National Casemix Office

NHS Digital

Grouper User Manual

HRG4+ 2016/17 Reference Costs Grouper

Published May 2017

RC 16/17

Information and technology
for better health and care

The Health and Social Care Information Centre is a non-departmental body created by statute, also known as NHS Digital

Contents

Purpose

The purpose of this document is to help users group patient activity and derive appropriate HRGs using the software application (the Groupers) provided by the National Casemix Office.

The Grouper will take input data from 7 data set types:

APC – Admitted Patient Care (In-patients)

NAC – Non-Admitted Consultations (Out-patients)

EM – Emergency Medicine (A&E activity)

ACC – Adult Critical Care

PCC – Paediatric Critical Care

NCC - Neonatal Critical Care

NRD – National Renal Dataset

Data can be loaded using a sample data format which gives the mandatory and required fields for grouping purposes, or users can create their own record definition files (RDFs) which similarly must contain the mandatory and required fields but will allow for flexibility of differential field ordering, including extra fields and bespoke field headers.

Non-mandatory fields can also be included using the sample data format if these are ordered after the mandatory/required fields.

Mandatory fields are those which must be populated with valid values, whereas required fields are those which must be present in the input data but can be left unpopulated (Blank).

Please note: The value NULL is not a valid value and must be removed prior to Grouping.

Quick Start

This section is intended to provide a brief overview of downloading, installing and starting to use the HRG4+ Grouper application. For more detailed information, you are advised to read the whole manual.

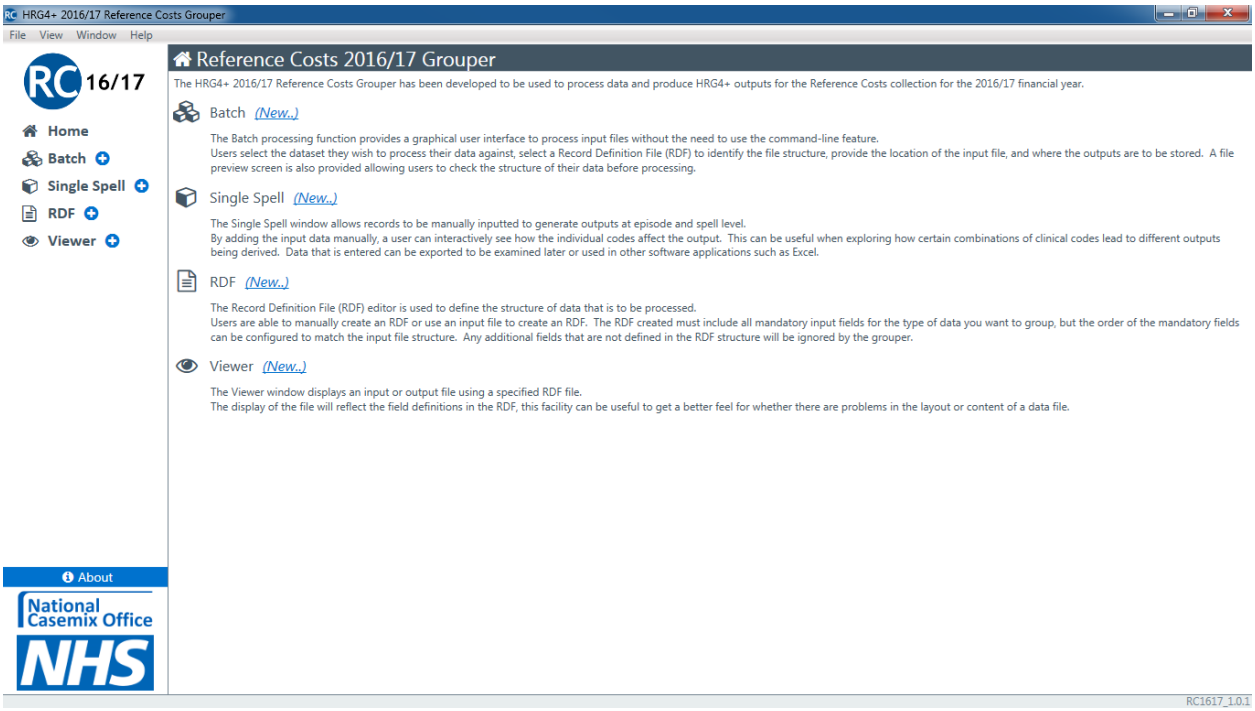
Download and Install the Grouper

- Go to http://content.digital.nhs.uk/casemix/downloads
- Click on Costing.
- There may be more than one version available. Older versions are on the archive pages.
- Click on the Grouper application link in the Download section at the right of the screen.
- Save the Zip file to an appropriate location on your computer.
- Open the Zip file and extract the installer.
- Double-click the installer file and follow the installation setup wizard.
- It is advised to install the Grouper to the default destination as provided in the installation set up wizard – **C:\Program Files\NHS Digital\HRG4+ 2016_17 Reference Costs Grouper**.
- If a user wishes to install the Grouper to a different destination to that provided in the installation set up wizard, then they can. It is strongly advised that the standard 21 Grouper application files including 5 folders (3 folders if Sample Data and Sample RDF folders are excluded), are extracted and maintained within a sub-directory.

Walkthrough Using Sample Data

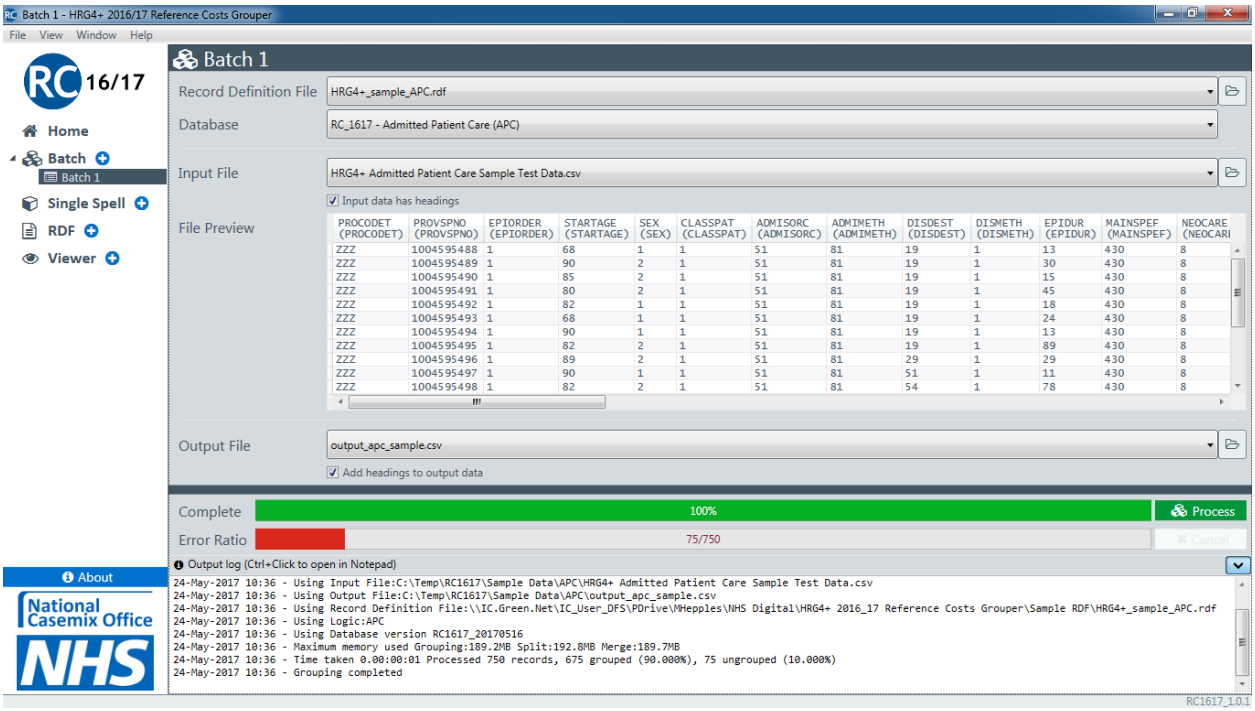
The following walkthrough describes processing a sample Admitted Patient Care (APC) input file, available from the Casemix website.

- Download the “HRG4+ 2016/17 Reference Costs Grouper Sample Data and Expected Results.zip” file from the same page as the Grouper.
- Extract the contents of the zip file to a suitable location.
- Open the Grouper application.



- In the navigation pane click the ⊕ icon next to **Batch** to open a new batch window.
- Click the 🗁 icon next to the **Record Definition File** drop-down list.
- In the **Open RDF File...** dialog box, browse for the sample RDF. The sample RDFs are located in the application’s installation folder in a sub-folder called `Sample RDF`.
- Open the file `HRG4+_sample_APC.rdf`. An alternative sample RDF file should be used for other database selections.
- Click the 🗁 icon to the right of the **Input File** field.
- In the **Open Input File...** dialog box, browse to and select the downloaded APC sample input file.
- Ensure that the **Input data has headings** box is checked.
- Click the 🗁 icon next to the **Output File** field.
- In the **Save Output File...** dialog box browse to the folder containing the input file and type `output_apc_sample` (or anything similarly appropriate) in the **File name** box.
- Click **Save**. The filename of the output file is displayed in the **Output File** field.

- Ensure that the **Add headings to output data** box is checked.
- Click the **Process** button.
- When finished, the log display area shows information about the grouping session, including the number of records grouped. These details can also be found in the file `hrg.log`. The log file can be opened from **Help** on the menu bar and selecting **View Log**.



- Open Windows Explorer and browse to the folder specified for the output file.
- Open the file `output_apc_sample_FCE.csv` in Microsoft Excel. The file is a copy of the input file with HRGs and other related columns appended. Other output files are explained in the Output Files section of this document.

Grouping Your Data

In order to group your own data you need to either:

- Ensure that the structure of your data is identical to that specified in an existing Record Definition File (RDF), or
- Create an RDF that matches the structure of your data.

Grouping can be initiated from the graphical user interface (GUI), or from the command line. Whichever approach is used the same input rules, grouper processing and output file structures and content will result.

Another approach for single records and for single admitted patient care (APC) multi-episode spells is to use the Single Spell Grouper, available from the navigation pane. To get an initial appreciation of grouping you are recommended to experiment with the Single Spell functionality.

Record Definition File (RDF)

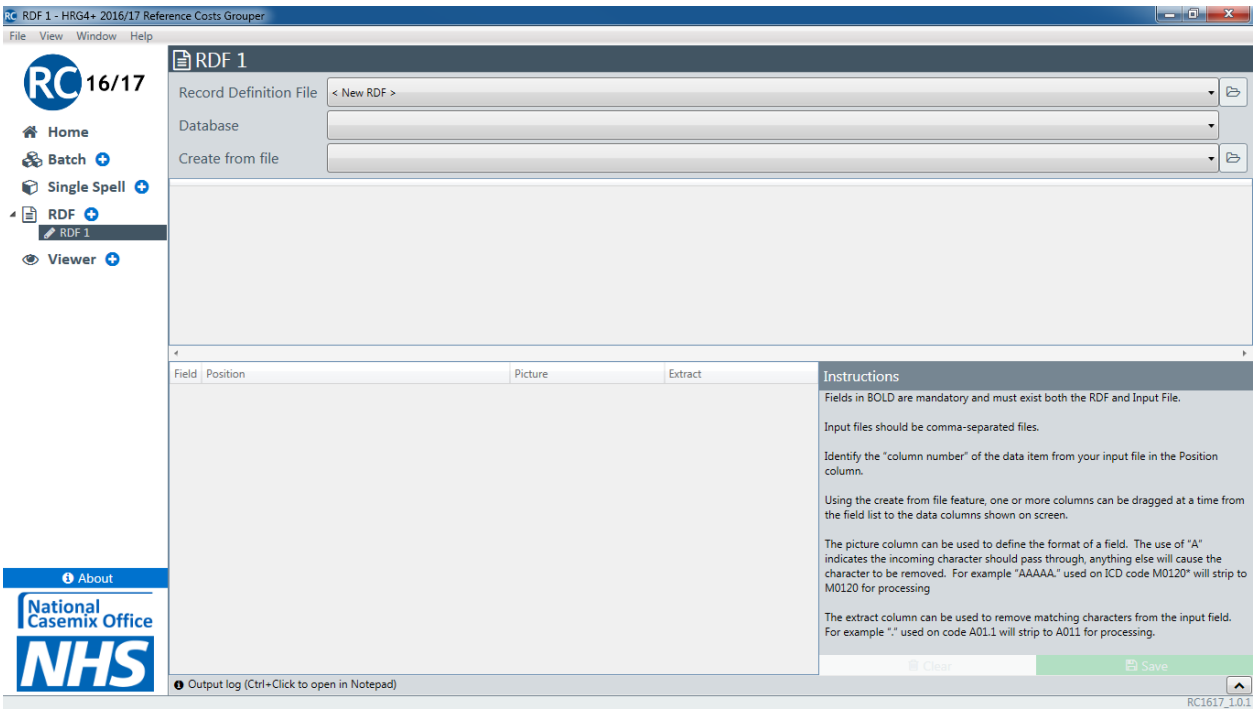
The RDF tells the grouper the structure of the data you want it to process. The field content of the RDF must include all mandatory input fields for the type of data you want to group. Although these fields are mandatory, you define their order. You can also include additional fields, which will be ignored by the Grouper.

The product comes with a set of standard RDFs that hold all the mandatory fields required to be able to group data. You can adapt your data to match the standard RDFs or create an RDF manually.

To create or modify an RDF the grouper has an interactive RDF module.

RDF Module

The RDF module allows a user to create an RDF from scratch or to modify an existing RDF.



The RDF columns are:

Name	Description
Field Name	A field name abbreviation for internal use; it cannot be modified.
Field Title	A field description which can be edited to display a locally-defined field name.
Field No	The field column number in the input file.
Picture	Specifies whether specific character positions in the field are used or ignored during processing. It must be specified as a string of “A” and “.” where “A” means keep the character in that position and “.” means ignore the character in that position. For example: “AAA.A” means ignore the 4th character and process characters in positions 1,2,3 and 5 as if it is a 4 character code, so that W58.1 gets processed as W581
Extract	Allows you to specify that specific characters (in any position) in the field are ignored during processing. For example “.” (without quote marks) will remove all full stops from the field.

Create an RDF manually – Create from file

Select the icon next to **RDF** on the navigation pane or select ‘**New’** or the icon on the main home screen or select ‘**File**’ on the menu bar, ‘**New**’ and then ‘**RDF**’.

When the RDF page is opened, the **Record Definition File** box is automatically set to **New RDF**.

Select the required **Database** you wish to create the RDF for (the type of input file with which the RDF will be used).

You can choose to create an RDF manually based just on the **Database** selection or **Create from file** which uses an input data file to base the RDF on.

Select **Create from file**, click the icon, and in the **Open** dialog box browse to and select your input data file (which should be format *.CSV).

The first 30 rows of the selected input file are displayed in a preview window under the **Create from file** drop down box. If the input file contains headers then this will be displayed in the preview window. Each field displayed in the preview window is assigned a number (e.g. 1, 2, 3).

Please note, if a database has not been selected, the preview scroll bars are disabled.

Depending on the data set type, the number of repeating occurrences of certain fields can be increased or decreased using the Number of fields function. For example, if the input file contains 20 diagnosis codes, the number of diagnosis fields can be increased to 20 using the toggle buttons rather than the default value of 14.

All field names highlighted bold are mandatory and must be assigned to fields in the input file. Fields are assigned by either dragging the field name from the field columns list to the appropriate field number in the file preview window or typing the field number directly into the **Position** column. Once a field is dragged into the file preview window the **Position** column will populate with the relevant field position.

Multiple fields can be selected by clicking with the mouse while holding down the SHIFT and/or CTRL keys. The selected fields can then be dragged to the file preview area.

