Systemic Anti-Cancer Therapy Dataset

Implementation User Guide v0.11

Amendment History:

Version	Date	Amendment History
0.1	13/05/10	Draft Stage submission
0.2	14/12/10	Amendments following feedback
0.3	21/03/11	Updating of Data Dictionary definitions
0.4	13/04/11	Addition of technical guidance section
0.5	20/04/11	Further amendments and additions
0.6	03/05/11	Full Stage Submission
0.7	23/06/11	Amendments following Appraisal Meeting
0.8	24/06/11	Amendments following XML discussions
0.9	27/06/11	Amendments following communications re Data Dictionary
0.10	19/09/13	Changes to dataset to include additional data item and reference data
0.11	20/09/13	Updated following comments received following the ISB appraisal

This version of the guidance incorporates Data Dictionary changes as below.

Health and Social Care Information Centre

NHS Data Model and Dictionary Service

Reference: Change Request 1158
Version No: 1.0
Subject: Systemic Anti-Cancer Therapy Data Set
Effective Date: 1 April 2014
Reason for Change: Change to Data Standards
Publication Date: 5 November 2013

Forecast Changes:

Anticipated Change	When
NHS NUMBER STATUS INDICATOR CODE added to the dataset in order to comply with general information standard rules	01/04/2014
TNM CATEGORY (FINAL PRETREATMENT) changed name to TNM STAGE GROUPING (FINAL PRETREATMENT) to align with the Cancer Outcomes and Services Dataset	01/04/2014
Addition guidance of ‘this also corresponds to the term “line of chemotherapy” expressed in many prescribing systems’ added to Implementation User Guide	01/04/2014

Additional nation code of 'D – Disease modification" and corresponding note added to DRUG TREATMENT INTENT	01/04/2014

Reviewers:

This document must be reviewed by the following:

Name	Signature	Title / Responsibility	Date	Version
CIU Board	CIU Board	CIU Board	27/08/2013	0.10
CIU Board	CIU Board	CIU Board	20/09/2013	0.11

Approvals:

This document must be approved by the following:

Name	Signature	Title / Responsibility	Date	Version
ISB Board				

Distribution:

- ISB distribution list.
- ISB website.

Document Status:

This is a controlled document.

Whilst this document may be printed, the electronic version maintained on the ISB website (www.isb.nhs.uk) is the controlled copy. Any printed copies of the document are not controlled.

Contents

1. Introduction

The Systemic Anti-Cancer Therapy (SACT) Information Standard and phased implementation of national data collection applies to all organisations providing cancer chemotherapy services in, or funded by, the NHS in England. The standard relates to all cancer patients, both adult and paediatric, in acute inpatient, daycase, outpatient settings and delivery in the community. It covers chemotherapy treatment for all solid tumour and haematological malignancies, including those in clinical trials.

The impact of the standard will vary, depending on the configuration of hospitals and services and the existing and planned implementation of electronic prescribing and other clinical electronic systems.

The contents of this User Guidance document should be made available to all staff groups involved in responding to the standard i.e. medical and nursing, pharmacy, information, IT and management staff. It is not intended that introduction of the standard should have any direct impact on the delivery of patient care. However, the above groups, which are involved in the local implementation of the information standard, need to take account of implications of the standard in their work area and develop a strategy to fully meet its requirements by the end of the implementation period.

If you are a new Provider of chemotherapy, as well as reading this Implementation User Guide, please contact the Chemotherapy Intelligence Unit Helpdesk at ciu@sph.nhs.uk / ciu@phe.gov.uk. Other useful recourses to support the collection of the SACT dataset, such as Frequently Asked Questions, can be found on our website: http://www.chemodataset.nhs.uk/home.

1.1 Background

The national collection of all cancer chemotherapy information in the NHS in England commenced in April 2012. This is in line with the requirements of the Department of Health's policy document Improving Outcomes: A Strategy for Cancer January 2011.

Chemotherapy is now a major part of cancer treatment, with new types of drugs being introduced capable of targeting individual cancers. Historically the recording of chemotherapy has only been held within individual patients' notes. Despite the considerable costs of cancer chemotherapy, estimated to be in the order of one billion pounds a year, there has been no comprehensive picture available of the number of patients being treated or details of their care. With the advent of electronic recording of treatment, and in particular electronic prescribing systems, national collection and analysis of cancer chemotherapy being provided within the NHS is now viable. The SACT Information Standard addresses the requirement to standardise the recording of chemotherapy treatment and outcomes through electronic systems.

1.2 Benefits

From April 2012, a staged monthly data collection commenced, initially from trusts with e-prescribing systems, though all organisations delivering any chemotherapy for cancer were expected to provide some information from September 2012. Sufficient

data have now been quality assured and analysed to enable initial reports to be issued to contributing providers, data collection and reporting processes are now firmly established. This is however a continuing process and requires careful governance and maintenance.

This is an important new initiative with a wide range of benefits in terms of understanding patterns of clinical management in cancer chemotherapy. This is already recognised as being very valuable for those providing and commissioning chemotherapy services, ensuring that services are both of high quality and delivered efficiently. Equally importantly, it will support patients and their clinical teams in choosing appropriate care, based on accurate knowledge of current practice and the corresponding benefits and toxicities of treatment. This will, therefore, support patient choice and empowerment in a way that has not previously been possible.

The SACT dataset is also integrated with the other clinical NHS datasets, ultimately enabling the outcome of the complete patient pathway to be understood.

For details of the implementation timetable refer to Appendix 2.

1.3 Chemotherapy Intelligence Unit

The national collection of chemotherapy data is held and analysed by a Chemotherapy Intelligence Unit, based at the National Cancer Registration Service (NCRS) Oxford, and responsible to the National Cancer Intelligence Network (NCIN) within Public Health England. Section 251 of the NHS Act 2006 bounds the data which the NCRS receives.

In order to provide an accurate and complete analysis of clinical practice, the data collected includes information on the patient and their condition, with details of every attendance for chemotherapy. It also records a summary of the outcome of treatment.

1.4 Information Governance

The dataset contains sensitive and patient-identifiable information items. The NHS Health Research Authority has confirmed that reporting of patient identifiable data to the CIU is covered by the National Cancer Registration Service existing support under the Health Service (Control of Patient Information) Regulations 2002. Reported data will be managed by the CIU, which is part of the National Cancer Registration Service where there is expertise in managing large volumes of confidential data.

In compliance with the fair processing requirement within the Data Protection Act, provider organisations are expected to inform patients of this purpose for reporting their information and of the potential use of the information for service development, analysis and statistical research.

Where patients have requested that their data is not shared, the provider organisation must ensure that their records are not included in the data downloads submitted to the CIU. It is suggested that a ‘no consent’ or similar flag is provided in local systems so that the record can then be omitted from the monthly upload.

If a patient discovers that their information has been uploaded to the central repository and they wish for this to be deleted, the organisation must complete a

Subject Deletion Request form (available on the Chemotherapy Upload Portal) and send this to the CIU to action. The CIU will then delete the record from the database along with any backup files. An updated Patient Information Leaflet is currently under development which will explain that individuals have the right to access and have their own data held by the National Cancer Registration Service deleted, and the process by which to do this. The NCRS are currently in the process of drafting the new leaflet and are looking to consult with patient groups on its content in October 2013. A final version of the leaflet will be tested with focus groups and made available to stakeholders for comment prior to a final version being published in early 2014.

1.5 Clinical Governance

Analysis of the clinical content of the data collected will provide previously impossible insights into the patterns of cancer chemotherapy being delivered by individual providers and to individual patient groups and communities.

The format and content of reporting will be matched to the reasonable requirements of the various recipients of the data and reports, and the confidence intervals applying to each analysis made clear. When an apparently unacceptable variation in clinical practice is revealed by analysis a formal staged process of investigation must be undertaken. This process will determine the following:

- Is this an issue of variation within acceptable range but with limited patient choice?
- Is this an acceptable practice but worrying trend?
- Is this an issue which requires action within an agreed timescale?
- Is this an issue of immediate clinical concern?

This will decide the urgency of appropriate action which will be managed by the Chemotherapy Information Group.

1.6 Mapping local data to the SACT Information standard

There is no requirement to modify local clinical practices or data recording, however local system managers will be required to map local nomenclature and data formats to that defined in the SACT information standard before transmission. Provider organisations are encouraged to review the content of the standard and consider whether making primary data recording consistent with the standard would benefit their services in terms of safety and efficiency. Examples of this are standardisation of chemotherapy cycle numbering, particularly relevant where patient management is transferred during treatment and the consistent completion of fields summarising the end of treatment.

1.7 Maintenance and updating

Any changes required to improve the functionality and changes required from time to time to ensure that the data standard remains consistent with need, will be co-ordinated through the Chemotherapy Information Group. This group reports to the National Cancer Intelligence Network's Steering Group. Provider organisations are

encouraged to submit comments or requests concerning the dataset, its collection and analysis to CIU@phe.gov.uk for consideration.

