Public Health England

Protecting and improving the nation's health

New quality and completeness checks on submissions of Systemic Anti-Cancer Therapies (SACT) data - in place June 2017

The SACT dataset held by the National Cancer Registration and Analysis Service (NCRAS) at Public Health England has been collecting chemotherapy activity data since 2012. From April 2014, providing the 43 dataset items became mandatory.

In June 2017 we are improving how our data is handled and processed, which will allow us to carry out better checks on the quality and completeness of data being uploaded through the SACT portal.

This document provides details of those changes and the support available to you.

Upcoming changes

The changes being made are limited to checks on data quality and completeness in the SACT data portal - there are no changes to the total number of dataset items.

New completeness checks

Five more of the 43 dataset items will need to have a value present for a record to be successfully submitted through the SACT data portal (see table 1 & 3).

This will mean that for these dataset items, such as cycle start date, if the data item is not provided for a record then it can't be uploaded to the national SACT dataset.

New quality checks

For 20 more of the 43 dataset items, the values provided for each record will need to match the existing national standards for the record to be uploaded to the national SACT dataset (See table 2 & 3).

This will mean that for these dataset items, such as treatment intent, if a value is present and doesn't match those listed in the national standards then the record can't be uploaded to the national SACT dataset.

24/03/2017

Timings and impact on trust data submissions to the SACT portal

From June 1st records in your SACT submissions that don't comply with the changes will be flagged as an error by the SACT portal with an error message shown against those records. As is the case now, the message will say what is causing the error.

Records with errors should be corrected as those records will not be accepted into the national dataset. Once that's done the whole file needs to be resubmitted.

If over 20% of the records in the upload file have errors caused by missing or invalid data (tables 1-3) your trust will not be able to submit any data until a sufficient number of those errors are corrected.

As well as the validations against national standards there are other checks in the portal on formats and date sequences – these won't change and are listed in table 3.

Why is this happening now?

Since the SACT dataset was established in 2012, it has been maintained separately from other datasets held by the NCRAS. We are bringing the SACT dataset into the main NCRAS system, which allows us to carry out more effective checks on SACT data submissions. This is an important step to reduce variations in data quality and completeness between trusts, which will help ensure that publications that use the SACT dataset are reliable.

Those improvements are vital for upcoming work that will use the dataset, which includes follow on studies to the last year's 30-day mortality Lancet paper and evaluations of treatments in the Cancer Drugs Fund.

The changes will also help us to link SACT data to with other information provided by trusts. For example, the Cancer Outcomes and Services Dataset (COSD) that collects information from sources such as multidisciplinary team software and patient administration systems and pathology systems. Those links will provide immediate benefits. They'll also help us gain insights on new processes that will inform a longer term review of the SACT dataset. The review will ensure the dataset continues to include only the highest value information that can't come from any other sources.

Feedback and support

We are currently assessing the impact of these changes on individual trusts using their recent upload files.

Where the changes are likely to have a big impact on a trust's data submissions and workload we will proactively offer support and develop solutions that will make the

changes easier to manage. This will include arranging a visit from one of our data liaison team.

We'd welcome any feedback on how we can make these changes easier for you to manage at your trust and please do let us know if you have concerns about the impact they will have on you. If it might help we'll arrange a visit to your trust before June to help address any concerns. We can also support trusts that would like to test submitting files with the new checks in place

Please contact SACT@phe.gov.uk if you'd like to get in touch with us.

Details of the changes

1. Dataset items that must have a value present

To pass the checks in the SACT data portal and submit trust data to the national dataset, the following data items will need to have a value present on each record. Five of these checks on completeness are new (shown in table 1).

Table 1 – Data items that must have a value present for records in SACT uploads

Item No	Dataset Item	Values must be present?
1	NHS Number	Yes
2	Date of birth	Yes
3	Gender - current	Yes (new)
5	Patient postcode	Yes
9	Organisational code of provider	Yes
10	Primary diagnosis	Yes
16	Regimen	Yes
22	Start date of regimen	Yes
26	Cycle number	Yes
27	Start date of cycle	Yes (new)
31	Drug name	Yes (new)
34	Administration date	Yes (new)

Impact on upload

If records in your trusts' submission to the national dataset are missing values for one or more of the 13 items then an error message will appear when your trusts' data file is submitted through the SACT data portal.

The message will say; "Error message: Critical Error" along with more details.

2. Dataset items where values will be checked against national standards

To pass the checks in the SACT data portal and submit trust data to the national dataset, any values included for the following dataset items will need to match the national standards. Twenty of these validations are new (shown in table 2).

Table 2 - Dataset items where values will be checked against national standards

Item No	Dataset Item	National standard	Values validated?	Values must be present?
1	NHS Number	Valid	Yes	Yes
3	Gender - current	NHS data dictionary	Yes (new)	Yes (new)
4	Ethnicity	NHS data dictionary[1]	Yes (new)	
5	Patient postcode	Valid	Yes (new)	Yes
6	Registered GP practice code	GP registration	Yes (new)	
7	Consultant GMC code	GMC registration	Yes (new)	
9	Organisational code of provider	NHSE valid	Yes (new)	
10	Primary diagnosis	WHO ICD-10	Yes	
11	Morphology	ICD-O-3	Yes	
15	Intent of treatment	National Code	Yes (new)	
19	Performance status at start of regimen	WHO / Lansky	Yes (new)	
20	Co-morbidity adjustment	Y or N	Yes (new)	
23	Clinical trial	Y or N	Yes (new)	
24	Chemo-radiation	Y or N	Yes (new)	
29	Performance status at start of cycle	WHO / Lansky	Yes (new)	
30	OPCS procurement code	OPCS list	Yes (new)	
33	SACT Administration route	SnoMed	Yes (new)	
35	Organisational code of provider	NHSE valid	Yes (new)	
36	OPCS delivery code	OPCS list	Yes (new)	
38	Regimen modification - dose reduction	Y or N	Yes (new)	
39	Regimen modification - time delay	Y or N	Yes (new)	
40	Regimen modification - stopped early	Y or N	Yes (new)	
41	Regimen outcome summary	National Code	Yes (new)	