If a field is erroneously assigned, it can be removed by selecting the field position from the **Position** column and pressing the **Delete** key. The field will then be removed from the file preview window. If you wish to clear all the field positions allocated then the **Clear** button will reset all the field positions.

When all mandatory fields have been assigned the RDF can be saved by selecting the **Save** button.

Create an RDF manually – Database

Following the steps above, an RDF can be created manually by either choosing to **Create from file** or just selecting a **Database**.

By selecting just a **Database** to create an RDF, the user is presented with the field list window. The file preview window will remain blank.

Depending on the data set type, the number of repeating occurrences of certain fields can be increased or decreased using the Number of fields function. For example, if the input file contains 20 diagnosis codes, the number of diagnosis fields can be increased to 20 using the toggle buttons rather than the default value of 14.

Assign a number to each of the field names in the **Position** column.

When all mandatory fields have been assigned the RDF can be saved by selecting the **Save** button.

Edit an Existing RDF

Rather than creating an RDF from scratch, a user can edit an already existing RDF.

Select the next to **RDF** on the navigation pane or select **New** or the icon on the main home screen or select '**File**' on the menu bar, '**New**' and then '**RDF**'.

The **Record Definition File** is automatically set to **New RDF**. Click the browse button (folder icon), and in the **Open RDF File...**dialog box browse and select the RDF you wish to edit.

The Database box will automatically populate with the relevant database depending on the mandatory fields contained in the RDF. If the RDF contains none of the mandatory fields for any of the datasets then the field position box will remain blank.

Once selected, the file preview window will display the fields for the relevant Database and the Field window will display the RDF field names and positions.

You can then modify the field positions and field attributes in the RDF. If you wish to increase or decrease some of the repeating occurrences of certain fields, the Number of fields function can be used. For example, if the input file contains 20 diagnosis codes, the number of diagnosis fields can be increased to 20 using the toggle buttons rather than the default value of 14.

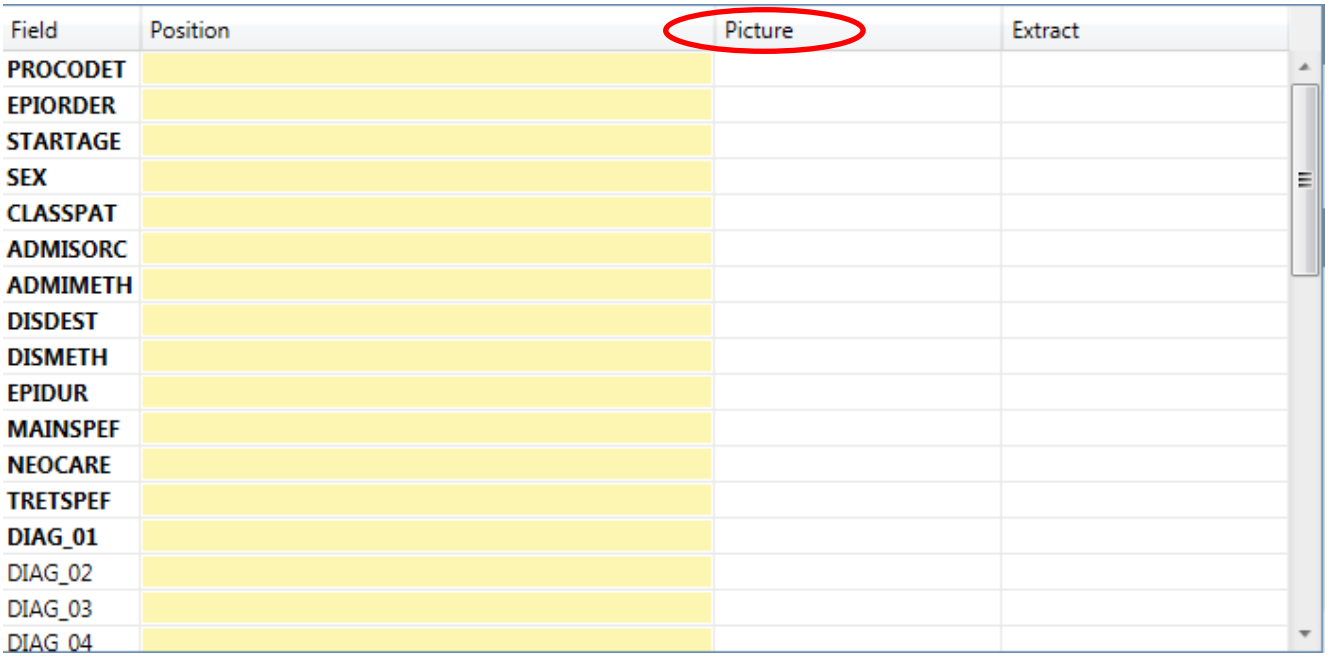
Extra fields can be added to the RDF using the blank row at the bottom or highlighting a field row, right-click and select **Insert Row**.

To delete a non-mandatory field, select the field row to be deleted and either selecting **Delete** or right-click and select **Remove Row**, the field will be removed from the field list. Mandatory fields cannot be deleted from an RDF.

Picture

Picture is a feature within the RDF that can be used to specify the inclusion or exclusion of specific character positions from input fields. It allows the user to provide filtering instructions for each field by describing a character template. During processing the application will apply this template to the field by selectively ignoring characters in specific positions within a field.

An “A” character indicates a character position to be included in grouping and a full-stop indicates a character position to be ignored in grouping. This ‘picture’ is then imposed on the field before grouping, meaning that the modified version of the field contents is processed by the Grouper.



To use Picture, enter a character template in the ‘Picture’ column of the Record Definition File using the RDF module.

- Characters should not be separated by spaces
- Quotation marks should not be used
- Can be used with alpha and numeric fields
- Where the number of characters in the field exceeds the number of characters specified in Picture, the application will implement the Picture on the left-most portion of the field. For this reason, care should be taken when using Picture with variable length fields as this may lead to unpredictable results.
- Picture is applied prior to data validation.
- Picture applies to file processing only; it does not affect single-spell grouping.

Example 1 AAA.AA Characters 1, 2, 3, 5 and 6 are used. Character 4 is ignored

Example 2 .A.AAA Characters 2, 4, 5 and 6 are used. Characters 1 and 3 are ignored

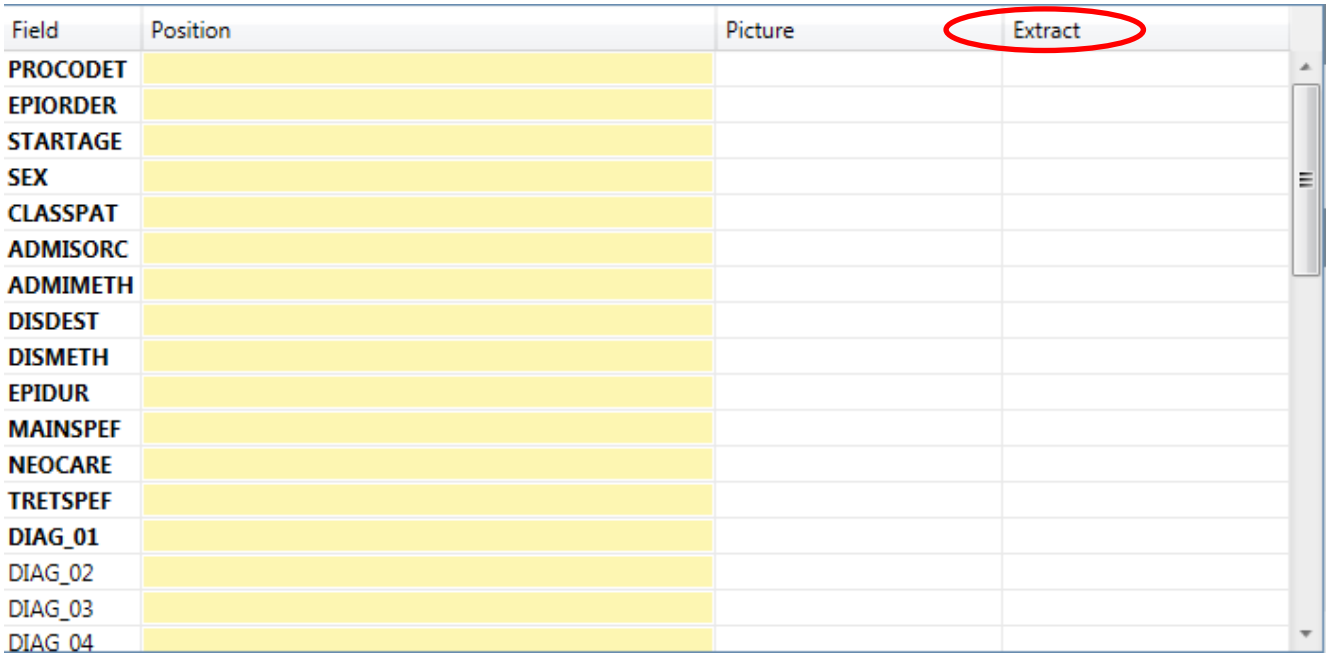
Example 3 .A. Character 2 is used. Characters 1 and 3 are ignored

Example 4 AA Characters 1 and 2 are used (Note that full stops are not required if consecutive characters from the beginning of the field are the only characters required).

Extract

Extract is a feature of the RDF that can be used to make the Grouper ignore specific characters in input fields, irrespective of character position.

Entering one or more characters in the 'Extract' column of the RDF using the RDF module will instruct the grouper that any occurrence of any of these characters in that field should be ignored in grouping. (Note: Commas can not be excluded because the input data files are comma-separated).



Field	Position	Picture	Extract
PROCODET			
EPIORDER			
STARTAGE			
SEX			
CLASSPAT			
ADMISORC			
ADMIMETH			
DISDEST			
DISMETH			
EPIDUR			
MAINSPEF			
NEOCARE			
TRETSPEF			
DIAG_01			
DIAG_02			
DIAG_03			
DIAG_04			

The application does not consider a series of characters as a string of characters to be ignored as a single entity. Each character entered is ignored wherever it appears in the input field.

- The characters specified will be ignored wherever they appear in the input data field.
- There is no relation between the position of a character in the extract field and the input field.
- All occurrences of the characters are affected.
- Extract is applied prior to data validation.
- Extract applies to file processing only and it does not affect single spell.

Extract Field Examples

The examples below are shown in quotation marks (“) to help identify the examples. Quotation marks should not be entered when using the Extract function (Unless a user wishes to “ignore” quotation marks).

Example 1

Entering an Extract character of “.” informs the application that all occurrences of a decimal place or full-stop should be ignored when processing that field. This is a convenient way to remove decimal points from a field in the input file. An input field containing the character string “abc…d” will be interpreted by the application as “abcd”.

Example 2

Entering the characters “+$” in this field informs the application that any occurrences of both the “+” character and the “$” character are to be ignored by the application when processing that field. An input field containing the character string “46+$$++” will be interpreted by the application as “46”.

Command Line Initiation

“Command line” refers to the character-based command shell user interface within Microsoft Windows, commonly known as the DOS Window. The Grouper can process a file using the command line with parameters shown below.

The parameters consist of a hyphen followed by a lower-case letter and, where applicable, a value (shown in italics). All parameters are mandatory except for the optional parameters shown enclosed in square brackets.

The command line takes the following form:

`hrggrouperc.exe -i` *`Input_File`* `-o` *`Output_File`* `-d` *`RDF_File`* `-l` *`Grouping_Logic`* `[-h] [-t] [-v] [-?] [>` *`Log_File`*`]`

Command line parameters:

`-i` *`Input_File`*	The path and filename of the **i**nput file.
`-o` *`Output_File`*	The path and filename of the **o**utput file.
`-d` *`RDF_File`*	The path and filename of the record **d**efinition file.
`-l` *`Grouping_Logic`*	The grouping **l**ogic. The available values are: APC, ACC, EM, NAC, PCC, NRD, NCC.
`-h`	Optional. Where present, indicates that the input file has a **h**eader row (field names). Omit this parameter where the first row in the file is a data row.
`-t`	Optional. Where present, stops generation of field names in the **t**op row of the output file. If this parameter is omitted then each of the output files will contain a header row.
`-v`	Optional. Increases verbosity of the log output.
`-?`	Optional. Where present, directs the Grouper to list the available command line parameters. This parameter cannot be combined with other parameters.
`>` *`Log_File`*	Optional. Where present, the Grouper redirects the log output to a specified file. If not used, the log output is written to the screen.

Where there are spaces in paths or filenames for parameter values, the parameter value must be enclosed within double-quote characters.

Running the Grouper

It is recommended to invoke the command line Grouper from a script (also known as a batch file). For example to group some APC data (with headings) in C:\Temp\data\apc.csv using a suitable Record Definition File (APC.rdf) in the same directory, then a suitable script is:

```
@echo off

cd /d "c:\Program Files\NHS Digital\HRG4+ 2016_17 Reference Costs Grouper"

hrggrouperc.exe -i "c:\Temp\data\apc.csv" -o "c:\Temp\data\output.csv" -
d "c:\Temp\data\apc.rdf" -l APC -h > "c:\Temp\data\hrg.log"

if %ERRORLEVEL% neq 0 echo Error in command, please check hrg.log

pause
```

Alternatively, as from Reference Costs 2016/17, it is possible to run the Grouper from a working directory by invoking hrggrouperc.exe as a fully qualified path:

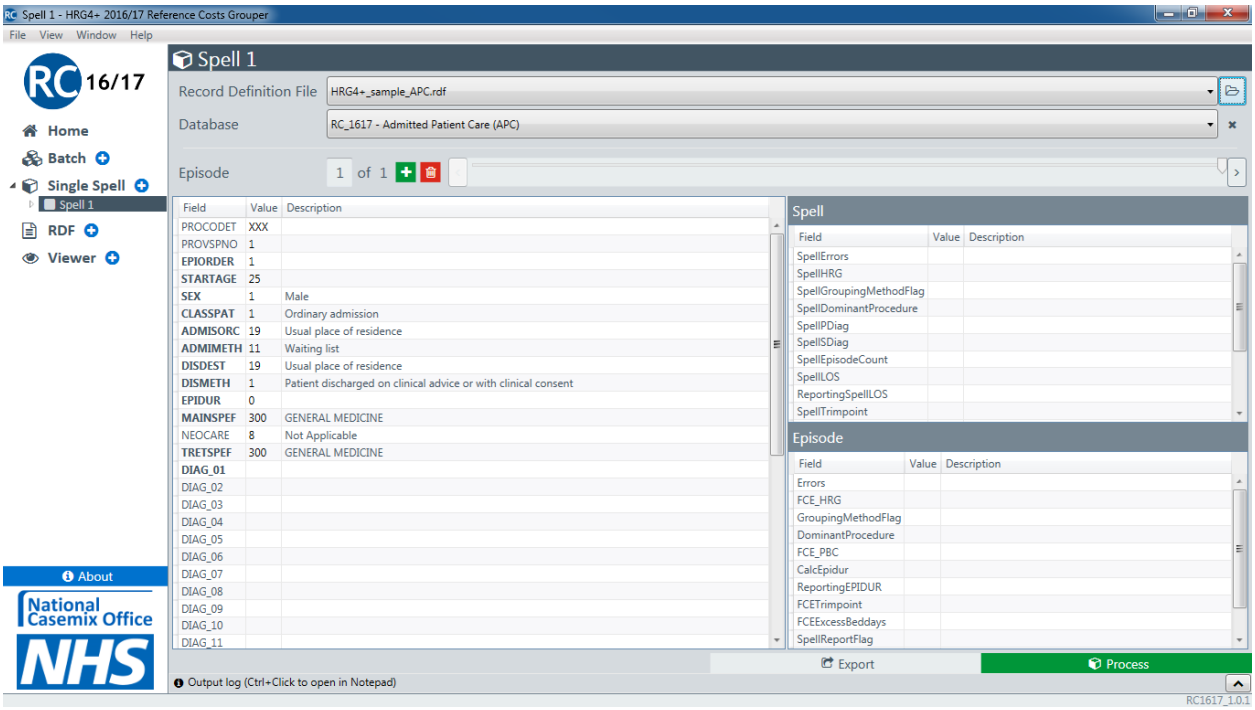
```
@echo off

cd /d "c:\Temp\data"

"c:\Program Files\NHS Digital\HRG4+ 2016_17 Reference Costs
Grouper\hrggrouperc.exe" -i "apc.csv" -o "output.csv" -d "apc.rdf" -l APC
-h > "hrg.log"

if %ERRORLEVEL% neq 0 echo Error in command, please check hrg.log

pause
```

The penultimate statement of these scripts, checks whether the Grouper run has successfully completed and prints a message if the run has failed.