Agreed changes or enhancements to the implementation of the data standard will be circulated to all contributors on a regular basis via the Chemotherapy Intelligence Unit.

2. Definitions for the National Systemic Anti-Cancer Therapy Data Set

With the advent of a National Systemic Anti-Cancer Therapy Data Set, it is important that field naming is consistent within hospital systems and the definitions of the fields are unambiguous and applied by all providers.

Where possible, field naming and definitions should either be aligned with those agreed for the Radiotherapy Dataset (ISB 0111), Cancer Outcomes and Services Dataset (ISB 1521) or avoided.

Definitions

The term “course” has not been used in the data set. The term is used variably and discussions have highlighted this as a potential risk. The term “regimen” plus the number of cycles has been substituted in the data set. The term programme has been added to mean the whole of a sequence of chemotherapy planned.

The relationships between programmes, regimens, cycles and administration dates are shown in the accompanying graphic and examples of dataset structures (pages 12-13).

Programme: The key factor in the definition of a programme is that it is a pre-planned sequence of treatment that may include one or more regimens. If the patient’s clinical situation changes, then subsequent treatment constitutes a new programme. (This is not applicable to dose reduction or time delay in administration)

- Where a curative programme is completed successfully but the patient subsequently develops recurrent disease, further treatment will constitute a new programme.

- Where a palliative treatment programme achieves the desired response but the patient subsequently relapses requiring further treatment this will constitute a new programme. *For example, a patient may receive four months of a taxane and is thought to have stable disease and the treatment is stopped. Two months later progressive disease is identified and the patient is started on Capecitabine, this constitutes a new programme.*

- Where a palliative treatment programme fails to achieve the desired response and is discontinued, with the formulation of a new treatment plan, further treatment will constitute a new programme. *For example, where a patient remains continuously on chemotherapy for a prolonged period, having a sequence of palliative regimens each in an attempt to control disease this would constitute a series of programmes, as it was not a planned sequence.*

In the management of the majority of adult solid tumours, the chemotherapy programme and regimen will be the same. Particularly in the management of haematological and paediatric tumours, two or more recognised regimens may be given concurrently or sequentially and constitute a single chemotherapy programme.

Programme number: Programmes will be numbered sequentially and the option to start from any number must be available to allow for prior management not recorded on the current system. This corresponds to the term "line of chemotherapy" expressed in many prescribing systems and constitutes a single chemotherapy programme.

Regimen: Conventionally this term is used to identify a standard or trial group of drugs given in a specific way and may include other instruction concerning the timing and parameters of treatment. The regimen title will be as agreed by the Oncology Regimen Steering Group and this will inform the OPCS Guidance for Clinical Coders.

In the management of the majority of adult solid tumours, the chemotherapy programme and regimen will be the same. Particularly in the management of haematological and paediatric tumours, two or more recognised regimens may be given concurrently or sequentially and constitute a single chemotherapy programme.

Regimen number: Where a patient has two or more regimens of chemotherapy within a programme, for a given cancer, they should be numbered sequentially, irrespective of intent. If two or more regimens commence on the same day, the regimen planned to be completed first should be given the lower number. The option to start from any number must be available to allow for prior management not recorded on the current system. If a patient develops a second cancer, the numbering will start again.

Cycle: Apart from continuous chemotherapy, a regimen normally contains identifiable repeating elements and each repeat should be identified and numbered. Some regimens have alternating repeating elements and some have consecutive sets of repeating elements. In all these cases the term “cycle” would be equally valid and help to identify the stage of progress of the patient through chemotherapy.

For continuous, normally oral chemotherapy, it will be necessary to agree an arbitrary equivalent. In order to align with the advice of the Oncology Regimen Steering Group, which informs the OPCS Guidance for Clinical Coders, a cycle will be 28 days from first administration.

Cycle number: These will be numbered sequentially within a regimen and the option to start from any number must be available to allow for prior management not recorded on the current system.

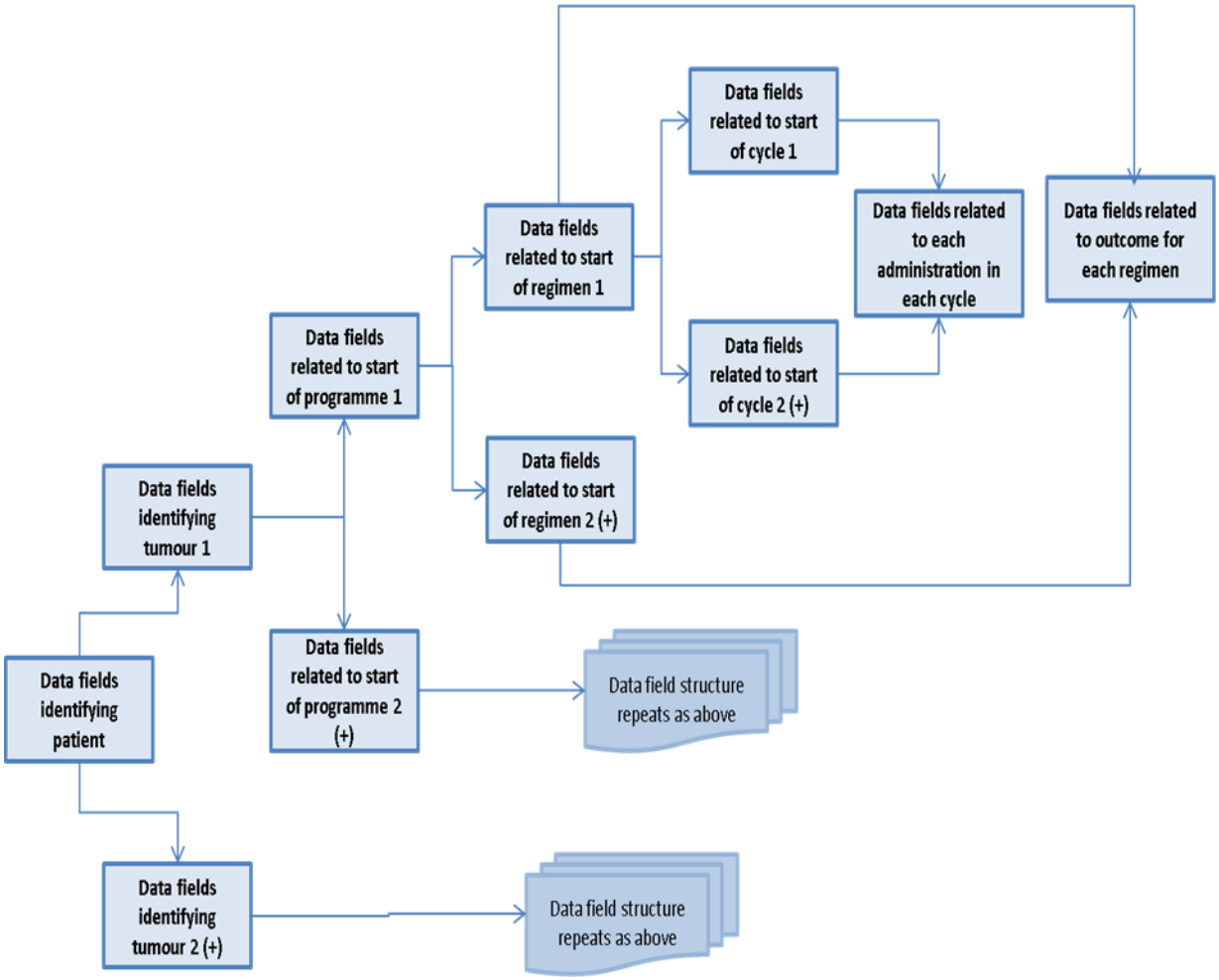
Administration date: Consistent terminology is required to identify each contact between the patient and the chemotherapy team when chemotherapy is administered. This will cover initial and subsequent contacts and needs to be recorded for inpatient treatment, chemotherapy clinic attendances, attendances in a primary care setting and domiciliary administration by a specialist service. In the case of infusions, the administration date will be the day the infusion was commenced.

For continuous oral chemotherapy, the administration date will be the first day of the nominal cycle i.e. one administration date per 28 days.

Date of final treatment: This is date of commencement of the final cycle (not the date of final administration).