---

[1] Z codes will be accepted

Impact on upload

Where they are present in your trusts SACT data submission then the values for these dataset items must be provided in a national standard. Values that don't match the national standards will cause an error message to appear when your trusts' data file is submitted through the SACT data portal.

The message will say; "Error message: Validation error" along with more details.

National standards

The national standards are summarised in tables 2 and 3 and more detail can be found through the links below.

- All national standards can be found in the Systemic Anti-Cancer Therapy Dataset Implementation User Guide
- The NHS data dictionary standards can be found on the NHS data dictionary website
- Extra information is available from our SACT FAQ document.

Mapping non-standard values provided by some trusts

We know that some trusts provide values for certain data items that are consistent but don't match the national standards.

Some examples are;

- Yes and No instead of Y and N
- For gender M and F instead of 1 and 2

Wherever it is reasonable to do so we will build these variations into our systems rather than requiring trusts to make changes to the reports they are generating.

Inclusion of the OPCS codes in the validations

We are aware that there are a number of issues with OPCS procurement and delivery codes and an increasing numbers of trusts are no longer using them.

To manage this you can continue to submit records with OPCS dataset items left blank. Errors will only occur where records contain values and those values do not match the current OPCS list. This change is being made to ensure that if information is being provided then it is also in line with the national standard.

Inclusion of outcome dataset items in the validations

We know there are currently issues that mean outcome data completeness is low. Please supply valid data wherever possible and leave these items blank if those issues prevent you from doing so.

We will be looking into solutions for those issues and reviewing the items included.

New and existing quality and completeness checks in the SACT data portal

The table below summarises the checks carried out by the SACT portal on each of the 43 dataset items. As well as validations against national standards it includes other checks in the portal on data formats and date sequences e.g. the start date of a cycle must be on or before the administration date. These checks aren't changing and records with errors will still need to be corrected.

Table 3 – All data item quality and completeness checks on SACT data uploads

Dataset area	Item no.	Dataset item	National standard	Values validated?	Values must be present?	Other checks
Demographics and Consultant	1	NHS number	Valid	Yes	Yes	
	2	Date of birth			Yes	Date sequence / order
	3	Gender - current	NHS data dictionary	Yes (new)	Yes (new)	
	4	Ethnicity	NHS data dictionary[2]	Yes (new)		
	5	Patient postcode	Valid	Yes (new)	Yes	
	6	Registered GP practice code	GP registration	Yes (new)		
	7	Consultant GMC code	GMC registration	Yes (new)		
	8	Consultant speciality code				Format (number)
	9	Organisation code of provider	NHSE valid	Yes (new)	Yes	
Clinical Status	10	Primary diagnosis (ICD-10)	WHO ICD-10	Yes	Yes	
	11	Morphology	ICD-O-3	Yes		
	12	Stage of disease (Final Pre-treatment)				Valid staging system
Programme and Regimen	13	SACT Programme number				Format (number)
	14	Regimen number				Format (number)
	15	Intent of treatment	National Code	Yes (new)		
	16	Regimen name (OPCS 4.6)			Yes	Local or national name
	17	Height at start of regimen				Meters
	18	Weight at start of regimen				Kilogrammes
	19	Performance status at start of regimen	WHO / Lansky	Yes (new)		

[2] Z codes will be accepted

Dataset area	Item no.	Dataset item	National standard	Values validated?	Values must be present?	Other checks
	20	Co-morbidity adjustment	Y or N	Yes (new)		
	21	Date decision to treat				Date sequence / order
	22	Start date of regimen			Yes	Date sequence / order
	23	Clinical trial	Y or N	Yes (new)		
	24	Chemo-radiation	Y or N	Yes (new)		
	25	Number of cycles planned				Format (number)
Cycle	26	Cycle number			Yes	Format (number)
	27	Start date of cycle			Yes (new)	Date sequence / order
	28	Weight at start of cycle				Kilogrammes
	29	Performance status at start of cycle	WHO / lansky	Yes (new)		
	30	OPCS procurement code	OPCS National list	Yes (new)		
Drug Details	31	Drug name			Yes (new)	Local or National (BNF)
	32	Actual dose per administration				Format (number)
	33	SACT Administration route	SnoMed Code	Yes (new)		
	34	Administration date			Yes (new)	Date sequence / order
	35	Organisational code of provider (for each administration)	NHSE valid	Yes (new)		
	36	OPCS delivery code	OPCS National list	Yes (new)		
Outcome	37	Start date of final treatment				Date sequence / order
	38	Regimen modification - dose reduction	Y or N	Yes (new)		
	39	Regimen modification - time delay	Y or N	Yes (new)		
	40	Regimen modification - stopped early	Y or N	Yes (new)		
	41	Regimen outcome summary	National Code	Yes (new)		
	42	Date of death				Date sequence / order
	43	NHS number status Indicator code				National Code