Single Spell Grouping

The Single Spell window is best used to help understand how outputs are derived for a particular spell, episode or attendance. By adding the input data manually a user can see how the individual codes affect the output in real time. This can be useful when exploring how certain combinations of clinical codes lead to different HRGs being derived. It can also be useful to quickly find the descriptions of codes and ascertain which fields are mandatory (expressed in **bold** text).

The Single Spell window can be accessed by selecting the icon next to **Single Spell** on the navigation pane or select **New** or the icon on the main home screen or select '**File**' on the menu bar, '**New**' and then '**Single Spell**'.



To add a Record you can either select a **Record Definition File** which will automatically choose the relevant **Database** or you can just choose a **Database**.

For some of the databases such as Admitted Patient Care it is possible to build up a spell consisting of multiple episodes by selecting the icon on the Single Spell window.

You can navigate between different episodes within a Spell by either clicking on the **Episode** in the navigation pane, entering an episode number in the **Episode Number box** or use the **Slider bar**. Episodes can be deleted by either using the icon or right-clicking on an episode in the navigation pane and select **Delete Episode**. To close the Spell, right-click on the relevant Spell number in the navigation pane and select **Close**.

Certain fields are populated with useful values to save some data input effort, and the user can edit these and populate other fields as appropriate. Changing or adding values to the input fields will automatically update the output values. If using a database that allows multiple episodes to be opened, any new episodes added will require the **Process** button to be selected or changes made to the input fields for the output values to be populated. You will see outputs for each record (episode) shown in the "**Episode**" window and see outputs for the entire Spell in the "**Spell**" window.

Once patient information is entered for an episode, all attributes other than non-primary diagnosis codes and procedure codes are repeated for a new episode within the spell saving the user from needing to change relevant fields for every episode. The Diag_01 field uses the same diagnosis code as the previous episode. All other diagnosis codes and the procedure codes start as blank fields in newly added episodes.

Entering values does not have to be monotonous. A user can copy a row or column of codes from applications such as Excel and paste them into the Single Spell. The Single Spell will recognise a string of codes and paste them appropriately transposing a row of codes automatically.

For example; after ensuring the columns are in the same order as the Single Spell rows, a user can simply copy a row of episode data from Excel:

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	PROCODET	PROVSPNO	EPIORDER	STARTAGE	SEX	CLASSPAT	ADMISORC	ADMIMETH	DISDEST	DISMETH	EPIDUR	MAINSPEF	NEOCARE	TRETSPEF	DIAG_01	DIAG_02	DIAG_03	DIAG_04
2	XXX	1	1	40	2	1	19	11	98	8	3	502	8	502	N804	N803	N805	N736
3	XXX	1	2	40	2	1	19	11	98	8	43	170	8	221	Q211			
4	XXX	1	3	40	2	1	19	11	98	8	0	420	8	253	D619			

A user can simply paste them into the Single Spell (Ctrl + V)

Field	Value	Description
PROCODET	XXX	
PROVSPNO	1	
EPIORDER	1	
STARTAGE	40	
SEX	2	Female
CLASSPAT	1	Ordinary admission
ADMISORC	19	Usual place of residence
ADMIMETH	11	Waiting list
DISDEST	98	Not applicable - Hospital Provider Spell not finished at episode end (i.e. not discharged) or current
DISMETH	8	Not applicable - Hospital Provider Spell not finished at episode end (i.e. not discharged) or current
EPIDUR	3	
MAINSPEF	502	GYNAECOLOGY
NEOCARE	8	Not Applicable
TRETSPEF	502	GYNAECOLOGY
DIAG_01	N804	Endometriosis of rectovaginal septum and vagina
DIAG_02	N803	Endometriosis of pelvic peritoneum
DIAG_03	N805	Endometriosis of intestine
DIAG_04	N736	Female pelvic peritoneal adhesions
DIAG_05		
DIAG_06		
DIAG_07		
DIAG_08		
DIAG_09		
DIAG_10		
DIAG_11		

Notice how it automatically transposes and pastes a row of codes into the column.

In addition to the copy and paste, the Grouper supports “Smart Pasting” which allows a user to copy a row of codes from applications such as Excel and along with the headers and paste them into the Single Spell. Smart Pasting recognises which fields to paste the codes into meaning that the ordering of the fields do not need to match the same structure of the RDF. In order for Smart Pasting to work the headers from which the codes are being copied must match the names in the Single Spell window.

For example; after copying a row of data from Excel with matching headings to those in Single Spell:

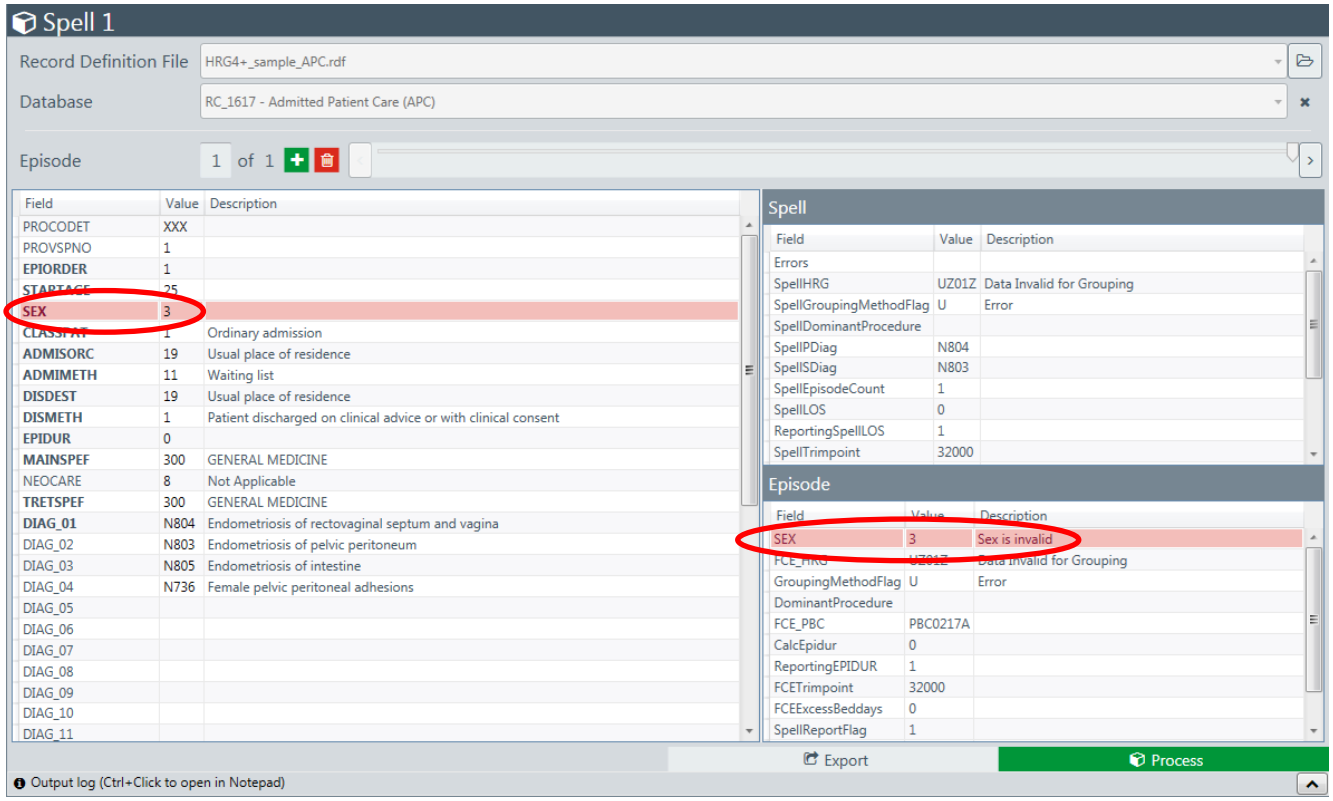
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R
1	PROCODET	PROVSPNO	EPIORDER	STARTAGE	SEX	ADMIMETH	ADMISORC	CLASSPAT	DISMETH	DISDEST	MAINSPEF	TRETSPEF	NEOCARE	EPIDUR	DIAG_01	DIAG_02	DIAG_03	DIAG_04
2	XXX	1	1	40	2	11	19	1	8	98	502	502	8	3	N804	N803	N805	N736
3	XXX	1	2	40	2	11	19	1	8	98	170	221	8	43	Q211			
4	XXX	1	3	40	2	11	19	1	8	98	420	253	8	0	D619			

A user can simply paste them into the Single Spell (Ctrl + V)

Field	Value	Description
PROCODET	XXX	
PROVSPNO	1	
EPIORDER	1	
STARTAGE	40	
SEX	2	Female
CLASSPAT	1	Ordinary admission
ADMISORC	19	Usual place of residence
ADMIMETH	11	Waiting list
DISDEST	98	Not applicable - Hospital Provider Spell not finished at episode end (i.e. not discharged) or current
DISMETH	8	Not applicable - Hospital Provider Spell not finished at episode end (i.e. not discharged) or current
EPIDUR	3	
MAINSPEF	502	GYNAECOLOGY
NEOCARE	8	Not Applicable
TRETSPEF	502	GYNAECOLOGY
DIAG_01	N804	Endometriosis of rectovaginal septum and vagina
DIAG_02	N803	Endometriosis of pelvic peritoneum
DIAG_03	N805	Endometriosis of intestine
DIAG_04	N736	Female pelvic peritoneal adhesions
DIAG_05		
DIAG_06		
DIAG_07		
DIAG_08		
DIAG_09		
DIAG_10		
DIAG_11		

For information on what the output values mean please jump to the Output Files section of this document.

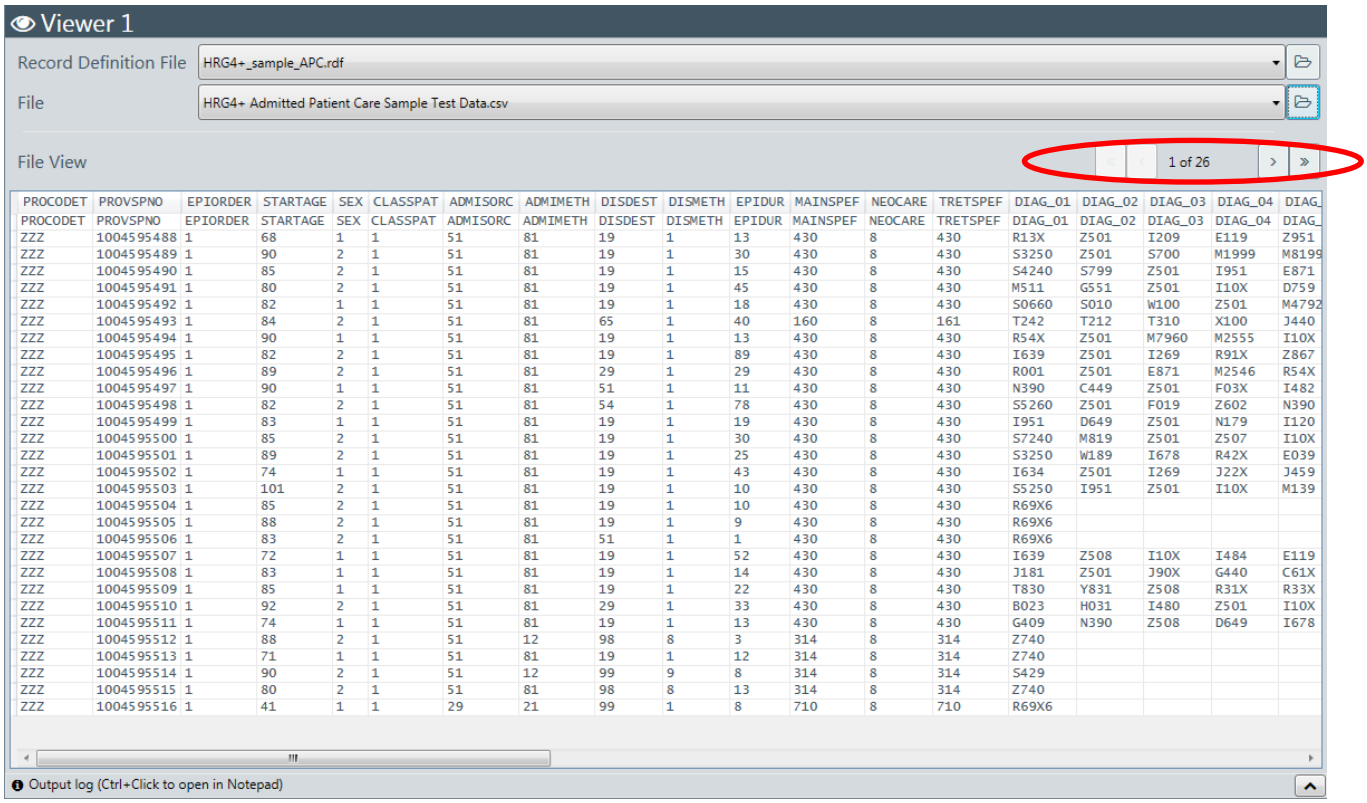
Any errors are highlighted in red in both the Input and Output sections.



Viewer

The Grouper provides a file viewer which can be used to view input or output files. To access the file viewer select **Viewer** from the navigation pane.

Selecting a **Record Definition File** and a **File** will display the input file data organised into the columns specified in the RDF and can be used as an aid to see if there are problems in the layout or content of a data file.



