Methodologies of counting Patient Continuous Inpatient Stays (CIS) using SMR01 inpatient data

<table>
<thead>
<tr>
<th>Document Control</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Version</td>
<td>0.3</td>
</tr>
<tr>
<td>Date Issued</td>
<td></td>
</tr>
<tr>
<td>Author</td>
<td>David Redpath</td>
</tr>
<tr>
<td>Comments to</td>
<td><a href="mailto:david.redpath@nhs.net">david.redpath@nhs.net</a></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Comment</th>
<th>Author</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.1</td>
<td>13/09/2017</td>
<td>Draft following comments from QI team members</td>
<td>David Redpath</td>
</tr>
<tr>
<td>0.2</td>
<td>20/09/2017</td>
<td>Draft following comments from QI team members</td>
<td>David Redpath</td>
</tr>
<tr>
<td>0.3</td>
<td>20/10/2017</td>
<td>Third draft incorporating comments from SAG – mental health CIS section, appendix with syntax. Update of Clinical profiles section following text from Catherine Drury. Inclusion of section on SMR04 CIS.</td>
<td>David Redpath</td>
</tr>
<tr>
<td>0.4</td>
<td>05/02/2018</td>
<td>Minor changes to font and formatting</td>
<td>David Redpath</td>
</tr>
</tbody>
</table>
## Contents

Methodologies and intricacies of counting Patient Continuous Inpatient Stays (CIS) using SMR01 inpatient data ................................................................. 1

INTRODUCTION .............................................................................................................. 3

BACKGROUND ............................................................................................................... 3

CIS DEFINITION and RULES ...................................................................................... 3

USES, METHODS AND THE RATIONALE OF CIS BASED ANALYSIS IN QI PRODUCTS ........................................................................................................ 4

General Rules, Conventions and Guidance on SMR01 data and the Creation of CISs in SMR01 ........................................................................................................ 4

Secondary Care - Information Requests and Parliamentary Questions ..................... 4

Secondary Care - Annual Publication Tables ............................................................... 6

CIS calculations in ACaDMe ....................................................................................... 6

Other Quality Indicators Products ............................................................................. 7

HSMR Publication and Dashboard ............................................................................ 7

Discovery .................................................................................................................... 7

Admissions indicators: ............................................................................................... 8

Elective Admissions per 1,000 population ................................................................. 8

Non Elective Admissions per 1,000 population ......................................................... 8

Elective Bed Days per 1,000 population .................................................................. 8

Non Elective Bed Days per 1,000 population ........................................................... 8

Potentially Preventable Admissions indicators ......................................................... 9

SMR04 Continuous Inpatient Stays ......................................................................... 9

RECOMMENDATIONS ON STANDARD USE WITHIN PHI .................................... 9

APPENDICES .................................................................................................................. 10

1. METHODOLOGIES .................................................................................................. 10

2. SPSS SYNTAX ......................................................................................................... 15

Acute activity publication – Stays by Specialty ....................................................... 15

Acute activity publication – Emergency Admissions ............................................... 19

Clinical Profiles ....................................................................................................... 21

SMR04 SPSS - CIS Syntax ......................................................................................... 24
INTRODUCTION

This paper follows from the previous ‘SAF Bulletin No.16 Special Edition - SMRA Guidance’ giving analysts guidance on the required sort order to be applied the records when being extracted from SMRA to mimic the sort order applied in ACaDMe¹ and that used in the old linked catalogue and flat files.

Following the Secondary Care Team’s(QI SCT) move from Primary and Secondary Care to Quality Indicators(QI), work has been done to assess the differences in methodologies used within QI products. One of the key measures that has been highlighted as having a variety of methodologies used to create it is the patient based Continuous Inpatient(CIS) stay. There have also been a variety of queries from ACaDMe users in the NHS Boards who have produced different results from their queries compared to published figures in a variety of QI outputs.

This has resulted in a QI investigation to see if a harmonisation of methodology is required across the service area in the first instance, and then following this throughout ISD more widely to ensure a consistent message and approach from ISD as a whole. The content of this paper is the result of this.