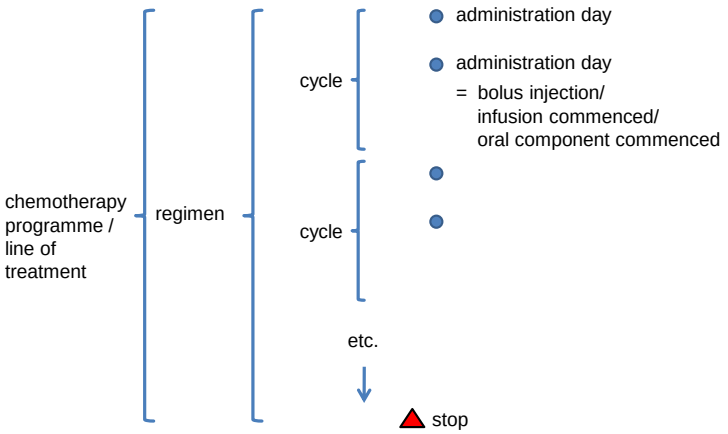
The dataset table is included as appendix 1.

2.1 SACT Data Model

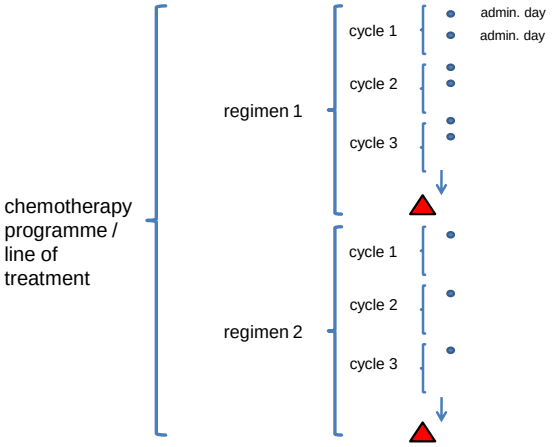


The Data Structures are described below.

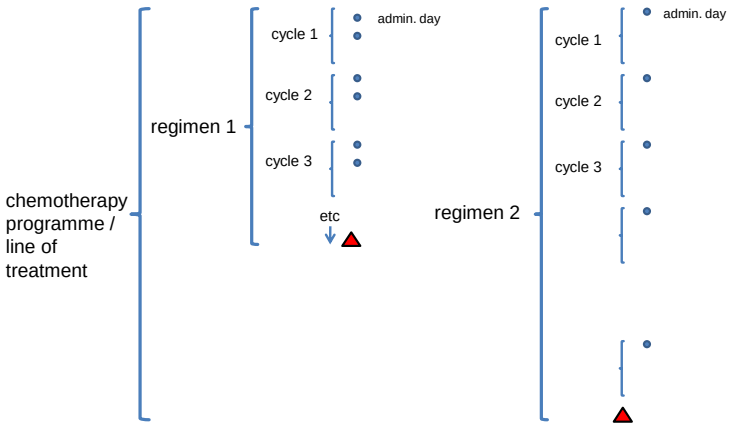
Dataset structure - example 1



Dataset structure - example 2



Dataset structure - example 3



3. Data Set Field Descriptions

Data item number and name
1. NHS number

Section
Demographics and consultant

SACT description
As NHS data dictionary

NHS data dictionary element
NHS NUMBER

The NHS NUMBER, the primary identifier of a PERSON, is a unique identifier for a PATIENT within the NHS in England and Wales. This will not vary by any ORGANISATION of which a PERSON is a PATIENT.

Format
n10

Relevant code and /or pick list
Not applicable

Schema specification
Mandatory

Purpose
Main identifier and essential for data linkage

Source
Hospital PAS

Comments
This is a fundamental field in the data set as the prime identifier.

The NHS NUMBER is 10 numeric digits in length. The tenth digit is a check digit used to confirm its validity. The check digit is validated using the Modulus 11 algorithm and the use of this algorithm is mandatory. There are 5 steps in the validation of the check digit:

Step 1 Multiply each of the first nine digits by a weighting factor as follows:

Digit Position
(starting from the left) Factor:

1	10

2	9
3	8
4	7
5	6
6	5
7	4
8	3
9	2

Step 2 Add the results of each multiplication together.

Step 3 Divide the total by 11 and establish the remainder.

Step 4 Subtract the remainder from 11 to give the check digit.

If the result is 11 then a check digit of 0 is used. If the result is 10 then the NHS NUMBER is invalid and not used.

Step 5 Check the remainder matches the check digit. If it does not, the NHS NUMBER is invalid.

Data item number and name
43. NHS number status indicator code

Section
Demographics and consultant

SACT description
NHS NUMBER STATUS INDICATOR CODE

NHS data dictionary element
NHS NUMBER STATUS INDICATOR CODE

Format
an2

Relevant code and /or pick list

01	Number present and verified
02	Number present but not traced
03	Trace required
04	Trace attempted - No match or multiple match found
05	Trace needs to be resolved - (NHS Number or patient detail conflict)
06	Trace in progress
07	Number not present and trace not required
08	Trace postponed (baby under six weeks old)

Schema specification
Mandatory

Purpose
The NHS NUMBER STATUS INDICATOR CODE indicates the verification status of the NHS number provided.

Source
Hospital PAS

Comments
No further comment

Data item number and name
2. Date of birth

Section
Demographics and consultant

SACT description
As NHS data dictionary

NHS data dictionary element
PERSON BIRTH DATE

The date on which a PERSON was born or is officially deemed to have been born.

Format
an10 ccyy-mm-dd

Relevant code and /or pick list
Not applicable

Schema specification
Mandatory

Purpose
This is additional identifier. It also allows analysis of provision by age

Source
Hospital PAS

Comments
This is a secondary identifier. It is generally well collected.

Data item number and name
3. Gender – current

Section
Demographics and consultant

SACT description
As NHS data dictionary

NHS data dictionary element
PERSON GENDER CODE CURRENT

A PERSON's gender currently

Format
an1

Relevant code and /or pick list
0 – not known
1 – male
2 – female
9 – not specified

Schema specification
Required

Purpose
To allow analysis by gender

Source
Hospital PAS system

Comments
Sex at birth would be a more fundamental data item but impractical to collect in some situations, therefore current gender has been included.

Data item number and name
4. Ethnicity

Section
Demographics and consultant

SACT description
As NHS data dictionary

NHS data dictionary element
ETHNIC CATEGORY

ETHNIC CATEGORY is the same as attribute ETHNIC CATEGORY CODE. The 16+1 ethnic data categories defined in the 2001 census is the national mandatory standard for the collection and analysis of ethnicity.

Format
an2

Relevant code and /or pick list
Office for National Statistic (ONS) 2001 categories 16+1

Schema specification
Required

Purpose
To allow analysis by ethnic category to reveal potential differences in uptake of treatment or types of treatment

Source
Hospital PAS

Comments
This field is not always well recorded and may be recorded differently by different sources. The incidence of some cancers may vary by ethnic group. This may be due to a combination of genetic, cultural and dietary factors.

Data item number and name
5. Patient postcode

Section
Demographics and consultant

SACT description
As NHS data dictionary

NHS data dictionary element
POSTCODE OF USUAL ADDRESS

The code allocated by the Post Office to identify a group of postal delivery points. A code used primarily for the delivery of correspondence to ADDRESSES. POSTCODES may also be used to define a GEOGRAPHIC AREA.

Format
max an8

Relevant code and /or pick list
Not applicable

Schema specification
Mandatory

Purpose
This is supportive identifier. It allows analysis by commissioner and geographical area, including generation of treatment rates by population. It allows demonstration of patient flows and provider catchments.

Source
Hospital PAS

Comments
This is an important field, since it is the only field that allows analysis by defined populations. The postcode may change during a patient's management either because the patient moves house or with changes in postcode allocation.

Data item number and name
6. Registered GP practice code

Section
Demographics and consultant

SACT description
As NHS data dictionary

NHS data dictionary element
The GENERAL MEDICAL PRACTICE CODE (PATIENT REGISTRATION) is an ORGANISATION CODE.
This is the CODE of the GP practice that the patient is registered with.

Format
an6

Relevant code and /or pick list
NHS list of code and name

Schema specification
Required

Purpose
To inform commissioning and allow analysis of patterns of care by commissioner

Source
Hospital PAS

Comments
GP practice code has been included as this is an established code which can be grouped for commissioning purposes.

Data item number and name
7. Consultant GMC code

Section
Demographics and consultant

SACT description
Code of consultant who initiated SACT programme

NHS data dictionary element
CONSULTANT CODE (INITIATED SYSTEMIC ANTI-CANCER THERAPY)

For the Systemic Anti-Cancer Therapy Data Set, this is the CONSULTANT CODE of the CONSULTANT who initiated the Systemic Anti-Cancer Therapy.

Format
an8

Relevant code and /or pick list
General Medical Council, unique number for each registered medical practitioner

Schema specification
Required

Purpose
It allows identification of consultant team responsible for initiating the programme and patterns of management provided.

Source
Hospital PAS or entered directly into prescribing system and code derived

Comments
In some specialty areas, several consultants may work as a team but an individual consultant must be identified as the consultant responsible for initiating the programme of chemotherapy.