PROCODET	PROVSPNO	EPIORDER	STARTAGE	SEX	CLASSPAT	ADMISORC	ADMIMETH	DISDEST	DISMETH	EPIDUR	MAINSPEF	NEOCARE	TRETSPEF	DIAG_01	DIAG_02	DIAG_03	DIAG_04	DIAG_
PROCODET	PROVSPNO	EPIORDER	STARTAGE	SEX	CLASSPAT	ADMISORC	ADMIMETH	DISDEST	DISMETH	EPIDUR	MAINSPEF	NEOCARE	TRETSPEF	DIAG_01	DIAG_02	DIAG_03	DIAG_04	DIAG_
ZZZ	1004595488	1	68	1	1	51	81	19	1	13	430	8	430	R13X	Z501	I209	E119	Z951
ZZZ	1004595489	1	90	2	1	51	81	19	1	30	430	8	430	S3250	Z501	S700	M1999	M8199
ZZZ	1004595490	1	85	2	1	51	81	19	1	15	430	8	430	S4240	S799	Z501	I951	E871
ZZZ	1004595491	1	80	2	1	51	81	19	1	45	430	8	430	M511	G551	Z501	I10X	D759
ZZZ	1004595492	1	82	1	1	51	81	19	1	18	430	8	430	S0660	S010	W100	Z501	M4792
ZZZ	1004595493	1	84	2	1	51	81	65	1	40	160	8	161	T242	T212	T310	X100	J440
ZZZ	1004595494	1	90	1	1	51	81	19	1	13	430	8	430	R54X	Z501	M7960	M2555	I10X
ZZZ	1004595495	1	82	2	1	51	81	19	1	89	430	8	430	I639	Z501	I269	R91X	Z867
ZZZ	1004595496	1	89	2	1	51	81	29	1	29	430	8	430	R001	Z501	E871	M2546	R54X
ZZZ	1004595497	1	90	1	1	51	81	51	1	11	430	8	430	N390	C449	Z501	F03X	I482
ZZZ	1004595498	1	82	2	1	51	81	54	1	78	430	8	430	S5260	Z501	F019	Z602	N390
ZZZ	1004595499	1	83	1	1	51	81	19	1	19	430	8	430	I951	D649	Z501	N179	I120
ZZZ	1004595500	1	85	2	1	51	81	19	1	30	430	8	430	S7240	M819	Z501	Z507	I10X
ZZZ	1004595501	1	89	2	1	51	81	19	1	25	430	8	430	S3250	W189	I678	R42X	E039
ZZZ	1004595502	1	74	1	1	51	81	19	1	43	430	8	430	I634	Z501	I269	J22X	J459
ZZZ	1004595503	1	101	2	1	51	81	19	1	10	430	8	430	S5250	I951	Z501	I10X	M139
ZZZ	1004595504	1	85	2	1	51	81	19	1	10	430	8	430	R69X6				
ZZZ	1004595505	1	88	2	1	51	81	19	1	9	430	8	430	R69X6				
ZZZ	1004595506	1	83	2	1	51	81	51	1	1	430	8	430	R69X6				
ZZZ	1004595507	1	72	1	1	51	81	19	1	52	430	8	430	I639	Z508	I10X	I484	E119
ZZZ	1004595508	1	83	1	1	51	81	19	1	14	430	8	430	J181	Z501	J90X	G440	C61X
ZZZ	1004595509	1	85	1	1	51	81	19	1	22	430	8	430	T830	Y831	Z508	R31X	R33X
ZZZ	1004595510	1	92	2	1	51	81	29	1	33	430	8	430	B023	H031	I480	Z501	I10X
ZZZ	1004595511	1	74	1	1	51	81	19	1	13	430	8	430	G409	N390	Z508	D649	I678
ZZZ	1004595512	1	88	2	1	51	12	98	8	3	314	8	314	Z740				
ZZZ	1004595513	1	71	1	1	51	81	19	1	12	314	8	314	Z740				
ZZZ	1004595514	1	90	2	1	51	12	99	9	8	314	8	314	S429				
ZZZ	1004595515	1	80	2	1	51	81	98	8	13	314	8	314	Z740				
ZZZ	1004595516	1	41	1	1	29	21	99	1	8	710	8	710	R69X6				

Depending on the number of rows of data in your data file, the Viewer will display the data over a number of pages. You can use the navigation buttons to move between pages in the Viewer window or move to the first or last page of the Viewer window.

Input File Preparation

This section provides guidance about preparing input files for processing by the Grouper.

Validation

Unless otherwise stated, the values in each field are validated against enumerated sets of values, predominantly based on NHS Data Dictionary definitions.

File Format

Input data must be in comma-separated value (CSV) format using the American Standard Code for Information Interchange (ASCII) character set, excluding the non-printing ASCII characters. The input file must not include qualifiers, such as quotes, surrounding the fields.

Text qualifiers in any original fields can be ignored by selecting the appropriate “Text Qualifiers < > none” when either extracting the CSV or importing into Excel prior to saving as a CSV file.

A Trouble Shooting section has been included in this document. If when loading your data you see the following error message – **“failure to set input file”,** this indicates that your file is not in the format currently accepted by the Grouper. The workaround to this issues is covered in detail within this section.

Using Excel for File Preparation

If some of the records end with empty fields then a file structure issue can arise when saving a text file using Microsoft Excel (for more information please refer to Microsoft Article ID 77295 “Column delimiters are missing in an Excel spreadsheet that is saved as text” [https://support.microsoft.com/en-us/kb/77295]). To prevent this issue arising, ensure that the final (rightmost) column of data is populated for every record in the file with dummy data such as “x”.

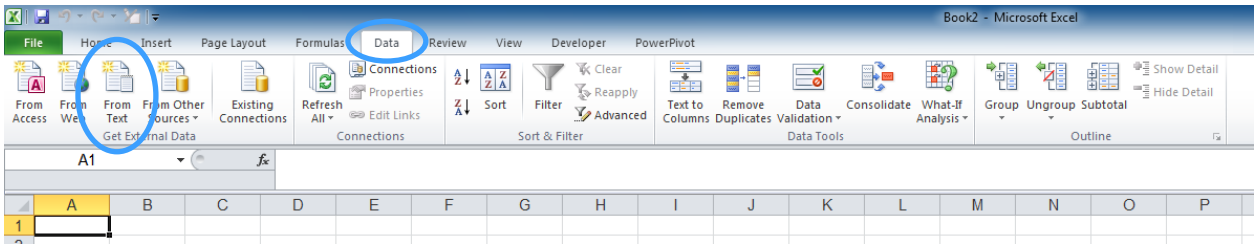
Data Set Types with Leading Zeros

Please be aware that unless properly imported, opening a CSV in Excel will cause changes to the data. A common occurrence is for leading zeros to be stripped when a CSV file is opened in Excel.

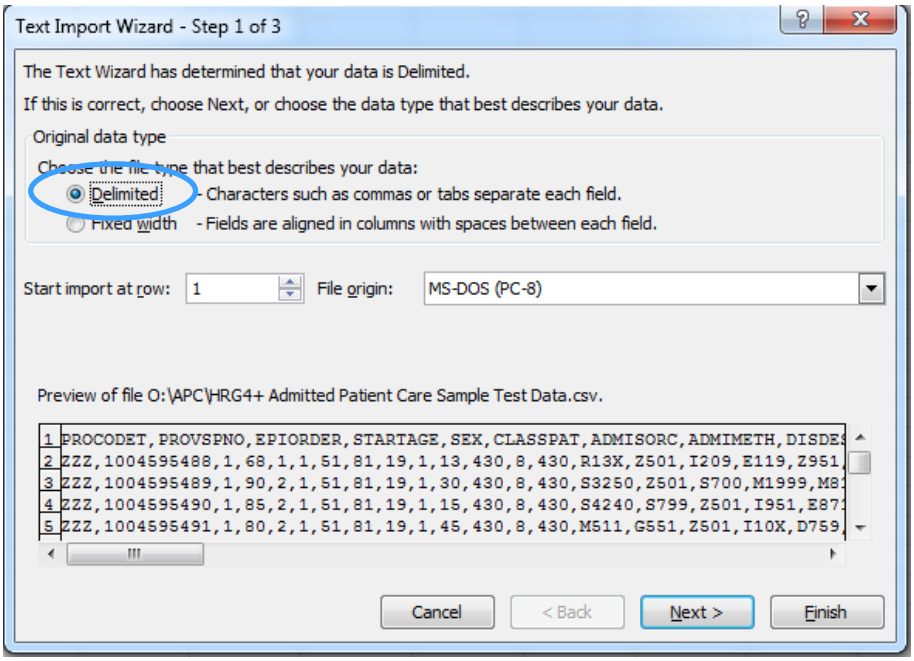
All Data Set Types utilised by the Grouper, with the exception of APC, use leading zeros in certain fields. If the leading zeros are not present then the record may fail validation or worse, will lead to incorrect grouping (this is especially important for EM grouping where Treatment Codes such as 011 and 11 are different).

The following screen shots illustrate how a CSV file can be imported into Excel for manipulation, while maintaining leading zeros.

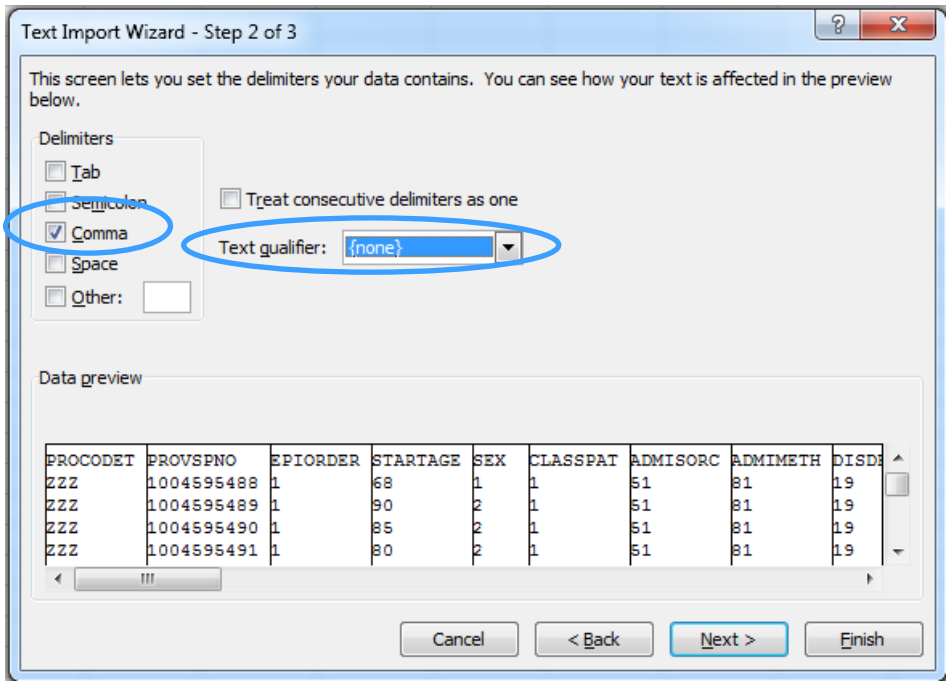
1. Open Excel. Click on the “Data” tab and click “From Text” in the ribbon.



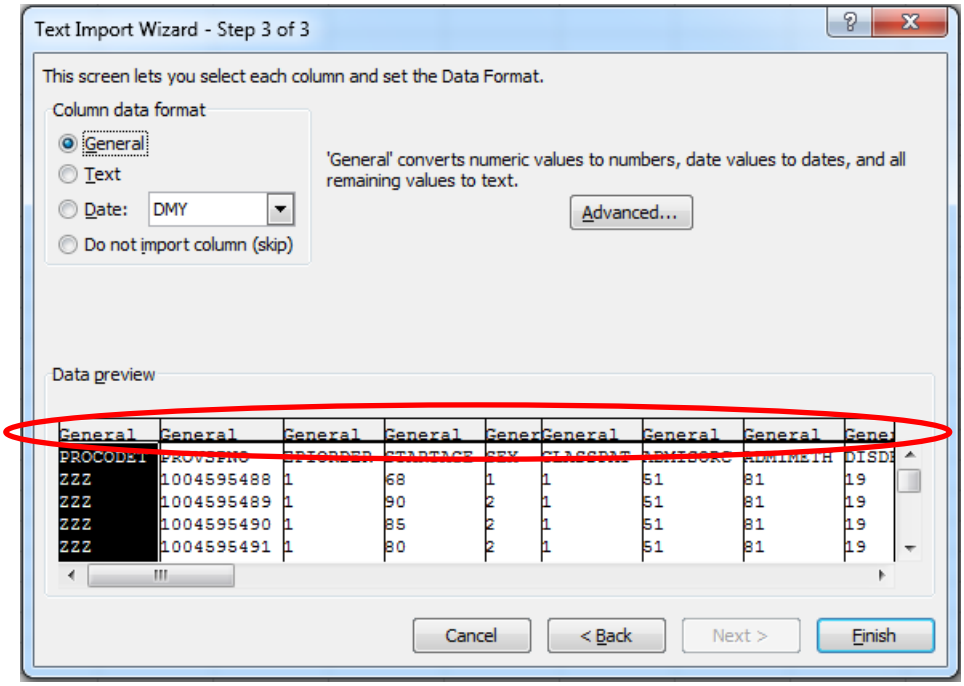
2. Navigate to your input file. In the Text Import Wizard select “Delimited” in step 1/3



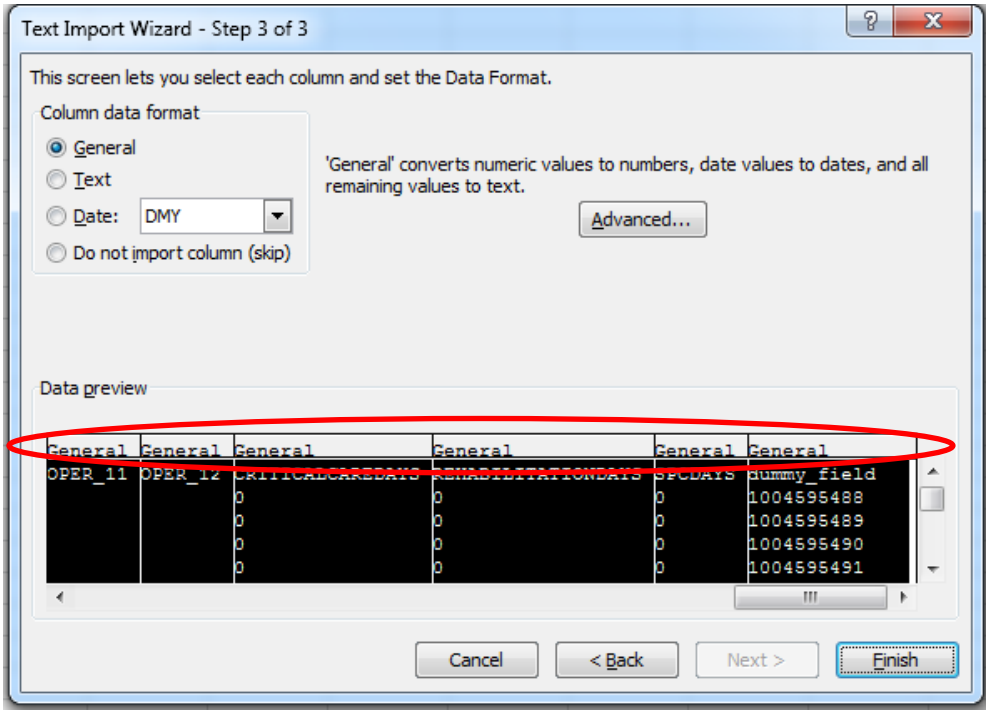
3. Then select “Comma” delimited in step 2/3



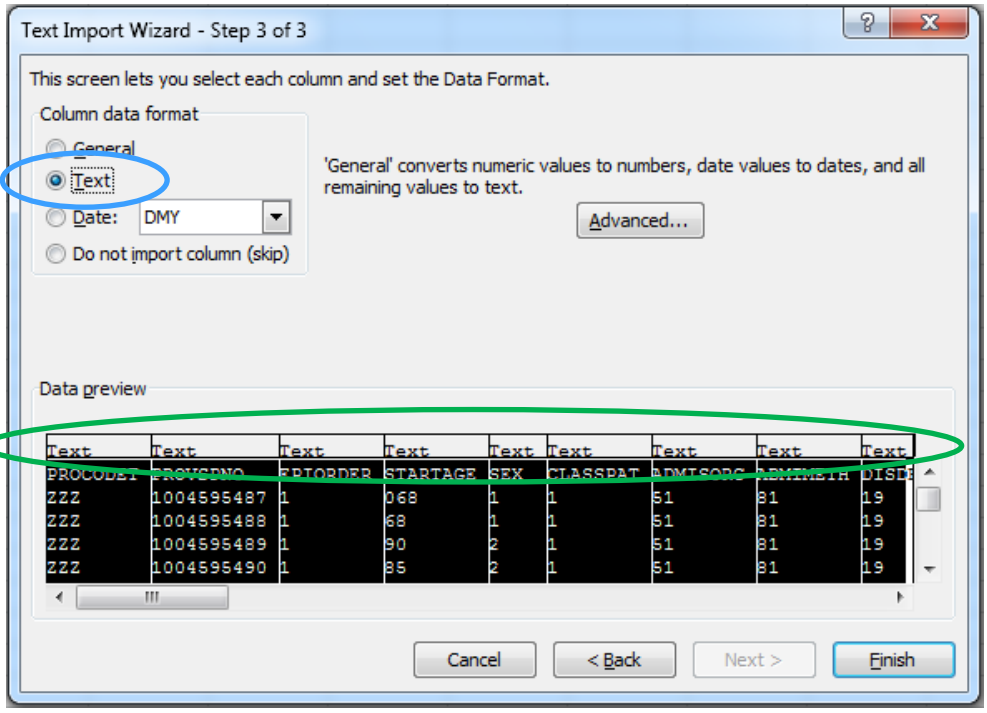
4. In step 3/3 notice Excel wants to pull the data into cells formatted as “general” not “text”.