BACKGROUND

CIS DEFINITION and RULES

The definition of a CIS stems from the pioneering work of Dr Steve Kendrick and Dr John Clarke in the early to mid 1990s² in the Medical Record Linkage Team as it was then. It was defined as thus:

“A continuous inpatient stay is an unbroken period of time that a patient spends as an inpatient.”

With this definition came a set of rules to help establish how to link the records together to create the stays, which are followed by ISD to this day in the creation of the CIS marker variable and how these are assigned to the records in the production of the SMRA tables and analysis views. However, as noted above while the CIS markers are applied using probability matching and these rules, the records in the tables do not follow the same sort order as applied in the flat files and the linked catalogue and this has to be done manually when reading in the data, as noted by the guidance in the ‘SAF Bulletin No.16 Special Edition - SMRA Guidance’ section 3. However, this is already done in ACaDMe as part of the load process.

¹ http://www.isdscotland.org/Products-and-Services/ACaDMe/
USES, METHODS AND THE RATIONALE OF CIS BASED ANALYSIS IN QI PRODUCTS

There are several different methodologies used to count the number of CIS which are utilised within the QI team. All of these methodologies are correct, as they are used to count slightly different things and time periods (years, quarters, months). Sometimes you might wish to assign a CIS to the first specialty of the stay, or to the first admission type. Or you may wish to count every specialty in the stay; all methods currently done in QI and are detailed below:

General Rules, Conventions and Guidance on SMR01 data and the Creation of CISs in SMR01

SMR01 data is episode based and uses episodes as the base unit of analysis. As noted in the definition above a CIS is a the unbroken period of care that the patient spends as an inpatient. The CIS is based on a series of individual episodes linked by the CIS marker.

When undertaking CIS based analysis the first and most important step is to ensure the episodes in the file are sorted in the recommended sort order, replicating the linked catalogue sort order. The date selections for the file are based on date of discharge, as is standard for all SMR01 analyses.

When conducting CIS based analysis, there are a couple of things that you must consider. When reading in the data from the SMRA analysis view to get the extract from which to count CIS, the period of lookback and lookforward used has an effect on the number of CISs that will be generated. The QI SCT usually uses a lookback of one year as its standard. However, other periods could also be used, depending on the request and the needs of the customer. The longer a lookback used, the more completed CISs will be generated (once the aggregation from episode level data has been applied) as some stays could be very long and have records which start more than a year ago. This conversely applies to the lookforward too. When extracting data from the SMR01 analysis view on SMRA the end date is left open ended. This ensures that no CISs are artificially cut off, thus giving an accurate reflection of number of stays.

Secondary Care - Information Requests and Parliamentary Questions

The SCT receives a variety of requests where, depending on the customer’s requirements, slightly different methods may be used depending on the circumstances.

If counting based on a certain admission type, say emergency admissions, the CIS is assigned an admission type based on the first episode of the stay. In the given example: all those which start with an emergency episode, will be counted as an emergency admission:

---

The CIS is assigned the designated time period of interest – i.e. year, financial year, quarter or month, based on the date of discharge on the first episode of the stay.

Health board, of treatment or residence is also selected based on the first episode in the CIS. This ensures that there is no double counting. This approach is used for most variables, however depending on the needs of the customer sometimes the last specialty or health board is taken if explicitly requested by the customer. In this example the health board for the whole stay is assigned to NHS Highland:

<table>
<thead>
<tr>
<th>Link_no</th>
<th>CIS</th>
<th>Episode</th>
<th>AdmType</th>
<th>Health Board of Treatment</th>
</tr>
</thead>
<tbody>
<tr>
<td>222</td>
<td>3</td>
<td>1</td>
<td>Emergency</td>
<td>NHS Highland</td>
</tr>
<tr>
<td>222</td>
<td>3</td>
<td>2</td>
<td>Transfer</td>
<td>NHS Grampian</td>
</tr>
</tbody>
</table>