Data item number and name
8. Consultant specialty code

Section
Demographics and consultant

SACT description
Specialty code of consultant who initiated SACT programme

NHS data dictionary element
CARE PROFESSIONAL MAIN SPECIALTY CODE

For the Systemic Anti-Cancer Therapy Data Set, this is the MAIN SPECIALTY CODE of the CONSULTANT who initiated the Systemic Anti-Cancer Therapy.

A unique code identifying each MAIN SPECIALTY designated by Royal Colleges. This is the same as the OCCUPATION CODES describing specialties. (Can be derived from consultant code).

Format
an3

Relevant code and /or pick list
HES item MAINSPEF

Schema specification
Required

Purpose
Identifies the specialty under which the patient is being managed.

Source
Organisation will derive from consultant code

Comments
This field can be derived from the consultant code but should be included as it provides an effective categorisation of clinical activity.

Data item number and name
9. Organisational code of provider

Section
Demographics and consultant

SACT description
As NHS data dictionary

NHS data dictionary element
ORANGISATION CODE (CODE OF PROVIDER)

See the "Organisation Default Codes" in the Default Codes Summary Table at http://www.datadictionary.nhs.uk/web_site_content/supporting_information/organisation_data_service_default_codes.asp?shownav=1.

Format
an3 or an5

Relevant code and /or pick list
NHS list of provider code and name. Additional lists will be required for non NHS providers, including home care delivery

Schema specification
Mandatory

Purpose
To allow analysis of care by provider and benchmarking between providers.

Source
Hospital PAS, other provider codes

Comments
This is a critical field in the data set as the provider of chemotherapy must be identified. This field shows the provider responsible for initiating the programme of chemotherapy.

Data item number and name
10. Primary diagnosis

Section
Clinical status

SACT description
Primary diagnosis at time of decision to treat

NHS data dictionary element
PRIMARY DIAGNOSIS (ICD AT START SYSTEMIC ANTI-CANCER THERAPY)

For the Systemic Anti-Cancer Therapy Data Set, this is the PRIMARY DIAGNOSIS at the start of the Systemic Anti-Cancer Therapy.

Format
an6

Relevant code and /or pick list
ICD-10

Schema specification
Mandatory

Purpose
To allow analysis by tumour site or group of tumour sites

Source
Several possible sources: PAS, prescribing system, MDT, linked pathology system

Comments
This field is essential for solid tumours as it defines the anatomical site of the primary tumour. Where a patient has more than one current cancer diagnosis the diagnosis recorded is the one for which treatment is being given.

Note: 10.Primary diagnosis and/or 11.Morphology can be submitted.

Data item number and name
11. Morphology

Section
Clinical status

SACT description
Morphology at time of decision to treat

NHS data dictionary element
MORPHOLOGY (ICD-O AT START SYSTEMIC ANTI-CANCER THERAPY)

This is the PATIENT DIAGNOSIS for the cell type of the malignant disease recorded as part of a Cancer Care Spell.

Format
min an5 – max an7

Relevant code and /or pick list
ICD-O3

Schema specification
Mandatory

Purpose
Identification of morphological subgroups of disease, not defined by ICD-10 e.g. varieties of lung cancer and haematological malignancies

Source
Several possible sources: prescribing system, MDT, linked pathology system

Comments
This field is more appropriate for haematological malignancy which is not primarily based on anatomical site. It also gives added information for some solid tumours e.g. lung and testis. Where a patient has more than one current cancer diagnosis the diagnosis recorded is the one for which treatment is being given.

Note: 11.Morphology and/or 10.Primary diagnosis can be submitted.

Data item number and name
12. TNM Stage Grouping (Final Pretreatment)

Section
Clinical status

SACT description
Stage of disease

NHS data dictionary element
TNM STAGE GROUPING (FINAL PRETREATMENT)
Record the overall clinical TNM stage grouping of the tumour, derived from each T, N and M component prior to treatment. This classification is based on all the evidence available to the clinician(s) with responsibility for assessing the patient and for the patient's treatment plan. Such evidence arises from physical examination, imaging, endoscopy, biopsy, surgical exploration and other relevant examinations.

Format
max an5

Relevant code and /or pick list
Site specific UICC (Union for International Cancer Control) coding is used

Schema specification
Required

Purpose
To allow analysis by stage of disease. Early stage disease will have better outcomes than more advanced disease.

Source
MDT

Comments
The stage to be recorded is the final pre-treatment stage as specified in the COSD dataset.

Data item number and name
13. SACT Programme number

Section
Programme and regimen

SACT Description
Programmes of chemotherapy are numbered according to their chronological order of commencement in the patient's disease management.

NHS data dictionary element
SYSTEMIC ANTI-CANCER THERAPY PROGRAMME NUMBER

The number of the Systemic Anti-Cancer Therapy Programme.

The SYSTEMIC ANTI-CANCER THERAPY PROGRAMME NUMBER is allocated locally.

Systemic Anti-Cancer Therapy Programmes are numbered according to their chronological order of commencement in the PATIENT's disease management. This corresponds to the term "line of chemotherapy" expressed in many prescribing systems.

Format
max n2

Relevant code and /or pick list
not applicable

Schema specification
Required

Purpose
To facilitate sequential analysis of patient care

Source
E-prescribing system, local recording, MDT

Comments
In the terminology of the SACT data standard, the programme is the pre-planned sequence of treatment which may include one or more regimens. Please refer to the definitions section. For example, if the patient's clinical situation changes e.g. from curative to palliative treatment, this would require the commencement of a new programme. Programmes will be numbered sequentially and the option to start from any number must be available to allow for prior management not recorded on the current system. If programme number is not available locally, it will be derived via an algorithm in the SACT data repository. This corresponds to the term "line of chemotherapy" expressed in many prescribing systems.

Data item number and name
14. Regimen number

Section
Programme and regimen

SACT description
Regimens are numbered according to their chronological order of commencement in the patient's treatment programme

NHS data dictionary element
ANTI-CANCER REGIMEN NUMBER

The number of the Anti-Cancer Drug Regimen, for example, Systemic Anti-Cancer Therapy Regimen.

Anti-Cancer Drug Regimens are numbered according to their chronological order of commencement in the treatment programme.

Format
max n2

Relevant code and /or pick list
not applicable

Schema specification
Required

Purpose
To facilitate sequential analysis of patient care

Source
E-prescribing system, local recording, MDT

Comments
Regimens will be numbered sequentially and the option to start from any number must be available to allow for prior management not recorded on the current system. If two regimens within a programme start concurrently, the one due to finish first should be given the lower number.

Data item number and name
15. Intent of treatment

Section
Programme and regimen

SACT description
Intent of SACT regimen

NHS data dictionary element
DRUG TREATMENT INTENT

A classification of the overall aim of the anti-cancer drug programme.

Format
an1

Relevant code and /or pick list
National codes as below

Schema specification
Required

Purpose
To allow analysis by treatment intent

Source
E-prescribing system, MDT

Comments
National Codes:
References:
Cancer Outcomes and Services Dataset
A Adjuvant
N Neo-adjuvant
C Curative
P Palliative
D Disease Modification - an anticipated clinical improvement of at least a year's duration

The list of options for intent was originally limited to the four options already included in the data dictionary. Developments in clinical practice in many speciality areas require the addition of an extra option - Disease modification (D). This is defined as "an anticipated clinical improvement of at least a year's duration". Many current treatment programmes are intended to control cancer, often for many years without the expectation of eradicating the disease. These situations were not covered adequately by the previous options of intent.

Data item number and name
16. Regimen

Section
Programme and regimen

SACT description
As NHS data dictionary

NHS data dictionary element
DRUG REGIMEN ACRONYM

The acronym derived from the drugs used in the Anti-Cancer Drug Regimen used to identify the drugs used in the regimen

Format
max an35

NOTE: Non-alphanumeric characters dash – and round brackets () are allowed as these may exist in regimen names. This field is not case-sensitive.

Relevant code and /or pick list
OPCS Classification of Interventions and Procedures version 4.6 Regimen Name (Dataset short version). NOTE: The local acronym may be submitted only where the regimen is currently not included in the OPCS classification.

Schema specification
Mandatory

Purpose
To allow analysis by individual regimen or drug

Source
E-prescribing system or local records

Comments
It is expected that there will be an annual update to the OPCS classification main list with 2 or 3 supplementary lists as required during the year.

Data item number and name
17. Height at start of regimen

Section
Programme and regimen

SACT description
Height in metres at start of SACT regimen

NHS data dictionary element
PERSON HEIGHT IN METRES

A PERSON'S height in metres

Format
n1.max n2

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
To confirm appropriate dose of chemotherapy and dose by metre^2

Source
E-prescribing system

Comments
This field is applicable where a drug dose is being calculated on the basis of a patient's height and weight.