It is essential to highlight all columns to change the format option to text. To do this scroll all the way to the right and press “shift” and click on the last column. This will select all columns.



Then choose the “Text” option.



Finally click “Finish” and choose what cell you want to pull the data into. This will ensure the data is imported as text for all columns.

It is recommended that you pull the data into cell A1 if you intend to save as CSV and run through the Grouper.

For more information on which fields have leading zeros please see the notes column in the below data set type tables.

Admitted Patient Care (APC)

The Grouper sorts APC data prior to processing so that records with the same Provider Code and Provider Spell Number are placed in Episode Number order.

The field list below shows the fields required by the Grouper. The field order of the input file is defined by the Record Definition File used for grouping; there is no significance to the order in which the fields are listed below.

Grouper Field Name	CDS/DD Field Name	Notes
PROCODET	ORGANISATION CODE (CODE OF PROVIDER)	A value must be supplied but is not validated. For NHS organisations use the first 3 characters for non-NHS organisations use all 5 characters
PROVSPNO	HOSPITAL PROVIDER SPELL NUMBER (or alternative)	A value must be supplied but is not validated. It is possible to use an alternative spell identifier as PROVSPNO, in place of the Hospital Provider Spell Number
EPIORDER	EPISODE NUMBER	Duplicates within a spell will create an error. EPIORDER values 98 and 99 are invalid for grouping
STARTAGE	Derived	Whole years rounded down: START DATE (EPISODE) – PERSON BIRTH DATE
SEX	PERSON GENDER CODE CURRENT	Must be same for all episodes in a spell
CLASSPAT	PATIENT CLASSIFICATION CODE	Must be the same for all episodes within a spell.
ADMISORC	SOURCE OF ADMISSION CODE (HOSPITAL PROVIDER SPELL)	Must be the same for all episodes within a spell
ADMIMETH	ADMISSION METHOD CODE (HOSPITAL PROVIDER SPELL)	Must be the same for all episodes within a spell
DISDEST	DISCHARGE DESTINATION CODE (HOSPITAL PROVIDER SPELL)	
DISMETH	DISCHARGE METHOD CODE (HOSPITAL PROVIDER SPELL)	
EPIDUR	Derived	Range: 0-99999. Whole days: END DATE (EPISODE) – START DATE (EPISODE)

Grouper Field Name	CDS/DD Field Name	Notes
MAINSPEF	CARE PROFESSIONAL MAIN SPECIALTY CODE	
NEOCARE	NEONATAL LEVEL OF CARE CODE	May be blank
TRETSPEF	ACTIVITY TREATMENT FUNCTION CODE	
DIAG_01	PRIMARY DIAGNOSIS (ICD)	Valid ICD-10 code. If this field is blank an error will be generated
DIAG_02 - DIAG_99	SECONDARY DIAGNOSES (ICD)	As above, but blank allowed
OPER_01 - OPER_99	PROCEDURE (OPCS)	Valid OPCS-4 codes or blank
CRITICALCAREDAYS	LENGTH OF STAY ADJUSTMENT (CRITICAL CARE, for all data set types)	Range: 0-99999 or blank. Count of distinct days within the episode when the patient was in Critical Care. If the patient was in Critical Care on the last day of an episode that was not the last episode of the spell, assign that day to the next episode
REHABILITATIONDAYS	LENGTH OF STAY ADJUSTMENT (REHABILITATION)	Range: 0-99999 or blank
SPCDAYS	LENGTH OF STAY ADJUSTMENT (SPECIALIST PALLIATIVE CARE)	Range: 0-99999 or blank

Each row of the input file represents a single episode, which may or may not be part of a multi-episode spell.

Non-Admitted Consultations (NAC)

Non-admitted consultations are outpatient attendances or attendances by patients for nursing care on a ward (ward attenders). Since 1 April 2005 the Outpatient Attendance CDS has contained both of these types of data.

The field list below shows the fields required by the Grouper. The field order of the input file is defined by the Record Definition File used for grouping; there is no significance to the order in which the fields are listed below.

Grouper Field name	CDS/DD Field Name	Notes
STARTAGE	Derived	Whole years rounded down: APPOINTMENT DATE – PERSON BIRTH DATE
SEX	PERSON GENDER CODE CURRENT	
MAINSPEF	CARE PROFESSIONAL MAIN SPECIALTY CODE	
TRETSPEF	ACTIVITY TREATMENT FUNCTION CODE	
FIRSTATT	FIRST ATTENDANCE CODE	
OPER_01 - OPER_99	PROCEDURE (OPCS)	Valid OPCS-4 codes or blank

Each row of the input file represents a single attendance.

Diagnosis Codes

Diagnosis coding is not utilised in the HRG4+ algorithm for non-admitted consultations.

Emergency Medicine (EM)

The field list below shows the fields required by the Grouper. The field order of the input file is defined by the Record Definition File used for grouping; there is no significance to the order in which the fields are listed below.

Grouper Field Name	CDS/DD Field Name	Notes
AGE	Derived	Whole years rounded down: ARRIVAL DATE – PERSON BIRTH DATE Validated, but not used in grouping
AEPATIENTGROUP	A AND E PATIENT GROUP	If populated must comply with NHS Data Dictionary standards
INV_01 - INV_99	ACCIDENT AND EMERGENCY INVESTIGATION – FIRST and ACCIDENT AND EMERGENCY INVESTIGATION – SECOND	Valid national code component (always 2 characters) or blank. The “Local Sub-Analysis” part should not be submitted. Leading zeros must be included where they form part of the national code component
TREAT_01 - TREAT_99	ACCIDENT AND EMERGENCY TREATMENT – FIRST and ACCIDENT AND EMERGENCY TREATMENT – SECOND	Valid national code component (2 or 3 characters) or blank. The “Local Sub-Analysis” part should not be submitted. Leading zeros must be included where they form part of the national code component

Each row of the input file represents one Accident and Emergency Attendance.

Although either of the Investigation and Treatment Code fields can be blank, where both of these field types are blank, the resultant HRG will be UZ01Z.

Renal Dialysis (NRD)

Renal Dialysis HRGs are generated using fields from the National Renal Dataset.

The field list below shows the fields required by the Grouper. The field order of the input file is defined by the Record Definition File used for grouping; there is no significance to the order in which the fields are listed below.

Grouper Field Name	DD Field Name	Notes
RENALMOD	RENAL TREATMENT MODALITY CODE	Leading zeros are significant
RENALSITE	RENAL TREATMENT PRIMARY SUPERVISION CODE	Leading zeros are significant
RENALACCESS	RENAL DIALYSIS ACCESS TYPE	Leading zeros are significant
HBV	HEPATITIS B ANTIGEN STATUS (RENAL CARE)	Accepted values are NEG, POS and UNK
HCV	HEPATITIS C ANTIBODY STATUS (RENAL CARE)	Accepted values are NEG, POS and UNK
HIV	HUMAN IMMUNODEFICIENCY VIRUS STATUS (RENAL CARE)	Accepted values are NEG, POS and UNK
AGE	Derived	The age of the patient in whole years at the start date of the session Range: 0-130

Each row of the input file represents either one haemodialysis session or one day of peritoneal dialysis.

Adult Critical Care (ACC)

The field list below shows the fields required by the Grouper. The field order of the input file is defined by the Record Definition File used for grouping; there is no significance to the order in which the fields are listed below.

Grouper Field Name	CDS/DD Field Name	Notes
CCUF	CRITICAL CARE UNIT FUNCTION	Leading zeros are significant
BCSD	BASIC CARDIOVASCULAR SUPPORT DAYS	Range: 0-99999
ACSD	ADVANCED CARDIOVASCULAR SUPPORT DAYS	Range: 0-99999
BRSD	BASIC RESPIRATORY SUPPORT DAYS	Range: 0-99999
ARSD	ADVANCED RESPIRATORY SUPPORT DAYS	Range: 0-99999
RSD	RENAL SUPPORT DAYS	Range: 0-99999
NSD	NEUROLOGICAL SUPPORT DAYS	Range: 0-99999
DSD	DERMATOLOGICAL SUPPORT DAYS	Range: 0-99999
LSD	LIVER SUPPORT DAYS	Range: 0-99999
CCL2D	CRITICAL CARE LEVEL 2 DAYS	Range: 0-99999
CCL3D	CRITICAL CARE LEVEL 3 DAYS	Range: 0-99999
CC_Start_Date	CRITICAL CARE START DATE	Format is YYYYMMDD. e.g. 14 March 2016 = 20160314
CC_Discharge_Date	CRITICAL CARE DISCHARGE DATE	Format is YYYYMMDD. e.g. 14 March 2016 = 20160314

Critical Care Start Date and Critical Care Discharge Date fields are used to calculate critical care days in the Grouper output file. They are not used in HRG derivation.

Each row of the input file represents one Adult Critical Care Period.

Paediatric Critical Care (PCC)

The Grouper sorts Paediatric Critical Care data prior to grouping so that records with the same provider code and local identifier are placed in activity date order.

The field list below shows the fields required by the Grouper. The field order of the input file is defined by the Record Definition File used for grouping; there is no significance to the order in which the fields are listed below.

Grouper Field Name	CDS/DD Field Name	Notes
PROCODET	ORGANISATION CODE (CODE OF PROVIDER)	A value must be supplied but is not validated. For NHS organisations use the first 3 characters for non-NHS organisations use all 5 characters
CCLocalID	CRITICAL CARE LOCAL IDENTIFIER	This and the provider field are the key that keeps records for the same patient together. They must be supplied
CCDate	ACTIVITY DATE (CRITICAL CARE)	Format is YYYYMMDD. e.g. 14 March 2016 = 20160314
DISDATE	CRITICAL CARE DISCHARGE DATE	Format is YYYYMMDD. e.g. 14 March 2016 = 20160314
DISMETH	DISCHARGE METHOD CODE (HOSPITAL PROVIDER SPELL)	
CCUF	CRITICAL CARE UNIT FUNCTION	Leading zeros are significant
CCAC_01	CRITICAL CARE ACTIVITY CODE	Valid CCAC code from list in Paediatric Critical Care Minimum Dataset (PCCMDS). If this field is blank an error will be generated
CCAC_02 - CCAC_20	CRITICAL CARE ACTIVITY CODE	May be blank
OPER_01 - OPER_20	HIGH COST DRUGS (OPCS)	Valid OPCS-4 codes or blank. The PCC MDS specifies two appropriate procedure codes only
DIAG_01 - DIAG_99	PRIMARY DIAGNOSIS (ICD) and SECONDARY DIAGNOSES (ICD)	Valid ICD-10 codes or blank

Each paediatric critical care day is represented by a data row in the input file. Where a child moves between units with different Critical Care Unit Function Codes, a new critical care period starts; this may result in the generation of more than one critical care record for the day of transfer and consequently more than one HRG for that day.

Neonatal Critical Care (NCC)

The Grouper sorts Neonatal Critical Care prior to grouping so that records with the same provider code and local identifier are placed in activity date order.

The field list below shows the fields required by the Grouper. The field order of the input file is defined by the Record Definition File used for grouping; there is no significance to the order in which the fields are listed below.

Grouper Field Name	CDS/DD Field Name	Notes
PROCODET	ORGANISATION CODE (CODE OF PROVIDER)	A value must be supplied but is not validated. For NHS organisations use the first 3 characters for non-NHS organisations use all 5 characters
CCLocalID	CRITICAL CARE LOCAL IDENTIFIER	This and the provider field are the key that keeps records for the same patient together. They must be supplied
CCDate	ACTIVITY DATE (CRITICAL CARE)	Format is YYYYMMDD. e.g. 14 March 2016 = 20160314
DISDATE	CRITICAL CARE DISCHARGE DATE	Format is YYYYMMDD. e.g. 14 March 2016 = 20160314
CCUF	CRITICAL CARE UNIT FUNCTION	Leading zeros are significant
AGE_DAYS	Derived	Whole days, rounded down: ACTIVITY DATE (CRITICAL CARE) - PERSON BIRTH DATE
DISMETH	DISCHARGE METHOD CODE (HOSPITAL PROVIDER SPELL)	
GestLen	GESTATION LENGTH (AT DELIVERY)	
PERWT	PERSON WEIGHT	Kilograms, to 3 decimal places. Range: greater than zero kg and less than 10 kg. Leading zeros are accepted
CCAC_01	CRITICAL CARE ACTIVITY CODE	Valid CCAC code from list in Neonatal Critical Care Minimum Dataset (NCCMDS). If this field is blank an error will be generated
CCAC_02 - CCAC_20	CRITICAL CARE ACTIVITY CODE	May be blank

Each neonatal critical care day is represented by a data row in the input file. Where a baby moves between units with different Critical Care Unit Function Codes, a new critical care period starts; this may result in the generation of more than one critical care record for the day of transfer and consequently more than one HRG for that day.

Output Files

In the text below and the tables in the following pages, the output file name supplied by the user during processing is referred to as `[name]`; the remainder of the file name is a standard suffix that is appended by the Grouper.

A single input file produces a number of output files. User requirements determine which files are used; some users may not require certain output files. In most cases an output field appears in more than one file thus enabling users to select the files that best suit their needs.

Output files are produced as comma-separated text. These files may be opened with Microsoft Excel but very large files may exceed the maximum number of rows for Excel. In these cases an error message such as “File not loaded completely” will be displayed when attempting to open the file. The output files may be opened with a variety of applications, including Windows Notepad.

Relational Outputs

Relational output files can be identified by the presence of “`_rel`” as part of the filename. They are characterised by the following features:

- They include row number references in order to allow files to be linked
- Where items are repeated (e.g. unbundled HRGs or error messages) this is represented by the addition of rows rather than columns, i.e. the data is normalised

The relational output files are supplied to support users who wish to import the Grouper output into a relational database; other users may choose to ignore them, and vice versa.

RowNo Field

RowNo is an identifier that can be used to link rows in output files to rows in other output files. This is particularly useful in the case of the relational output files. Use of a system generated row number avoids problems using input values where they may not be unique. The **RowNo** values in the `[name]_spell_rel.csv` file may contain non-consecutive values where the input data contains multi-episode spells.

Iteration Field

Iteration appears in some relational output files where it is used to distinguish between occurrences within the same key value. There is no significance to the numerical values of ‘Iteration’; the values are assigned in order, being numbered 1, 2, 3 etc.

Admitted Patient Care (APC)

There are thirteen output files.