For diagnosis a variety of different types of CIS analysis can be done. Last diagnosis is predominantly used for assigning a diagnosis to the stay, however if the customer want to know about stays and what the initial reason for admission the diagnosis assigned to the stay would be based on the first one. However, in cases where the customer was more interested on the diagnosis immediately before discharge, the last diagnosis would be the one selected. There are also instances where we count all diagnoses in the stay, e.g. publication tables or for specific requests where it might be relevant.
Secondary Care - Annual Publication Tables

In the Acute Activity: Inpatients, Outpatients and Beds Annual publication:

- **IPDC table:** An approach is adopted that to capture specialty within a stay, therefore each specialty, diagnosis within the CIS is counted. In the methodology for this approach, the specialty is included in the initial aggregate which turns the episodes into CIS, it is not assigned to the specialty of the first or last episode. This allows a measure of the activity across all specialties to be gauged, where as using an alternative method, may potentially mask the true level of activity in each specialty. When this is presented the figure for all CIS is also calculated, so the total number of CIS presented is not over counted.

- **Emergency Admissions and multiple Emergency Admissions Tables**
  This method is trying to count the number of CISs for each admission type and it provides analysis by age, sex, Health Board of Residence, Council Area, Health and Social Care Partnership, Scottish Index of Multiple Deprivation. This is calculated by first sorting the file in the SAF order and then, in the CIS aggregation selecting the characteristics based on the first episode of the stay. The length of stay (LOS), also referred to as ‘bed days’, for the CIS is calculated by summing up the constituent LOS variables from the episodes in the CIS. This can cause issues as it will provide a slightly different count of CISs to the ones produced by ACaDMe due to the assignation of the CIS to a particular time period.

**CIS calculations in ACaDMe**

ACaDMe has the records in the correct sort order as part of the load process as noted above so it you don’t have to sort the data when using it. In ACaDMe there is a predefined measure to count the number of CIS. This measure in ACaDMe counts CIS is a slightly different way. When using SPSS syntaxes, in QI SCT IRs as noted above the CIS is assigned to a time period based on the first date of discharge in the stay. ACaDMe doesn’t do this, but assigns CIS across all the years it covers, not based on the first date of admission as when using SPSS or R syntax. It is possible to assign different elements to the first or the last episode in ACaDMe. This is important to understand how to replicate especially if customers compare analysis supplied by ISD to their own analysis using BOXI.
Other Quality Indicators Products

HSMR Publication and Dashboard

The Hospital Standardised Mortality Ratios (HSMR) publication calculates CIS for the purpose of producing a risk-adjusted mortality rate for specific hospitals. This is done by counting the number of deaths that occur within 30 days of admission to hospital, including in-hospital deaths and deaths which occur within the community. The information for the variables used in the risk-adjustment (patient age, sex, primary diagnosis etc) is all taken from the first episode within a CIS. This is done by applying the linked catalogue sort order to the data, and then using the lag command to assign an episode number to every episode within a CIS. Unlike most CIS-based analyses, patients are assigned a time period – usually quarter for HSMR – based on the date of admission, rather than date of discharge. As patients can be admitted to hospital again within 30 days of admission, in order to avoid double counting deaths for patients who have died within 30 days of admission for multiple hospital stays, only the first stay where the death is recorded is counted. All subsequent stays are discarded.

Discovery

Discovery contains a suite of indicators for benchmarking performance to support performance and quality improvement. Some of these indicators calculate CIS differently, as they are used for different purposes:

Clinical Profiles Indicators

Within the suite of Clinical Profiles, which currently includes Emergency Laparotomy, Sepsis, Venous Thromboembolism (VTE), Urology and Atrial Fibrillation (AF), the methodology for counting CISs varies according to indicator. The three indicators that are typically included in a Clinical Profile are: length of stay, emergency readmission within 30 days of discharge, and mortality within 30 days of admission.