Data item number and name
18. Weight at start of regimen

Section
Programme and regimen

SACT description
Weight in kilogrammes at start of SACT regimen

NHS data dictionary element
PERSON WEIGHT

A PERSON'S weight in kilogrammes

Format
max n3.max n3

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
To confirm appropriate dose of chemotherapy and dose by metre^2

Source
E- prescribing system

Comments
This field is applicable where a drug dose is being calculated on the basis of a patient's height and weight.

Data item number and name
19. Performance status at start of regimen

Section
Programme and regimen

SACT description
A person's status relating to activity / disability at start of SACT regimen

NHS data dictionary element
PERFORMANCE STATUS FOR ADULTS
PERFORMANCE STATUS CODE FOR YOUNG PERSON

A World Health Organisation classification indicating a PERSON's status relating to activity / disability

The Lansky Play - Performance Scale indicating a young PERSON's status relating to activity / disability. This scale is used for young PERSONS aged 16 years and under.

Format
an1 or an2

Relevant code and /or pick list
WHO codes for adults 0-4

Lansky for children codes 00-11:

00	100% = Fully active, normal
01	90% = Minor restrictions in physically strenuous activity
02	80% = Active, but tires more quickly
03	70% = Both greater restriction of, and less time spent in, play activities
04	60% = Up and around, but minimal active play; keeps busy with quieter activities
05	50% = Gets dressed but lies around much of the day; no active play; able to participate in all quiet play and activities
06	40% = Mostly in bed; participates in quiet activities
07	30% = In bed; needs assistance even for quiet play
08	20% = Often sleeping; play entirely limited to very passive activities
09	10% = No play; does not get out of bed
10	5% = Unresponsive
11	0% = Dead

Schema specification
Required

Purpose

To allow for casemix adjusted analysis. Patients with poor performance status are less likely to tolerate or complete rigorous treatment.

Source
MDT

Comments
WHO categories 1-4 are a match to the Eastern Cooperative Oncology Group (ECOG) categories and should be used for adults (above 16 years). For birth to16 years the Lansky scale should be used.

Data item number and name
20. Co-morbidity adjustment

Section
Programme and regimen

SACT description
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 from the start of treatment

NHS data dictionary element
CO-MORBIDITY ADJUSTMENT INDICATOR

An indication of whether a PATIENT's overall physical state (i.e. other diseases and conditions) was a significant factor in deciding on the type, dose or scheduling of Anti-Cancer Drug Regimen, for example a Systemic Anti-Cancer Therapy Regimen.

Format
an1

Relevant code and /or pick list
Y/N

Schema specification
Required

Purpose
To allow for casemix adjusted analysis. Patients with co-morbidity are less likely to tolerate or complete rigorous treatment.

Source
MDT

Comments
This differs from the ACE 27 co-morbidity data item in the Cancer Outcomes and Services Dataset. The field in the SACT data set records whether the chemotherapy treatment chosen has been modified because of the patient's overall clinical condition. This includes treatment with an alternative regimen, or varying the dose or treatment interval from the start of treatment.

Data item number and name
21. Date decision to treat

Section
Programme and regimen

SACT description
As NHS data dictionary

NHS data dictionary element
DECISION TO TREAT DATE (ANTI-CANCER DRUG REGIMEN)

The date on which it was decided that the PATIENT required a specific Planned Cancer Treatment.

This is the date that the consultation between the PATIENT and the clinician took place and a Planned Cancer Treatment was agreed.

Format
an10 ccyy-mm-dd

Relevant code and /or pick list
not applicable

Schema specification
Required

Purpose
To allow analysis of wait before start of treatment

Source
Cancer Waiting Times

Comments
The Cancer Waiting Times dataset requires the decision date of every treatment to be recorded. This information may therefore be taken from a generic software system used to record all Cancer Waiting Times information.

Data item number and name
22. Start date of regimen

Section
Programme and regimen

SACT description
This is the first administration date of the first cycle of a regimen

NHS data dictionary element
START DATE (ANTI-CANCER DRUG REGIMEN)

Format
an10 ccyy-mm-dd

Relevant code and /or pick list
not applicable

Schema specification
Mandatory

Purpose
To allow analysis by time period

Source
E-prescribing system

Comments
In practice this will be the same date as the start date of the first cycle in a regimen. It is the date of the first administration of chemotherapy. This information may therefore be taken from the e-prescribing system or may be recorded on a specific system used for chemotherapy treatments

Data item number and name
23. Clinical trial

Section
Programme and regimen

SACT description
As NHS data dictionary

NHS data dictionary element
CLINICAL TRIAL INDICATOR

For the SYSTEMIC ANTI-CANCER THERAPY PROGRAMME NUMBER, this identifies if a PATIENT's Chemotherapy treatment is within a CLINICAL TRIAL.

Format
an2

Relevant code and /or pick list
01 PATIENT is taking part in a CLINICAL TRIAL
02 PATIENT is not taking part in a CLINICAL TRIAL
99 Not known

Schema specification
Required

Purpose
To identify chemotherapy given within clinical trials

Source
E-prescribing system or local records

Comments
This field is simply to indicate whether a regimen is within a clinical trial which would not be clear otherwise, if it was the standard arm of the trial.

Data item number and name
24. Chemo-radiation

Section
Programme and regimen

SACT description
This field identifies regimens which are given as part of a combined treatment with radiation

NHS data dictionary element
CHEMO-RADIATION INDICATOR

An indication of whether a regimen, such as a Systemic Anti-Cancer Therapy Regimen, is given as part of a combined treatment with radiation.

Format
an1

Relevant code and /or pick list
Y/N

Schema specification
Required

Purpose
To identify use of chemo-radiation only used where this is a recognised treatment regimen

Source
E-prescribing system or local records

Comments
This field is used to record if a regimen is part of a recognised combined treatment, the radiotherapy and chemotherapy may be concurrent or sequential. The regimen name may indicate that it is a combined treatment.

Data item number and name
25. Number of cycles planned

Section
Programme and regimen

SACT description
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.

NHS data dictionary element
NUMBER OF SYSTEMIC ANTI-CANCER THERAPY CYCLES PLANNED

The number of Systemic Anti-Cancer Therapy Cycles specified in the CHEMOTHERAPY PRESCRIPTION.

This may be the number of Systemic Anti-Cancer Therapy Cycles in the standard Systemic Anti-Cancer Therapy Regimen or be modified by the prescriber.

Format
max n2

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
To allow comparison with number of cycles actually given.

Source
E-prescribing system or local records

Comments
Many regimens are prescribed with a stated number of cycles; this may be specified in a protocol but may be varied by the prescriber. Some prescriptions will not have a fixed number prescribed at the outset; this is particularly the case with some palliative treatments.

Data item number and name
26. Cycle number

Section
Cycle

SACT description
Cycles numbered sequentially within each regimen

NHS data dictionary element
ANTI-CANCER DRUG CYCLE IDENTIFIER

A unique identifier for an Anti-Cancer Drug Cycle within an Anti-Cancer Drug Regimen.

Anti-Cancer Drug Cycle is a CLINICAL INTERVENTION where the CLINICAL INTERVENTION TYPE is National Code 02 '*Anti-Cancer Drug Cycle*'.

Format
max n2

Relevant code and /or pick list
Not applicable

Schema specification
Mandatory

Purpose
Indicates a patient's progress through the regimen and to support analysis between years

Source
E-prescribing system

Comments
Cycles will be numbered sequentially within a regimen and the option to start from any number must be available to allow for prior management not recorded on the current system.

Data item number and name
27. Start date of cycle

Section
Cycle

SACT description
Date of first drug administration in each cycle

NHS data dictionary element
START DATE (SYSTEMIC ANTI-CANCER DRUG CYCLE)

The date of the first drug administration in each Systemic Anti-Cancer Therapy Cycle.

Format
an10 ccyy-mm-dd

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
To identify treatment patterns and to support analysis between years.

Source
E-prescribing and local records

Comments
No additional comment

Data item number and name
28. Weight at start of cycle

Section
Cycle

SACT description
A PERSON'S weight in kilogrammes at start of cycle

NHS data dictionary element
PERSON WEIGHT

A PERSON'S weight in kilogrammes

Format
max n3.max n3

Relevant code and /or pick list
Not applicable

Schema specification
Optional

Purpose
Where relevant to confirm appropriate dose of chemotherapy

Source
E- prescribing system

Comments
This is only relevant where weight change during a regimen triggers a change in drug dosage.

Data item number and name
29. Performance status at start of cycle

Section
Cycle

SACT description
A person's status relating to activity / disability at start of cycle

NHS data dictionary element
PERFORMANCE STATUS FOR ADULTS
PERFORMANCE STATUS CODE FOR YOUNG PERSON

A World Health Organisation classification indicating a PERSON's status relating to activity / disability

The Lansky Play - Performance Scale indicating a young PERSON's status relating to activity / disability. This scale is used for young PERSONS aged 16 years and under.