File Name/Field Name		**Description**
[name].csv		A list of all the other output files
[name]_sort.csv		A copy of the input data after it has been sorted by Provider Code, Spell Number and Episode Number prior to grouping
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	The generated row number of the record after sorting. This will match the equivalent fields in other output files from the same grouper run
[name]_FCE.csv		Contains both episode and spell output fields. The spell fields are repeated for each episode in the spell; care must be taken to avoid double-counting when using spell fields
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	FCE_HRG	The episode HRG
	GroupingMethodFlag	The grouping method used for the spell: U=Error, M=Multiple trauma, B=Burns-driven, P=Procedure-driven, G=Global exception, D=Diagnosis-driven
	DominantProcedure	The dominant procedure
	FCE_PBC	Programme Budgeting Code for the episode
	CalcEpidur	The calculated episode duration. This is the input Episode duration minus the number of days in the input fields CRITICALCAREDAYS, REHABILITATIONDAYS and SPCDAYS. If the sum of these deductions is larger than the episode duration, then CalcEpidur will be zero
	ReportingEPIDUR	Only populated in Reference Costs Groupers. For ordinary admissions (Patient Classification = 1) if CalcEpidur is zero days ReportingEPIDUR is set to one day. Otherwise ReportingEPIDUR is the same as CalcEpidur
	FCETrimpoint	Only populated in Reference Costs Groupers. The trimpoint for the episode HRG
	FCEExcessBeddays	Only populated in Reference Costs Groupers. The number of days by which CalcEpidur exceeds the trimpoint
	SpellReportFlag	Populated with 1 if the episode contains the grouping variable used for deriving the spell HRG (Dominant Episode) or 0 for other episodes in the spell
	FCESSC_Ct	A count of derived SSCs at episode level

File Name/Field Name		**Description**
	FCESSC1 to FCESSC7	SSCs derived at episode level
	SpellHRG	The spell HRG
	SpellGroupingMethodFlag	The grouping method used for the spell. P=Procedure driven, D=Diagnosis driven, M=Multiple trauma, B=Burns driven, G=Global exception, U=Error
	SpellDominantProcedure	The dominant procedure for the spell
	SpellPDiag	The primary diagnosis used when spell grouping
	SpellSDiag	The first secondary diagnosis used when spell grouping
	SpellEpisodeCount	The number of episodes in the spell
	SpellLOS	The spell duration used for grouping. This is the total of the CalcEpidur fields for the episodes in the spell
	ReportingSpellLOS	Only populated in Reference Costs Groupers. The total of the ReportingEPIDUR fields for the episodes in the spell
	SpellTrimpoint	Only populated in Reference Costs Groupers. The trimpoint for the spell HRG
	SpellExcessBeddays	Only populated in Reference Costs Groupers. The number of days by which the SpellLOS exceeds the trimpoint
	SpellCCDays	The number of critical care days in the spell
	SpellPBC	Programme Budgeting Code for the spell
	UnbundledHRGs	Variable number of fields containing unbundled HRGs (episode level). There is no significance to the order in which unbundled HRGs are output. For Specialist Palliative Care and Rehabilitation unbundled HRGs, where there is more than one daily HRG the field consists of an HRG followed by an asterisk followed by the number of days, rather than HRG being repeated for every day. For example VC18Z*20 indicates VC18Z for 20 days
[name]_spell.csv		Contains one row per spell
	RowNo	Matches the run generated row number
	PROCODET	The organisation code from the input file
	PROVSPNO	The hospital provider spell number from the input file
	SpellHRG	The spell HRG
	SpellGroupingMethodFlag	The grouping method used for the spell: U=Error, M=Multiple trauma, B=Burns-driven, P=Procedure-driven, G=Global exception, D=Diagnosis-driven
	SpellDominantProcedure	The dominant procedure for the spell

File Name/Field Name		**Description**
	SpellPDiag	The primary diagnosis used for spell grouping
	SpellSDiag	The first secondary diagnosis used for spell grouping
	SpellEpisodeCount	The number of episodes in the spell
	SpellLOS	The spell duration used for grouping
	ReportingSpellLOS	Only populated in Reference Costs Groupers. The total of the ReportingEPIDUR fields for the episodes in the spell
	SpellTrimpoint	Only populated in Reference Costs Groupers. The trimpoint for the spell HRG
	SpellExcessBeddays	Only populated in Reference Costs Groupers. The number of days by which the SpellLOS exceeds the trimpoint
	SpellCCDays	The number of critical care days in the spell
	SpellPBC	Programme Budgeting Code for the spell
	SpellSSC_Ct	Number of distinct SSCs produced for the spell
	SpellSSC1 to SpellSSC7	Only populated in Local Payment Groupers. Candidate Specialised Service Codes for the spell
	SpellBP_Ct	Only populated in Local Payment Groupers. Number of distinct BPTs produced for the spell
	SpellBP1 to SpellBP7	Only populated in Local Payment Groupers. Candidate Best Practice Tariff codes for the spell
	SpellFlag_Ct	Only populated in Local Payment Groupers. Number of distinct other flags produced for the spell
	SpellFlag1 to SpellFlag7	Only populated in Local Payment Groupers. Other flags for the spell
	UnbundledHRGs	Variable number of fields containing unbundled HRGs (spell level). There is no significance to the order in which unbundled HRGs are output. For Specialist Palliative Care and Rehabilitation unbundled HRGs, where there is more than one daily HRG the field consists of an HRG followed by an asterisk followed by the number of days, rather than HRG being repeated for every day. For example VC18Z*20 indicates VC18Z for 20 days
[name]_quality.csv		Contains a row for each episode that contains an error. Where one or more episodes within a multi-episode spell contain errors, all the episodes from the spell, including those that do not contain errors, are included in the quality file. Thus there are records in the quality file that do not include error messages
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number

File Name/Field Name		**Description**
	Error Message	A variable number of fields, each consisting of Code Type, Code and Error Message, separated by pipe (|) symbols
[name]_FCE_rel.csv		Episode-level output in relational form
	RowNo	Matches the run generated row number
	FCE_HRG	The episode HRG
	GroupingMethodFlag	The grouping method used for the spell: U=Error, M=Multiple trauma, B=Burns-driven, P=Procedure-driven, G=Global exception, D=Diagnosis-driven
	DominantProcedure	The dominant procedure
	FCE_PBC	Programme Budgeting Code for the episode
	CalcEpidur	The calculated episode duration. This is the input Episode duration minus the number of days in the input fields CRITICALCAREDAYS, REHABILITATIONDAYS and SPCDAYS
	ReportingEPIDUR	Only populated in Reference Costs Groupers. For ordinary admissions (Patient Classification = 1) if CalcEpidur is zero days ReportingEPIDUR is set to one day. Otherwise ReportingEPIDUR is the same as CalcEpidur
	FCETrimpoint	Only populated in Reference Costs Groupers. The trimpoint for the episode HRG
	FCEExcessBeddays	Only populated in Reference Costs Groupers. The number of days by which CalcEpidur exceeds the trimpoint
	SpellReportFlag	Identifies the dominant episode
[name]_spell_rel.csv		Spell-level output in relational form; one row per spell
	RowNo	Matches the run generated row number
	PROCODET	The organisation code from the input file
	PROVSPNO	The hospital provider spell number from the input file
	SpellHRG	The spell HRG
	SpellGroupingMethod Flag	The grouping method used for the spell: U=Error, M=Multiple trauma, B=Burns-driven, P=Procedure-driven, G=Global exception, D=Diagnosis-driven
	SpellDominant Procedure	The dominant procedure for the spell
	SpellPDiag	The primary diagnosis used when spell grouping
	SpellSDiag	The first secondary diagnosis used when spell grouping
	SpellEpisodeCount	The number of episodes in the spell
	SpellLOS	The spell duration used for grouping
	ReportingSpellLOS	Only populated in Reference Costs Groupers. The total of the ReportingEPIDUR fields for the episodes in the spell

File Name/Field Name		**Description**
	SpellTrimpoint	Only populated in Reference Costs Groupers. The trimpoint for the spell HRG
	SpellExcessBeddays	Only populated in Reference Costs Groupers. The number of days by which the SpellLOS exceeds the trimpoint
	SpellCCDays	The number of critical care days in the spell
	SpellPBC	Programme Budgeting Code for the spell
[name]_quality_rel.csv		Relational format includes a row for each error for each episode
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the quality message
	Code Type	The type of code that has failed validation
	Code	The value of the code that has failed validation. Blank where the code is missing from the input data
	Error Message	Description of the error
[name]_flag_rel.csv		This file is populated by Payment Groupers only
	RowNo	Matches the run generated row number
	PROCODET	The organisation code from the input file
	PROVSPNO	The hospital provider spell number from the input file
	Iteration	The ordinal number of the SSC, BPT or other flag
	SpellFlag	Includes all distinct Specialised Service Codes (SSCs), Best Practice Tariff (BPT) codes and other flags generated for the spell
[name]_FCE_flag_rel.csv		SSC output in relational form
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the SSC
	FCEFlag	Includes all distinct Specialised Service Codes (SSCs)
[name]_ub_rel.csv		The unbundled HRGs. There are no entries for episodes that do not have any unbundled HRGs
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the unbundled HRG
	UnbundledHRGs	The unbundled HRGs. There is no significance to the order in which unbundled HRGs are output. For Specialist Palliative Care and Rehabilitation, HRGs are generated on a per diem basis, so there will be a row for each day
[name]_summary.csv		A single-row file containing details about the grouping session
	Grouper Version	Version of the Grouper that produced the output files
	Database Version	The Grouper's internal HRG database version
	FCE Count	The number of episodes submitted

File Name/Field Name		**Description**
	Spell Count	The number of spells submitted
	FCE Error Count	The number of episodes having errors
	Spell Error Count	The number of spells having errors
	Run Start Date/Time	The date and time that the grouping session started
	Run End Date/Time	The date and time that the grouping session finished
	Input Filename	The path and filename of the input file
	Output Filename	The path and filename selected by the user
	RDF path and name	The path and filename of the Record Definition File used for grouping
[name]_report.csv		This file is populated by Reference Costs Groupers only. An aggregated report file
	Type	Contains one of three values: APC-E – Episode level aggregates, APC-S – Spell level aggregates, APC-U – Unbundled HRG aggregates
	Provider	The PROCEDET input field
	HRG	The FCE_HRG, SpellHRG or UB_HRG output fields as appropriate for the Type
	TretSpef	The TRETSPEF input field. For APC-S type it is the TRETSPEF of the dominant episode in the spell
	ClassPat	The CLASSPAT input field. For APC-S type it is the CLASSPAT of the first episode in the spell
	AdmiMeth	The ADMIMETH input field. For APC-S type it is the ADMIMETH of the first episode in the spell
	TrimPoint	
	Mean LOS	The mean of the input EPIDUR
	Adj Mean LOS	The mean of the output CalcEPIDUR
	Inlier Beddays	The sum of the output CalcEPIDUR or SpellLOS minus the output FCEExcessBedDays or SpellExcessBeddays as appropriate
	Excess Beddays	The sum of the output FCEExcessBeddays or SpellExcessBeddays as appropriate
	Count	The count of episodes, spells or records

Non-Admitted Consultations (NAC)

There are nine output files.

File Name/Field Name		**Description**
[name].csv		A list of other output files
[name]_attend.csv		The main grouped output file
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	The generated row number of the record
	NAC_HRG	The HRG
	GroupingMethodFlag	Grouping method used: U=Error, P=Procedure-driven, G=Global exception, O=Outpatient default
	DominantProcedure	The dominant procedure
	AttendanceHRG	Only populated in Local Payment Groupers. An alternative attendance HRG. This may be used for tariffing purposes when the NAC_HRG does not have a mandatory tariff, and is equivalent to the SUS_HRG in the National System
	AttendSSC_Ct	Only populated in Local Payment Groupers. The count of distinct SSCs produced for the attendance
	AttendSSC1 to AttendSSC5	Not currently populated
	AttendBP_Ct	Only populated in Local Payment Groupers. The count of distinct BPT flags produced for the attendance
	AttendBP1 to AttendBP5	Not currently populated
	AttendFlag_Ct	The count of distinct other flags produced for the attendance
	AttendFlag1 to AttendFlag5	Not currently populated
	UnbundledHRGs	A variable number of fields containing unbundled HRGs appended to the end of each record. There is no significance to the order in which unbundled HRGs are output
[name]_quality.csv		Contains a row for each attendance that contains errors
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	Error Message	A variable number of fields, consisting of Code Type, Code and Error Message, separated by pipe (|) symbols
[name]_attend_rel.csv		Output in relational form
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	NAC_HRG	The attendance HRG

File Name/Field Name		**Description**
	GroupingMethodFlag	Grouping method used: U=Error, P=Procedure-driven, G=Global exception, O=Outpatient default
	DominantProcedure	The dominant procedure
	AttendanceHRG	Only populated in Payment Groupers. An alternative attendance HRG. This may be used for tariffing purposes when the NAC_HRG does not have a mandatory tariff, and is equivalent to the SUS_HRG in the National System
[name]_quality_rel.csv		All error messages in relational form
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the quality message
	Code Type	The type of code that has failed validation
	Code	The code that failed validation. Blank if code missing from input data
	Error Message	Description of the error
[name]_flag_rel.csv		This file is not currently populated
	RowNo	
	Iteration	
	AttendFlag	
[name]_ub_rel.csv		The unbundled HRGs. There are no entries for attendances without unbundled HRGs
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the unbundled HRG
	UnbundledHRGs	The unbundled HRGs. There is no significance to the output order
[name]_summary.csv		A single-row file containing details about the grouping session
	Grouper Version	The version of the Grouper that produced the output files
	Database Version	The Grouper's internal HRG database version
	Attendance Count	The number of records submitted
	Attendance Error Count	The number of records having errors
	Run Start Date/Time	The date and time that the grouping session started
	Run End Date/Time	The date and time that the grouping session finished
	Input Filename	The path and filename of the input file
	Output Filename	The path and filename selected by the user
	RDF path and name	The path and filename of the Record Definition File used for grouping
[name]_report.csv		This file is populated by Reference Costs Groupers only. An aggregated report file

File Name/Field Name		**Description**
	Type	Contains one of two values: NAC – Attendance aggregates, NAC-U – Unbundled HRG aggregates
	Provider	Field not populated
	HRG	The NAC_HRG or UB_HRG output fields as appropriate for the Type
	TretSpef	The TRETSPEF input field
	ClassPat	Field not populated
	AdmiMeth	Field not populated
	TrimPoint	Field not populated
	Mean LOS	Field not populated
	Adj Mean LOS	Field not populated
	Inlier Beddays	Field not populated
	Excess Beddays	Field not populated
	Count	The count of attendances or records

Emergency Medicine (EM)

There are six output files.

File Name/Field Name		**Description**
[name].csv		A list of other output files
[name]_attend.csv		The main grouped output file
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	The generated row number of the record
	EM_HRG	The attendance HRG
[name]_quality.csv		Contains a row for each input record that contains errors
	<input data>	All of the input data including any non-mandatory fields
	RowNo	Matches the run generated row number
	Error Message	A variable number of fields, consisting of Code Type, Code and Error Message, separated by pipe (|) symbols
[name]_quality_rel.csv		All error messages in relational form
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the quality message
	Code Type	The type of code that has failed validation
	Code	The code that failed validation. Blank if code missing from input data
	Error Message	Description of the error
[name]_summary.csv		A single-row file containing details about the grouping session
	Grouper Version	The version of the Grouper that produced the output files
	Database Version	The Grouper's internal HRG database version
	Attendance Count	The number of records submitted
	Attendance Error Count	The number of records having errors
	Run Start Date/Time	The date and time that the grouping session started
	Run End Date/Time	The date and time that the grouping session finished
	Input Filename	The path and filename of the input file
	Output Filename	The path and filename selected by the user
	RDF path and name	The path and filename of the Record Definition File used for grouping
[name]_report.csv		This file is populated by Reference Costs Groupers only. An aggregated report file
	Type	Contains the value EM
	Provider	Field not populated
	HRG	The EM_HRG
	TretSpef	Field not populated
	ClassPat	Field not populated

File Name/Field Name		**Description**
	AdmiMeth	Field not populated
	TrimPoint	Field not populated
	Mean LOS	Field not populated
	Adj Mean LOS	Field not populated
	Inlier Beddays	Field not populated
	Excess Beddays	Field not populated
	Count	The count of attendances

Renal Dialysis (NRD)

There are six output files

File Name/Field Name		**Description**
[name].csv		A list of other output files
[name]_renal.csv		The main grouped output file
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	The generated row number of the record
	NRD_HRG	The HRG for the dialysis record
[name]_quality.csv		Contains one row for each record that contains errors
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	Error Message	A variable number of fields, consisting of Code Type, Code and Error Message, separated by pipe (|) symbols
[name]_quality_rel.csv		All error messages in relational form
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the quality message
	Code Type	The type of code that has failed validation
	Code	The code that failed validation. Blank if code missing from input data
	Error Message	Description of the error
[name]_summary.csv		A single-row file containing details about the grouping session
	Grouper Version	The version of the Grouper that produced the output files
	Database Version	The Grouper's internal HRG database version
	NRD Record Count	The number of records submitted
	NRD Record Error Count	The number of records having errors
	Run Start Date/Time	The date and time that the grouping session started
	Run End Date/Time	The date and time that the grouping session finished
	Input Filename	The path and filename of the input file
	Output Filename	The path and filename selected by the user
	RDF path and name	The path and filename of the Record Definition File used for grouping
[name]_report.csv		This file is populated by Reference Costs Groupers only. An aggregated report file
	Type	Contains the value NRD.
	Provider	Field not populated
	HRG	The NRD_HRG
	TretSpef	Field not populated

File Name/Field Name		**Description**
	ClassPat	Field not populated
	AdmiMeth	Field not populated
	TrimPoint	Field not populated
	Mean LOS	Field not populated
	Adj Mean LOS	Field not populated
	Inlier Beddays	Field not populated
	Excess Beddays	Field not populated
	Count	The count of records

Adult Critical Care (ACC)

There are six output files.

