-----------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------------</td>
<td>------------</td>
<td>-----------------------------------</td>
</tr>
<tr>
<td>CR patient Identifier</td>
<td>e_cr_patid</td>
<td>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.</td>
<td>ID</td>
<td>Number</td>
</tr>
<tr>
<td>Pseudonymised Tumour Identifier</td>
<td>e_pseudo_merged_tumour_id</td>
<td>Tumour Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Pseudonymised Regimen identifier</td>
<td>e_pseudo_merged_regimen_id</td>
<td>Regimen Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Pseudonymised Cycle Identifier</td>
<td>e_pseudo_merged_cycle_id</td>
<td>Unique identifier for a Cycle within a Regimen (used to link to drug detail table).</td>
<td>NUMBER</td>
<td>Max length 7 characters</td>
</tr>
<tr>
<td>Cycle Identifier</td>
<td>cycle_number</td>
<td>Number of cycle</td>
<td>NUMBER</td>
<td></td>
</tr>
<tr>
<td>Start Date (Cycle)</td>
<td>start_date_of_cycle</td>
<td>Date of the first drug administration in each Cycle.</td>
<td>DATE</td>
<td>ddmmyyyy</td>
</tr>
<tr>
<td>Month of start date of cycle</td>
<td>month_of_start_date_of_cycle</td>
<td>Month</td>
<td>DATE</td>
<td>mm</td>
</tr>
<tr>
<td>Year of start date of cycle</td>
<td>year_of_start_date_of_cycle</td>
<td>Year</td>
<td>DATE</td>
<td>yyyy</td>
</tr>
<tr>
<td>Person Weight</td>
<td>weight_at_start_of_cycle</td>
<td>Weight in kilograms at the start of the Cycle. Note that the completion of this variable is optional as required for local purposes.</td>
<td>NUMBER</td>
<td>Range: 1 - 985 Recorded to 3 d.p.</td>
</tr>
<tr>
<td>Performance Status</td>
<td>performance_status_at_start_of_cycle_clean</td>
<td>Performance status at start of cycle (adult)</td>
<td>TEXT</td>
<td>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</td>
</tr>
</tbody>
</table>
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
<table>
<thead>
<tr>
<th>Column description</th>
<th>Column Name</th>
<th>Details</th>
<th>Field Type</th>
<th>Valid Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR patient Identifier</td>
<td>e_cr_patid</td>
<td>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.</td>
<td>ID</td>
<td>Number</td>
</tr>
<tr>
<td>Pseudonymised Tumour Identifier</td>
<td>e_pseudo_merged_tumour_id</td>
<td>Tumour Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Pseudonymised Regimen identifier</td>
<td>e_pseudo_merged_regimen_id</td>
<td>Regimen Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Pseudonymised Systemic Anti-Cancer Therapy Cycle Identifier</td>
<td>e_pseudo_merged_cycle_id</td>
<td>Cycle Identifier is a unique identifier for an Anti-Cancer Drug Cycle within an Anti-Cancer Drug Regimen.</td>
<td>NUMBER</td>
<td>Max length 7 characters</td>
</tr>
<tr>
<td>Drug Detail identifier</td>
<td>e_pseudo_merged_drug_detail_id</td>
<td>Drug Detail Identifier</td>
<td>NUMBER</td>
<td>Max length 7 characters</td>
</tr>
<tr>
<td>Drug Analysis Grouping</td>
<td>drug_group</td>
<td>This is a cleaned version of the drug name. Note includes 'Not Chemo', 'Not matched'.</td>
<td>TEXT</td>
<td>Max Length 30 characters</td>
</tr>
<tr>
<td>Actual Dose</td>
<td>actual_dose_per_administration</td>
<td>The actual chemotherapy dose can be given in milligrams (mg) or other applicable units, for example, 400 milligrams, 200 units, 1.5 grams etc.</td>
<td>NUMBER</td>
<td>Range: Min -1, Max 6575</td>
</tr>
<tr>
<td>The unit of measure used for each SACT drug administration</td>
<td>admin_measure_per_actual_dose</td>
<td>Units of measurements used for each administration in a SACT cycle. This descriptive field should be provided in combination with Actual_Dose_Per_Administration. Not available prior to 2018.</td>
<td>NUMBER</td>
<td>Max length 2 characters 01=mg 02=Mcg 03=g 04=Units 05=Cells 06=x10^6 PFU 07=x10^8 PFU 98=Other 99=Unknown</td>
</tr>
<tr>
<td>SACT drug route of Administration</td>
<td>administration_route</td>
<td>Route of administration for SACT drug</td>
<td>NUMBER</td>
<td>01 = Intravenous 02 = Oral</td>
</tr>
<tr>
<td>SACT Administration Date</td>
<td>administration_date</td>
<td>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.</td>
<td>DATE</td>
<td>ddmmyyyyy</td>
</tr>
<tr>
<td>--------------------------</td>
<td>----------------------</td>
<td>--------------------------------------------------------------------------------------------------------------------------------</td>
<td>------</td>
<td>---------</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

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
### 6. Outcome

<table>
<thead>
<tr>
<th>Column description</th>
<th>Column Name</th>
<th>Details</th>
<th>Field Type</th>
<th>Valid Content</th>
</tr>
</thead>
<tbody>
<tr>
<td>CR patient Identifier</td>
<td>e_cr_patid</td>
<td>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.</td>
<td>ID</td>
<td>Number</td>
</tr>
<tr>
<td>Pseudonymised Tumour Identifier</td>
<td>e_pseudo_merged_tumour_id</td>
<td>Tumour Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Pseudonymised Regimen identifier</td>
<td>e_pseudo_merged_regimen_id</td>
<td>Regimen Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Pseudonymised Outcome Identifier</td>
<td>e_pseudo_merged_outcome_id</td>
<td>Outcome Identifier</td>
<td>NUMBER</td>
<td>Max length 6 characters</td>
</tr>
<tr>
<td>Regimen modification indicator (dose reduction)</td>
<td>regimen_modification_dose_reduction</td>
<td>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.</td>
<td>TEXT</td>
<td>Y = Yes, N = No</td>
</tr>
<tr>
<td>Regimen outcome summary – curative (not completed as planned) reason</td>
<td>Regoutsum_cur_not_com_plan</td>
<td>Reason for not completing regimen as planned</td>
<td>TEXT</td>
<td>1 - Progressive/recurrent cancer  2 - Toxicity  3 - Death  4 - Patient choice  5 - Other</td>
</tr>
<tr>
<td>Regimen outcome summary – non-curative</td>
<td>Regoutsum_non_curat</td>
<td>An indicator of whether the patient benefited from the non-curative treatment</td>
<td>TEXT</td>
<td>Y/N</td>
</tr>
<tr>
<td>Regimen outcome summary - toxicity</td>
<td>Regoutsum_toxic</td>
<td>An indicator of the toxicity of the regimen</td>
<td>TEXT</td>
<td>Y/N</td>
</tr>
<tr>
<td>Regimen outcome summary – curative (completed as planned)</td>
<td>Regoutsum_cur_com_plan</td>
<td>An indicator of whether the patient completed the regimen as planned</td>
<td>TEXT</td>
<td>Y/N</td>
</tr>
</tbody>
</table>