Format
an1 or an2

Relevant code and /or pick list
WHO codes for adults 0-4

Lansky for children codes 00-11:

00	100% = Fully active, normal
01	90% = Minor restrictions in physically strenuous activity
02	80% = Active, but tires more quickly
03	70% = Both greater restriction of, and less time spent in, play activities
04	60% = Up and around, but minimal active play; keeps busy with quieter activities
05	50% = Gets dressed but lies around much of the day; no active play; able to participate in all quiet play and activities
06	40% = Mostly in bed; participates in quiet activities
07	30% = In bed; needs assistance even for quiet play
08	20% = Often sleeping; play entirely limited to very passive activities
09	10% = No play; does not get out of bed
10	5% = Unresponsive
11	0% = Dead

Schema specification
Required

Purpose
To assess the patient's suitability for further treatment.

Source
E-prescribing system

Comments
WHO categories 1-4 are a match to the Eastern Cooperative Oncology Group (ECOG) categories and should be used for adults (above 16 years). For birth to16 years the Lansky scale should be used. This field is only relevant in some patients where the performance status changes during the chemotherapy treatment.

Data item number and name
30. OPCS procurement code

Section
Cycle

SACT description
As NHS data dictionary

NHS data dictionary element
PRIMARY PROCEDURE (OPCS)

OPCS-4 code of an OPERATIVE PROCEDURE

Format
an4

Relevant code and /or pick list
OPCS 4.6

Schema specification
Required

Purpose
To allow analysis by cost group

Source
Hospital PAS. Used to support Payment by Results (PbR).

Comments
Normally entered onto the hospital system by clinical coders.

Data item number and name
31. Drug name (this is repeated for each anti-cancer drug in the regimen)

Section
Drug details

SACT description
BNF or trial name

NHS data dictionary element
SYSTEMIC ANTI-CANCER DRUG NAME

The name of the Systemic Anti-Cancer Therapy drug given to a PATIENT during an Anti-Cancer Drug Regimen. The name is taken from British National Formulary chapter 8.

Format
max an35

Relevant code and /or pick list
British National Formulary (BNF), Virtual Therapeutic Moiety (VTM) list

Schema specification
Required

Purpose
To identify drug usage

Source
E-prescribing system or local record

Comments
This is the approved name in the British National Formulary (BNF). This is equivalent to the NHS Dictionary of Medicines and Devices Virtual Therapeutic Moiety (VTM) (SNOMED CT concept identifier). Drug names may be held as code within e-prescribing systems.

Data item number and name
32. Actual dose per administration

Section
Drug details

SACT description
Dose in mg or other applicable unit for each administration in a SACT cycle.

NHS data dictionary element
CHEMOTHERAPY ACTUAL DOSE

The actual Chemotherapy dose given in milligrams or other applicable unit for each administration in a Systemic Anti-Cancer Therapy Cycle.

Format
max n7

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
To allow cumulative analysis of drug use by patient and global analysis

Source
E-prescribing system

Comments
This will normally be in milligrams but a small number of drugs may be prescribed using other units. For oral regimens this is the total dose for a single day.

Data item number and name
33. SACT Administration route

Section
Drug details

SACT description
The prescribed method of delivery for each administration in a SACT cycle

NHS data dictionary element
SYSTEMIC ANTI-CANCER THERAPY DRUG ROUTE OF ADMINISTRATION

The prescribed method of delivery for each administration in a Systemic Anti-Cancer Therapy Cycle.

Format
an2

Relevant code and /or pick list
National codes should be used for this data but the SNOMED CT preferred term has been matched to this along with the corresponding SNOMED CT code to facilitate future change to SNOMED CT coding.

National Codes	Routes of administration	Definition	SNOMED CT Code
01	Intravenous	Injection of a medicinal product into a vein.	47625008
02	Oral	Taking a medicinal product by means of swallowing.	26643006
03	Intrathecal	Injection of a medicinal product through the dura to the subarachnoid cavity.	72607000
04	Intramuscular	Injection of a medicinal product into muscular tissue.	78421000
05	Subcutaneous	Injection of a medicinal product directly underneath the skin.	34206005
06	Intraarterial	Injection of a medicinal product into an artery.	58100008
07	Intraperitoneal	Injection of a medicinal product into the peritoneal cavity.	38239002
09	Intra-Vesicular Intravesical	Administration of a medicinal product to the urinary bladder.	372471009
10	Intratumour Intralesional	Administration by injection or any other means of a medicinal product directly to a lesion.	372466002
11	Topical Cutaneous	Administration of a medicinal product to the skin and/or cutaneous wounds and/or nails and/or hair in order to obtain a local effect.	6064005
12	Intradermal	Injection of a medicinal product into the dermis.	372464004

Purpose
To allow analysis by route of administration and identify critical areas e.g. intrathecal chemotherapy

Schema specification
Required

Source
E-prescribing system or local record

Comments
The list above is the list currently agreed by the Chemotherapy Information Group. The above definitions are from the NHS Dictionary of Medicines and Devices Virtual Therapeutic Moiety (VTM) (SNOMED CT concept identifier).

Data item number and name
34. Administration date

Section
Drug details

SACT Description
The date on which the anti-cancer drug was administered to a patient, an infusion commenced, or an oral drug initially dispensed to the patient

NHS data dictionary element
SYSTEMIC ANTI-CANCER THERAPY ADMINISTRATION DATE

The date on which the Systemic Anti-Cancer Therapy drug was administered to a PATIENT, an infusion commenced, or an oral drug was initially dispensed to the PATIENT.

Format
an10 ccyy-mm-dd

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
Defines the date of actual administration.

Source
E-prescribing system

Comments
No additional comment

Data item number and name
35. Organisational code of provider (for each administration)

Section
Drug details

SACT description
Code of provider for each administration in a SACT cycle

NHS data dictionary element
ORANGISATION CODE (CODE OF PROVIDER)

See the "Organisation Default Codes" in the Default Codes Summary Table at http://www.datadictionary.nhs.uk/web_site_content/supporting_information/organisation_data_service_default_codes.asp?shownav=1.

Format
an3 or an5

Relevant code and /or pick list
NHS list of provider code and name. Additional lists will be required for non NHS providers, including home care delivery

Schema specification
Required

Purpose
To allow analysis of care by provider and benchmarking between providers

Source
Hospital PAS, other provider codes

Comments
This is a critical field in the data set as the provider of chemotherapy must be identified. Patients may move between providers during their chemotherapy treatment.

Data item number and name
36. OPCS delivery code

Section
Drug details

SACT description
Delivery code for each administration

NHS data dictionary element
PRIMARY PROCEDURE (OPCS)

OPCS-4 code of an OPERATIVE PROCEDURE

Format
an4

Relevant code and /or pick list
OPCS 4.6

Schema specification
Required

Purpose
To allow analysis by cost group

Source
Hospital PAS. Used to support Payment by Results (PbR).

Comments
Normally entered onto the hospital system by clinical coders.

Data item number and name
37. Date of final treatment

Section
Outcome

SACT Description
The date of the start of the final cycle of SACT treatment within a regimen

NHS data dictionary element
START DATE (FINAL SYSTEMIC ANTI-CANCER THERAPY)

The Start Date of the final cycle of Systemic Anti-Cancer Therapy within a Systemic Anti-Cancer Therapy Regimen. This is defined as the End Date of the Systemic Anti-Cancer Therapy treatment.

Format
an10 ccyy-mm-dd

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
To register the completion or stopping of a regimen.

Source
E-prescribing system or local records

Comments
This has been made consistent with the definition in the National Confidential Enquiry into Patient Outcome and Death (NCEPOD) report 2008 www.ncepod.org.uk. It is the most practical date to record.

Data item number and name
38. Regimen modification – dose reduction

Section
Outcome

SACT Description
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

NHS data dictionary element
SYSTEMIC ANTI-CANCER THERAPY REGIMEN MODIFICATION INDICATOR (DOSE REDUCTION)

An indication of whether a Systemic Anti-Cancer Therapy Regimen was modified by reducing the dose administered.

Format
an1

Relevant code and /or pick list
Y/N

Schema specification
Required

Purpose
To allow a measurement of regimen toxicity

Source
E-prescribing system or local record

Comments
This field may also be generated automatically and is one of three fields recording changes in the regimen.

Data item number and name
39. Regimen modification – time delay

Section
Outcome

SACT Description
Identifies if a regimen was modified by extending the time between administration dates at any point in the regimen after commencement of the regimen.

NHS data dictionary element
SYSTEMIC ANTI-CANCER THERAPY REGIMEN MODIFICATION INDICATOR (TIME DELAY)

An indication of whether a Systemic Anti-Cancer Therapy Regimen was modified by extending the time between Systemic Anti-Cancer Therapy Administration Dates.
Note: Time delays of 5 days or fewer are discounted to allow for bank holidays or other incidental interruptions not related to drug tolerance.