File Name/Field Name		**Description**
[name].csv		A list of other output files
[name]_acc.csv		The main grouped output file
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	The generated row number of the record
	ACC_HRG	The unbundled HRG for the Adult Critical Care period
	Calc_CC_Days	The number of Critical Care days calculated as: CC Discharge Date - CC Start Date + 1 This will be set to -1 if there are problems with the dates
	CC_Warning_Flag	Flag to indicate the result of validation of dates and respiratory support day's fields. Failure does not prevent HRG derivation. Blank indicates passing validation. F indicates date validation failure; applied if any of the following are true: Calc_CC_Days = -1. (This indicates that CC Discharge Date is before CC Start Date, or CC Start Date or CC Discharge Date is blank, is not a valid date or does not match the required format.) CCL2 Days + CCL3 Days > Calc_CC_Days ARSD + BRSD > Calc_CC_Days ARSD + BRSD > CCL2 Days + CCL3 Days
[name]_quality.csv		Contains a row for each input record that contains errors
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	Error Message	A variable number of fields, consisting of Code Type, Code and Error Message, separated by pipe (|) symbols
[name]_quality_rel.csv		All error messages in relational form
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the quality message
	Code Type	The type of code that has failed validation
	Code	The code that failed validation. Blank if code missing from input data
	Error Message	Description of the error
[name]_summary.csv		A single-row file containing details about the grouping session
	Grouper Version	The version of the Grouper that produced the output files
	Database Version	The Grouper's internal HRG database version
	ACC Period Count	The number of records submitted

File Name/Field Name		**Description**
	ACC Period Error Count	The number of records having errors
	Run Start Date/Time	The date and time that the grouping session started
	Run End Date/Time	The date and time that the grouping session finished
	Input Filename	The path and filename of the input file
	Output Filename	The path and filename selected by the user
	RDF path and name	The path and filename of the Record Definition File used for grouping
[name]_report.csv		This file is populated by Reference Costs Groupers only. An aggregated report file
	Type	Contains the value ACC
	Provider	Field not populated
	HRG	The ACC_HRG
	TretSpef	Field not populated
	ClassPat	Field not populated
	AdmiMeth	Field not populated
	TrimPoint	Field not populated
	CC Days	The output Calc_CC_Days field
	Adj Mean LOS	Field not populated
	Inlier Beddays	Field not populated
	Excess Beddays	Field not populated
	Count	The count of periods

Paediatric Critical Care (PCC)

There are seven output files.

File Name/Field Name		**Description**
[name].csv		A list of other output files
[name]_sort.csv		A copy of the input data after it has been sorted
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	The generated row number of the record after sorting
[name]_pcc.csv		The main grouped output file
	<input data>	All of the input data including any non-mandatory fields
	RowNo	Matches the run generated row number
	PCC_HRG	The unbundled HRG for the Paediatric Critical Care day
[name]_quality.csv		Contains one row for each record that contains errors
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	Error Message	A variable number of fields, consisting of Code Type, Code and Error Message, separated by pipe (|) symbols
[name]_quality_rel.csv		All error messages in relational form
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the quality message
	Code Type	The type of code that has failed validation
	Code	The code that failed validation. Blank if code missing from input data
	Error Message	Description of the error
[name]_summary.csv		A single-row file containing details about the grouping session
	Grouper Version	The version of the Grouper that produced the output files
	Database Version	The Grouper's internal HRG database version
	PCC Record Count	The number of records submitted
	PCC Record Error Count	The number of records having errors
	Run Start Date/Time	The date and time that the grouping session started
	Run End Date/Time	The date and time that the grouping session finished
	Input Filename	The path and filename of the input file
	Output Filename	The path and filename selected by the user
	RDF path and name	The path and filename of the Record Definition File used for grouping
[name]_report.csv		This file is populated by Reference Costs Groupers only. An aggregated report file
	Type	Contains the value PCC

File Name/Field Name		**Description**
	Provider	The PROCEDET input field
	HRG	The PCC_HRG
	TretSpef	Field not populated
	ClassPat	Field not populated
	AdmiMeth	Field not populated
	TrimPoint	Field not populated
	CC Days	Always 1, for single day PCC records
	Adj Mean LOS	Field not populated
	Inlier Beddays	Field not populated
	Excess Beddays	Field not populated
	Count	The count of records

Neonatal Critical Care (NCC)

There are seven output files.

File Name/Field Name		**Description**
[name].csv		A list of other output files
[name]_sort.csv		A copy of the input data after it has been sorted
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Generated row number of the record after sorting
[name]_ncc.csv		The main grouped output file
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	NCC_HRG	The unbundled HRG for the Neonatal Critical Care day
[name]_quality.csv		Contains one row for each record that contains errors
	<input data>	All of the input data, including any non-mandatory fields
	RowNo	Matches the run generated row number
	Error Message	A variable number of fields, consisting of Code Type, Code and Error Message, separated by pipe (|) symbols
[name]_quality_rel.csv		All error messages in relational form
	RowNo	Matches the run generated row number
	Iteration	The ordinal number of the quality message
	Code Type	The type of code that has failed validation
	Code	The code that failed validation. Blank if code missing from input data
	Error Message	Description of the error
[name]_summary.csv		A single-row file containing details about the grouping session
	Grouper Version	The version of the Grouper that produced the output files
	Database Version	The Grouper's internal HRG database version
	NCC Record Count	The number of records submitted
	NCC Record Error Count	The number of records having errors
	Run Start Date/Time	The date and time that the grouping session started
	Run End Date/Time	The date and time that the grouping session finished
	Input Filename	The path and filename of the input file
	Output Filename	The path and filename selected by the user
	RDF path and name	The path and filename of the Record Definition File used for grouping
[name]_report.csv		This file is populated by Reference Costs Groupers only. An aggregated report file
	Type	Contains the value NCC

File Name/Field Name		**Description**
	Provider	The PROCEDET input field
	HRG	The NCC_HRG
	TretSpef	Field not populated
	ClassPat	Field not populated
	AdmiMeth	Field not populated
	TrimPoint	Field not populated
	CC Days	Always 1, for single day NCC records
	Adj Mean LOS	Field not populated
	Inlier Beddays	Field not populated
	Excess Beddays	Field not populated
	Count	The count of records

Errors and Validation

As part of the grouping process the Grouper carries out validation checks on many of the mandatory input fields. Where one or more fields fail validation the Grouper derives the HRG UZ01Z (Data Invalid for Grouping).

For single-episode spells, where the episode fails validation, the error HRG is derived at both episode and spell level.

For multi-episode spells that contain a mixture of episodes that pass validation and episodes that fail validation, the error HRG is derived for each episode that fails validation and at spell level.

Details of all errors are reported in the output quality file. The listings in the Output Files section of this document include a description of the quality file for each of the data set types processed by the Grouper. The Input File Preparation section provides further information about valid values for various input fields.

Details of clinical coding validation can be found in the chapter summary documentation for chapter **UZ Undefined Groups** which is available for download from the Casemix website.

Error Message Format

Error messages in the quality file are displayed in three sections, separated by a pipe character:

`Code Type|Code|Error Message`

`Code Type` identifies the field or field type

`Code` is the unrecognised value. Where an error is due to the absence of a code, this section is left blank.

`Error Description` explains the nature of the error

An example error message is:

`DIAG_01|P102|UZ03 - Diagnosis Conflicts with Age`

Error Types

There are three types of errors:

Field Validation Errors are generated where field values are missing or are outside of the accepted range of values. Example error messages of this type are:

`DIAG_01||Primary Diagnosis is blank DIAG_01|R69X6|Primary Diagnosis is invalid`

`OPER_02|C992|Procedure is invalid`

Spell Validation Errors result from cross checks across records in a spell. For example, where consistency checks for sex or age fail:

`SEX||Sex is inconsistent in spell`

`STARTAGE||Age rises by more than expected in the spell`

Clinical Coding Errors are mainly due to the use of clinical codes that violate clinical coding convention or are not useful resource indicators for HRG grouping.

Clinical coding errors produce error messages that contain the error category codes described in the following section.

Error Categories

Error Categories are used mainly for clinical coding errors. Critical care grouping error descriptions also include error category codes.

UZ01 Invalid Primary Diagnosis

Diagnosis code is present but should not be used in a primary position according to clinical coding conventions.

```
DIAG_01|Z509|UZ01 - Invalid Primary Diagnosis
```

(Z509: Care involving use of rehabilitation procedure, unspecified)

UZ02 Poorly Coded Primary Diagnosis

Diagnosis code is valid as a primary diagnosis but is too vague to determine the resource use.

```
DIAG_01|T140|UZ02 - Poorly Coded Primary Diagnosis
```

(T140: Superficial injury of unspecified body region)

UZ03 Diagnosis Conflicts with Age

A paediatric primary diagnosis has been recorded for an adult patient (age 19 years and over).

```
DIAG_01|P704|UZ03 - Diagnosis Conflicts with Age
```

(P704: Other neonatal hypoglycaemia)

UZ04 Diagnosis Conflicts with Anatomical Site

Indicates an invalid combination of primary diagnosis and anatomical site. This only applies to specific musculoskeletal codes entered at 5th digit level.

```
DIAG_01|M7217|UZ04 - Diagnosis Conflicts with Anatomical Site
```

(M7217: Knuckle pads: Ankle and foot)

UZ05 Invalid procedure for Casemix grouping purposes

Indicates invalid dominant procedure (e.g. an anatomical site) and includes the entire Y (methods of operations) and Z (anatomical sites) codes and also a number of codes in the main body system chapters.

```
OPER_02|Y841|UZ05 - Invalid procedure for Casemix grouping purposes
```

`OPER_02|W450|UZ05 - Invalid procedure for Casemix grouping purposes`

(Y841: Gas and air analgesia in labour)

(W540: Conversion from previous prosthetic replacement of articulation of bone NEC)

UZ06 Poorly coded procedure for Casemix grouping purposes

Indicates a dominant or unbundled procedure that is too vague and unspecific to determine resource use from an HRG design perspective.

`OPER_02|A579|UZ06 - Poorly coded procedure for Casemix grouping purposes`

`OPER_01|U019|UZ06 - Poorly coded procedure for Casemix grouping purposes`

(A579: Unspecified operations on spinal nerve root)

(U019: Unspecified diagnostic imaging of whole body)

UZ11 Neonatal Critical Care Error

This is a general purpose grouping error for Neonatal Critical Care, generated when the input record does not meet any of the criteria in the neonatal critical care grouping algorithm.

UZ13 ACC Grouping Error

This is a general purpose grouping error for Adult Critical Care, generated when the input record does not meet any of the criteria in the adult critical care grouping algorithm.

UZ14 Renal (NRD) Error

This is a general error for grouping renal activity using the national renal dataset and is generated when conditions in the grouping algorithm have not been met.

UZ15 Burns Error

Indicates where a burns primary diagnosis code of unspecified body region or total body surface area (TBSA) is recorded, or burns diagnosis code recorded, in any position, with no subsequent TBSA code present.

UZ21 CCAC Inappropriate in NCC

Generated when the Critical Care Activity Code is inappropriate for the Neonatal Critical Care HRG algorithm.

Notes

When one error is found in a record the Grouper does not stop the validation process. The grouping software aims to identify all errors and output them together.

In Admitted Patient Care errors cascade up to the next level. If an episode HRG is UZ01Z then the spell HRG will be UZ01Z. If an unbundled HRG is UZ01Z then both the episode and the spell HRGs will be UZ01Z. Errors do not cascade downwards, so if one episode gets UZ01Z then it will not stop the other episode HRGs from being generated.

In Admitted Patient Care data the primary diagnosis is always validated.

All clinical codes are validated against the Grouper's internal database of codes.

Codes used that are not on this list will result in the generation of a UZ01Z HRG.

ICD10 codes that are not on the list are classified as invalid but will not result in a specific error message. These will be output in the DQ report as follows;

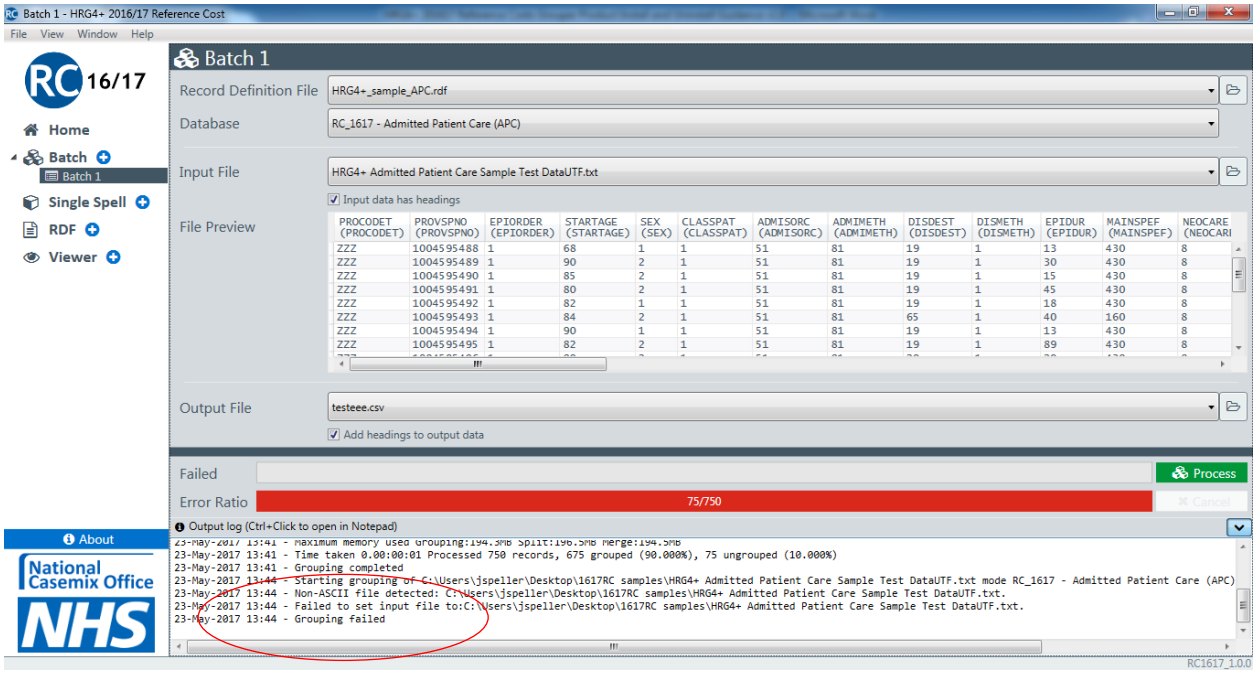
```
DIAG_XX|XXXX|Diagnosis is invalid
```

OPCS codes that are not on the list are similarly classified as invalid but will not result in a specific error message. These will be output in the DQ report as follows;

```
OPER_XX|XXXX|Procedure is invalid
```

Trouble Shooting

If when loading your data you see the following error message – **“failure to set input file”,** this indicates that your file is not in the format currently accepted by the Grouper.