Format
an1

Relevant code and /or pick list
Y/N

Schema specification
Required

Purpose
To allow a measurement of regimen toxicity

Source
E-prescribing system or local record

Comments
This field may also be generated automatically and is one of three fields recording changes in the regimen. Time delays in any cycle of 5 days or fewer should be discounted to allow for bank holidays or other incidental interruptions not related to drug tolerance.

Data item number and name
40. Regimen modification – stopped early

Section
Outcome

SACT Description
Identifies if a regimen was modified by reducing the administration days below the number planned.

NHS data dictionary element
SYSTEMIC ANTI-CANCER THERAPY REGIMEN MODIFICATION INDICATOR (DAYS REDUCED)

An indication of whether a Systemic Anti-Cancer Therapy Regimen was modified by reducing the administration days below the number planned.
Note: This is only applicable where a fixed number of cycles were specified at the start of treatment.

Format
an1

Relevant code and /or pick list
Y/N

Schema specification
Required

Purpose
To allow a measurement of regimen toxicity

Source
E-prescribing system or local record

Comments
This field is one of three fields recording changes in the regimen. It is only applicable where a fixed number of cycles were specified at the start of treatment.

Data item number and name
41. Regimen outcome summary

Section
Outcome

SACT description
To record the immediate outcome of the treatment

NHS data dictionary element
PLANNED TREATMENT CHANGE REASON

An indicator of whether the treatment within an Anti-Cancer Drug Programme was completed as planned, and if not, the reason why.

Format
an1

Relevant code and /or pick list
National Codes:
0 Treatment completed as prescribed

Treatment not completed
1 PATIENT died
2 Progressive disease during chemotherapy
3 Acute chemotherapy toxicity
4 Technical or organisational problems
5 PATIENT choice (stopped or interrupted treatment)

Schema specification
Required

Purpose
To allow outcome analysis

Source
E-prescribing system or local records

Comments
This is a fundamental field required by the National Chemotherapy Advisory Group (NCAG) Report August 2009 'Chemotherapy Services in England: Ensuring Quality and Safety' www.ncat.nhs.uk. Although this field is available in e-prescribing systems, there is frequently a failure to complete the field.

Data item number and name
42. Date of death

Section
Outcome

SACT description
As NHS data dictionary description

NHS data dictionary element
PERSON DEATH DATE

The date on which a PERSON died or is officially deemed to have died.

Format
an10 ccyy-mm-dd

Relevant code and /or pick list
Not applicable

Schema specification
Required

Purpose
To estimate 30-day mortality or analyse survival after chemotherapy

Source
Office for National Statistic (ONS)

Comments
This field will only be filled directly for the submission if a patient dies in hospital or the hospital is informed by the GP. For analysis purposes the CIU will draw on ONS data and match death data to the SACT by the Chemotherapy Intelligence Unit.

4. Technical Guidance for data extraction and submission

4.1 Use of XML for Chemotherapy Data Transmission

The NHS will move towards XML as a standard for data transmission in the future and the SACT data set and systems using it need to be able to make this transition. The Chemotherapy Intelligence Unit (CIU) is able to handle both CSV files and XML data in parallel from the outset but initially it is anticipated that the vast majority of returns will be in CSV format.

XML data can be imported easily into the existing database structure though cost and timescales for suppliers being able to output XML are likely to vary considerably. The system will therefore be able to handle both CSV and XML files in parallel for some time.

Timescales

XML schema tested between pilot Site (Addenbrookes) and CIU database	July 2013
XML schema published	August 2013
Suppliers to confirm SACT XML is available	March 2014:

4.2 Data extraction in CSV format

Data files are required to be submitted monthly, within 7 working weeks of the end of the calendar month, e.g. submissions of April 2013 chemotherapy data (01/04/2013 - 30/04/2013) to be uploaded to CIU by 15th June 2013. The timetable for monthly data submissions is around the 15th of each calendar month. The CIU provides an annual timetable for data submissions to all providers which contains exact dates, this is available on the website www.chemodataset.nhs.uk.

Data will be extracted from electronic prescribing and other electronic systems by system software suppliers working with local IT staff in constructing extraction routines.

The database import process requires files to be in a consistent format as outlined below:
Extracted data files should be a single Comma Separated Values (CSV) only, with a .csv file extension. A CSV file template will be available from the Chemotherapy Intelligence Unit (CIU) for data suppliers and software system developers. Note that CSV files must be of the windows type rather than Unix, with carriage returns at the end of each line as well as linefeeds

None of the data required is case sensitive.

CSV files should be saved with a text delimiter set to the double-quote character in order to allow the use of commas in data values.

The first row of the CSV file should consist of the Column Headers with the column names in exactly the format shown (i.e. including underscore characters). CSV files should not be compressed or packaged in any way.

CSV files should contain only, and all of, the following column headers in the following order, regardless of the data items that can be supplied. The mapping to data set items is shown by the Column Number.

Column Header	Column Number\data set Item number
NHS_number	1
Date_of_birth	2
Gender_current	3
Ethnicity	4
Patient_postcode	5
Registered_GP_Practice_Code	6
Consultant_GMC_code	7
Consultant_specialty_code	8
Organisation_code_of_provider	9
Primary_diagnosis	10
Morphology	11
Stage_of_disease	12
Programme_number	13
Regimen_number	14
Intent_of_treatment	15
Regimen	16
Height_at_start_of_regimen	17
Weight_at_start_of_regimen	18
Performance_status_at_start_of_regimen	19
Comorbidity_adjustment	20
Date_decision_to_treat	21
Start_date_of_regimen	22
Clinical_trial	23
Chemo_radiation	24
Number_of_cycles_planned	25
Cycle_number	26
Start_date_of_cycle	27
Weight_at_start_of_cycle	28
Performance_status_at_start_of_cycle	29
OPCS_procurement_code	30
Drug_name	31
Actual_dose_per_administration	32

Administration_route	33
Administration_date	34
Organisation_code_of_provider_administration	35
OPCS_delivery_code	36
Date_of_final_treatment	37
Regimen_modification_dose_reduction	38
Regimen_modification_time_delay	39
Regimen_modification_stopped_early	40
Regimen_outcome_summary	41
Date_of_death	42
NHS_number_status_indicator_code	43

4.3 File submission via the Chemotherapy Intelligence Unit (CIU) web portal

When a CSV file is ready for submission to the national database, staff at the treatment supplier will connect to the CIU chemotherapy web portal via a whole host of browsers, including Internet Explorer, Firefox, Safari, Chrome and others, as well as being able to access it on a Mac, PC, iPhone, iPad, tablets and even mobile phones.

The URL for the web portal is https://www.chemodataset.nhs.uk. The portal requires each registered user to agree to the site's terms and conditions. User logins are held within the repository database along with encrypted passwords for authentication.

Once users have logged in to the portal they will be presented with links to a choice of pages:

- Upload data page
- Validation and data quality reports pages
- User support pages and contact details

4.4 Data submission and file naming

The following file naming convention is to be used for submissions:

UnitID-yyyymmdd-yyyymmdd.CSV

Where UnitID is an agreed unique identifier for the supplying chemotherapy provider and matches the user login's unit code and yyyymmdd is the start date of the data (date of earliest treatment) followed by the end date (date of final treatment).

The name of the file will be created by the user on the web site during the submission process regardless of what the file is called locally on the treatment provider's computers.

Date picker controls will allow the user to select the date range of the file's data based on drug administration dates. The unit's unique identifier will be taken from the user's login credentials held in the database. The web portal will then display the

proposed filename for user approval preventing user errors in file naming. Upload instructions will be available on this page for users.

Files are transferred using the secure web based Hypertext Transfer Protocol Secure (HTTPS) / Secure Socket Layer (SSL) encrypted protocol, which is used on a daily basis for online shopping, online banking, etc. No extra action is required at the data suppliers end to establish this apart from being on an N3 network connection.