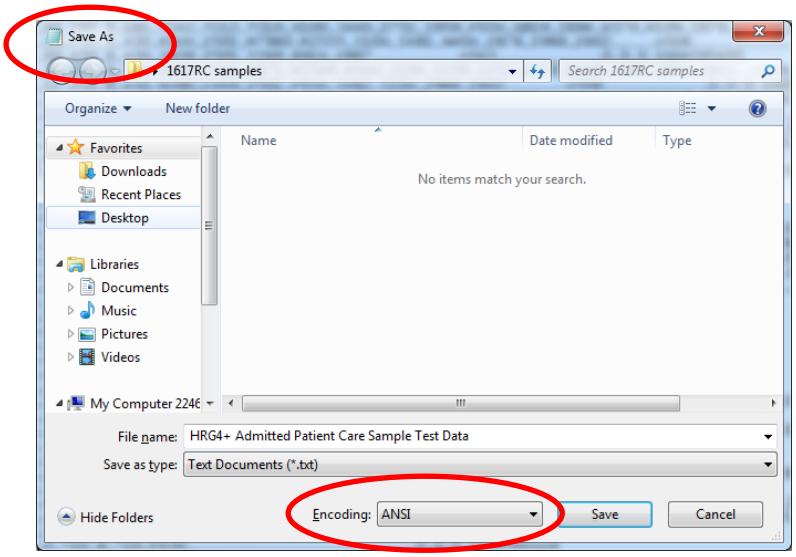
Previous Groupers (pre April 2017) did accept certain characters outside this Character set but the decision was taken to enforce the rule to only accept characters in range. We are currently looking at expanding this requirement for future product releases.

The file preview may very well indicate an issue with the file showing greyed ? marks. This indicates characters not accepted by the current products but this may not always be the case. If you wish to check to see if your file is encoded in an unsupported format this can be done by using Notepad or Excel. These are discussed below.

Workaround

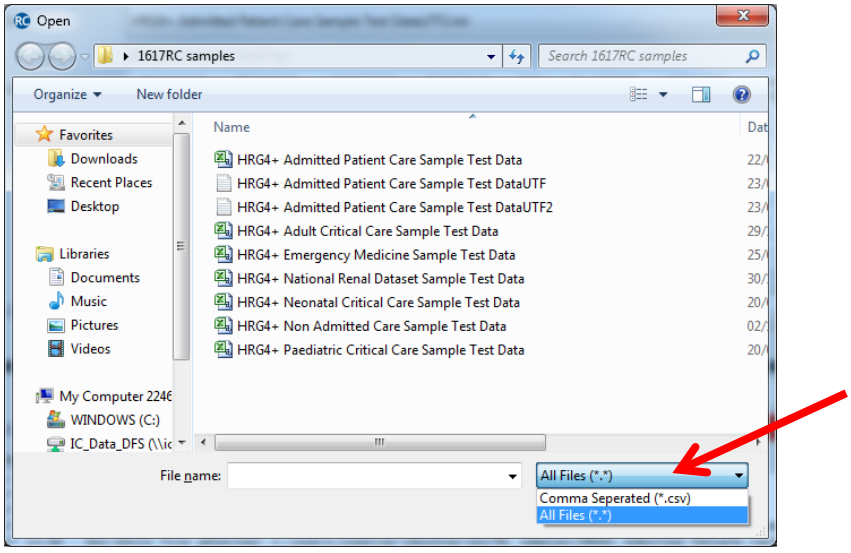
There are no plans to release an updated version of either the LP2017/18 or RC2016/17 Groupers, so simple workarounds will need to be enacted.

If you get this error message you need to either convert your file to be ANSI (ASCII) encoded by opening in Notepad and then saving as an ASCII encoded file.



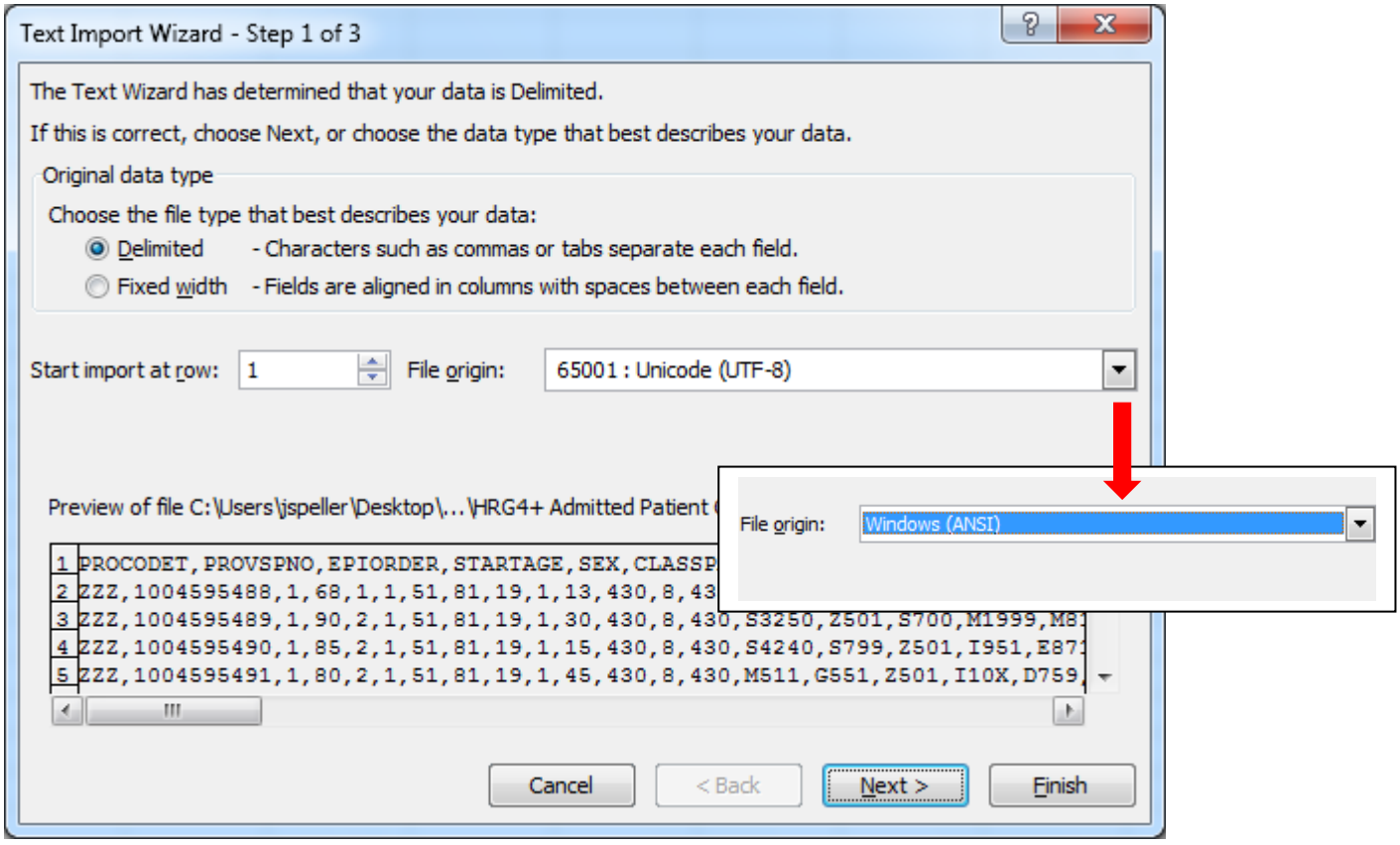
If your file was encoded in other formats the encoding drop down will indicate what this was. Simply choose ANSI and save.

The file will be saved as a **.txt** file, but this can be opened by the current Grouper without conversion to CSV or altering the extension. You will just need to browse for the input file using the **All Files (*.*)** option rather than the default **Comma Separated (*.csv**) when selecting your input when using the Grouper. Alternatively you can select to save as type "**All Files" and add the extension .csv onto the end of the file name.**

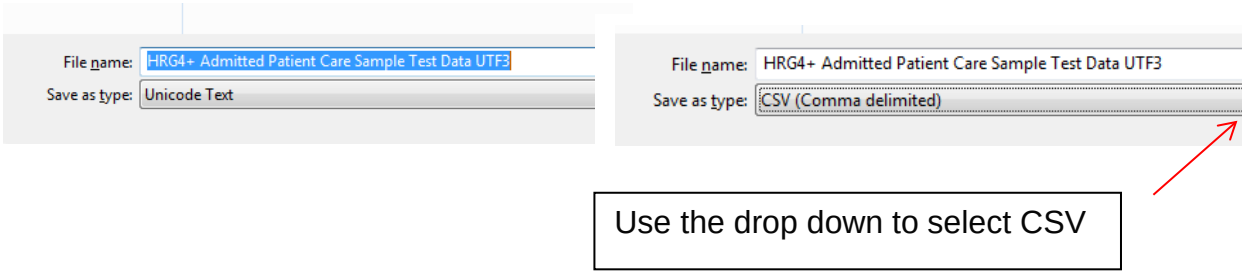


Alternatively the CSV file can be imported into excel using the "**Get external data – From TEXT**" under the **Data** tab – this takes you to the import wizard. Using the import wizard is advised so as to avoid loss of leading zeros within the original file (this is explained in more detail in the Grouper User manual).

This method will also allow you to see what encoding is used within the file. Simply change this to **Windows (ANSI)** and complete the import process. The file then can be saved as a CSV and loaded into the Grouper as normal.



Alternatively you can open the text file using "**open**" and this will take you through the same import steps but will not necessarily show the original encoding of the file as the open function screen defaults to Windows (ANSI). When completing the import, the file is likely to show the following **Save_as** screen. You will need to change the **Save asType** to CSV. The file will then save as an ANSI encoded CSV file and can be loaded as normal into the Grouper.



Use the drop down to select CSV

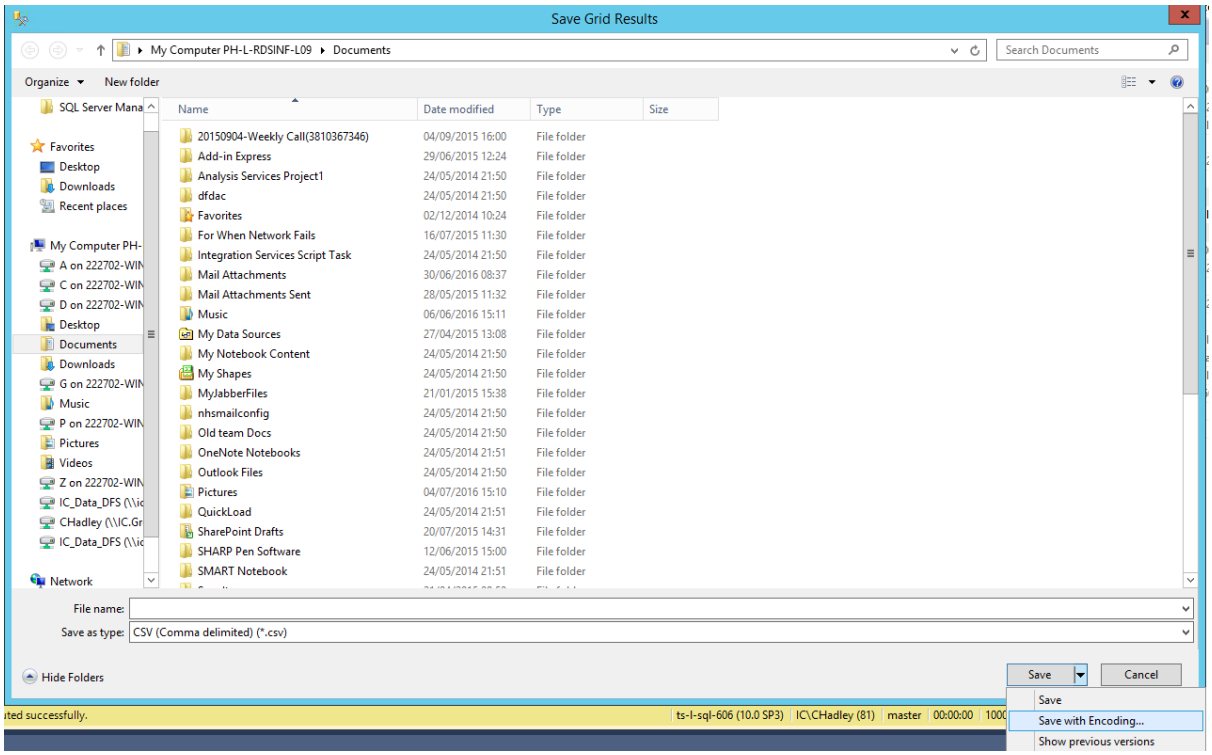
SQL database Extraction

For those users extracting from SQL databases, care should be taken when using the “save as” Option rather than using the export wizard.

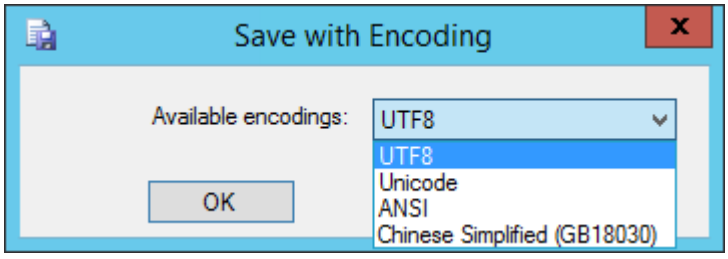
The default setting is set to save using UTF8.

As discussed above, these files cannot be loaded currently.

The workaround is simply that the default setting needs to be changed to ANSI as per the screen shots below.



Takes you to the following screen, where you select ANSI.



This will resolve the issue.

The Documentation Suite

Below is a list of the various documents which are available to download from the National Casemix Office website http://content.digital.nhs.uk/casemix/downloads.

This Documentation Suite provides a comprehensive resource to enable users to understand design concepts and logic, as well as practical use of the Grouper.

- The **Casemix Companion** is a starting point and general reference guide for anyone interested in learning about the casemix classification system used by the NHS in England. The document provides an introduction to HRGs, groupers, HRG4+ design concepts and grouping logic, and it contains links to additional resources
- The **Grouper User Manual** provides instructions on how to prepare and group data using the Grouper software application. Sample data with expected results is provided. This document is updated with every grouper release.
- The **Summary of Changes** document provides an overview of the main differences between the current grouper design and its relevant predecessor.
- The **Chapter Summaries** document provides an overview of the scope, composition and relevant grouping logic of individual HRG subchapters, and highlights significant changes to the latest HRG design.
- The **Code to Group Workbook** is a spreadsheet that embodies the casemix design. It provides details of the constituent elements that contribute to HRG grouping, and it contains reference data such as the ICD-10 and OPCS-4 codes utilised in the design. It contains the procedure and diagnosis hierarchies pertinent to a specific design, and the Complication and Comorbidities lists for HRG subchapters. The spreadsheet also includes information on Programme Budgeting Category (PBC) mapping, as well as a comprehensive list of HRG codes and labels.
- The **Code to Group User Manual** explains how to make best use of the information found in the Code to Group Workbook. Specifically, the manual clarifies the grouping logic found in the workbook's Code to Group tab.
- The **ICD-10 5th Edition Update** outlines the changes to the HRG4+ 2016/17 Reference Costs Grouper made as a result of the 5th Edition update to the International Statistical Classification of Diseases and Related Health Problems – Tenth Revision (ICD-10), which is effective from 1 April 2016.