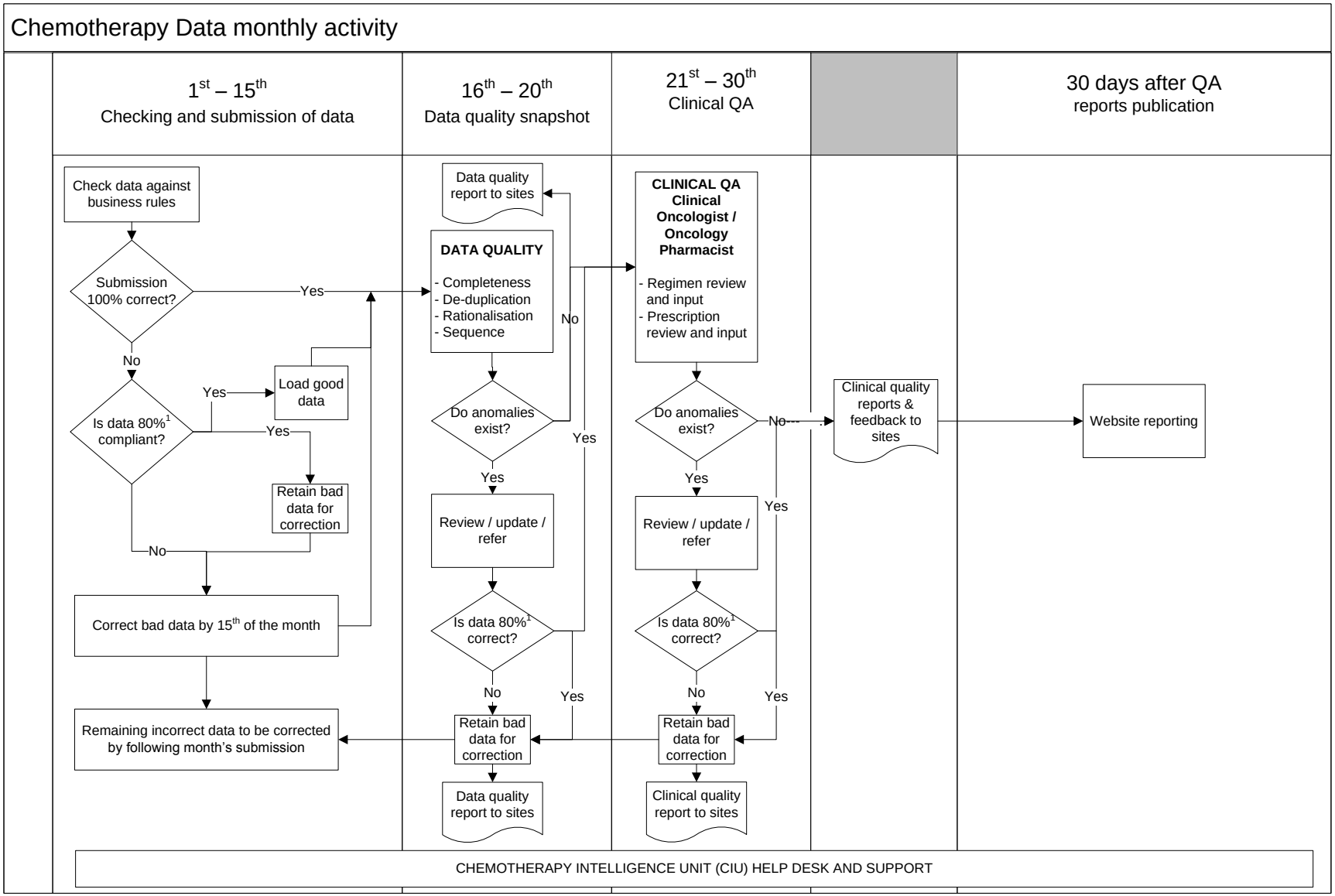
4.5 File Validation and Data Quality Reports

Files submitted are processed one at a time. The web portal should be able to provide validation results for any uploaded file within one hour of submission. This will be via a report generated for the data supplier on the web portal using logged validation data from the database.

File validation reports will be available for each file uploaded by the treatment supplier. Users will only be able to view reports related to their own data. Each report should display the following information for each file uploaded and processed:

Column name	Description of column
Filename	The name of the file that has been validated/uploaded
Uploaded By	The portal username who uploaded the file
Date range	The date and time of the extract.
Date uploaded	The date and time that the file was uploaded
Total Records	The count of non-blank records in the file
Valid	The count of successfully validated records in the file
Invalid	The count of invalid records in the file
Load%	The percentage of records loaded (i.e. valid records divided by number of records)
DQ%	The data quality percentage (i.e. number of records with no errors or warnings divided by number of records)
Error counts	Error counts by rule
Warning counts	Warning counts by rule
Informational errors	Informational error counts by rule
File Status	The current status of file – validated, rejected or loaded

Note that if the file fails validation of mandatory fields above a certain threshold it will be rejected and therefore not reach the data quality checks. Therefore the warning and informational counts will be empty.



[1] Recommended initial threshold, subject to review based on data quality issues encountered between April 2011 and April 2012. The threshold will increase over time.

Chemotherapy data monthly activity

Continuous help desk support will be provided by the Chemotherapy Intelligence Unit (CIU). As suppliers become more familiar with their monthly SACT data quality submission issues their success rate in having their monthly submission accepted first time (Submission 100% correct) will increase.

- **1st to 15th day of the month**
  - **Objectives –**
    - To test the data against the agreed business rules and ensure that mandatory fields are completed.
    - To ensure that data in all fields satisfies the required format and size.
    - Where requirements are satisfied to submit the data through the secure portal for data quality assurance.

  - **Testing data against business rules –**
    - Where the data tested is 100% correct it may be submitted through the secure portal.
    - Where at least 80%[1] of the data satisfies the business rules, the correct data may be submitted through the portal. The remaining data is retained for correction.
    - Where less than 80%[1] of the data satisfies the business rules, all of the data is retained for correction.
    - Sites should aim to correct bad data by the 15th day of the month to be able to submit data through the secure portal.
    - Where incorrect data remains, the data should be corrected in time for the following month's submission.

- **16th to 20th day of the month**
  - **Objectives –**
    - To ensure that the data submitted satisfies completeness and sequential requirements.
    - Where necessary rationalise and remove duplicate records.
    - To submit the qualified data for clinical quality assurance.

  - **Data quality process –**
    - Where the data which has been quality assured is 100% correct it may be submitted through for clinical QA and a data quality report sent to the site.

[1] Recommended initial threshold, subject to review based on data quality issues encountered between April 2011 and April 2012. The threshold will increase over time.

- Where anomalies exist, the data is to be reviewed and updated. This will include the correction of obvious errors, rationalisation of data and de-duplication.
- A data quality report will be returned to the site highlighting any changes.
- Where the data is at least 80%[1] correct, the correct data may be submitted for clinical QA. The remaining data is retained for correction by the site.
- Where less than 80%[1] of the data is correct, the data is retained for correction by the site.
- A data quality report is generated and sent to the site.
- The incorrect data is returned to the site for correction in time for the following month's submission.

- **21st to 30th day of the month**
  - **Objectives –**
    - The quality assurance of all submissions for the regimen and prescription against diagnosis.
    - A review of all data by Clinical Oncologist and Oncology Pharmacist.
    - Generation and provision of clinical QA reports to all sites.
    - Submission of data for analysis and report generation.

  - **Clinical quality assurance process –**
    - Where no anomalies exist the data is forwarded for analysis and the generation of the CIU reports.
    - Feedback reports on clinical quality generated and sent to sites
    - Where anomalies exist the data is reviewed by the Clinical Oncologist and Oncology Pharmacist through an iterative approach involving the treatment site where necessary.
    - If at least 80%[1] of the data is correct, the correct data may be submitted for analysis and report generation.
    - Incorrect data is retained for correction and referred back to the treatment site for correction in time for the following month's submission.

[1] Recommended initial threshold, subject to review based on data quality issues encountered between April 2011 and April 2012. The threshold will increase over time.

5. Supporting Information

Further information from:
https://www.chemodataset.nhs.uk

Help desk email:
ciu@sph.nhs.uk / ciu@phe.gov.uk

Appendix 1: Systemic Anti-Cancer Therapy Dataset

The Mandatory, Required or Optional (M/R/O) column indicates the recommendation for the inclusion of data.
M = Mandatory: this data element is mandatory; the message will be rejected if this data element is absent
R = Required: data is required as part of NHS business rules and must be included where available or applicable
O = Optional: the flow of this data is optional. It should be included at the discretion of the submitting organisation and their commissioners as required for local purposes.

In the case of fields 10 and 11, the requirement will be satisfied by one of the two fields being completed.

The SACT dataset V 2.0 is available on the ISB Website.

Appendix 2: Implementation timetable

<table>
<tr><th>Current situation</th><th>April 2013 – March 2014</th><th>October 2013</th><th>From April 2014</th><th>April - October 2014</th></tr>
<tr><td>Trusts with fully implemented e-prescribing systems</td><td>Continue full downloads</td><td rowspan="4">All software suppliers and NHS Trusts issued with ISN for modified SACT v2.0</td><td>Continue full downloads</td><td rowspan="4">Upgrades to all software systems with provider trusts implemented. Trusts submitting amended SACT data set</td></tr>
<tr><td>Trusts with partially implemented e-prescribing systems</td><td>Continue full downloads</td><td>Continue full downloads</td></tr>
<tr><td>Electronic clinical system but no e-prescribing</td><td>Continue partial downloads</td><td>Start full downloads</td></tr>
<tr><td>Basic hospital systems only</td><td>Continue partial downloads</td><td>Start full downloads</td></tr>
</table>