The methodology for counting CISs has been developed specifically for reporting on specific diagnoses or procedures. The first step, once the data extract has been sorted, is to select all of the CISs that contain the relevant diagnosis code (ICD-10) or procedure code (OPCS-4) for the Clinical Profile. For example, in the Profile for sepsis, all the CISs which contain a diagnosis code of A40 or A41 are selected. For each CIS, an ‘index’ episode is then identified and flagged. The ‘index’ episode is the first episode in the CIS that meets the selection criteria (i.e. diagnosis of sepsis). For all three indicators, the hospital, reporting time period (e.g. quarter) and the variables used for standardisation (age, sex and SIMD) are extracted from the index episode.

For the length of stay indicator, the number of CISs is a count of the stays where the patient was initially admitted as an inpatient and contains an ‘index’ episode. A patient can be counted more
than once in a quarter if they have had more than one inpatient admission to hospital with the relevant diagnosis or procedure recorded e.g. sepsis.

For the emergency readmission indicator, the methodology is similar to length of stay however both day case and inpatient admissions are included. A flag is created to identify if the time between the CIS date of discharge (for a CIS with an index event) and the CIS date of admission of the following emergency admission episode is 30 days or less. In addition, patients who died within the readmission follow up period (30 days) without having been readmitted prior to death are excluded, and episodes where the patient was discharged dead are excluded. The table below shows how CISs are identified for the emergency readmission indicator.

<table>
<thead>
<tr>
<th>Link No.</th>
<th>Admission Date</th>
<th>Qtr</th>
<th>Discharge Date</th>
<th>Type of Admission</th>
<th>CIS</th>
<th>Episode</th>
<th>Sepsis diagnosis (Y/N)</th>
<th>Index episode (Y/N)</th>
<th>CIS flag</th>
</tr>
</thead>
<tbody>
<tr>
<td>222</td>
<td>01-Jan-16</td>
<td>Q1</td>
<td>02-Jan-16</td>
<td>Elective</td>
<td>1</td>
<td>1</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>222</td>
<td>02-Jan-16</td>
<td>Q1</td>
<td>06-Jan-16</td>
<td>Elective</td>
<td>1</td>
<td>2</td>
<td>Y</td>
<td>Y</td>
<td>1</td>
</tr>
<tr>
<td>222</td>
<td>06-Jan-16</td>
<td>Q1</td>
<td>08-Jan-16</td>
<td>Transfer</td>
<td>1</td>
<td>3</td>
<td>N</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td>222</td>
<td>12-Jan-16</td>
<td>Q1</td>
<td>15-Feb-16</td>
<td>Emergency</td>
<td>2</td>
<td>1</td>
<td>Y</td>
<td>Y</td>
<td>1</td>
</tr>
</tbody>
</table>

For the mortality indicator, this differs from the other indicators as the denominator is the number of patients, rather than CISs, as a patient should only be counted once in a quarter. If any individual patient has more than one index hospital stay in a quarter, the final continuous inpatient stay is the one selected. This ensures that the analysis is patient-based, within quarter, and that deaths are counted only once. Please refer to Appendix 2 for an example of the SPSS syntax written to select CISs for the emergency readmission indicator.

**Admissions indicators:**

**Elective Admissions per 1,000 population**

**Non Elective Admissions per 1,000 population**

**Elective Bed Days per 1,000 population**

**Non Elective Bed Days per 1,000 population**

These indicators in discovery are created using SQL scripts that have been translated to SPSS by the Business Intelligence team in IT.

These indicators share the same methodology for how they calculate CIS, assign admission type to each CIS and subsequently how the length of stays is counted for each. The one thing to note that is for these indicators, and many others on discovery use the ‘old tadm’ admission type to define and assign emergency and non emergency admissions. It uses the following criteria to assign the noted characteristics to the CIS.

- Age on admission – From **first** episode (within CIS)
- Date of Discharge – From **last** episode (within CIS)
- Speciality – From **last** episode (within CIS)
- Type of Admission - From **first** episode (within CIS)
Thus this methodology has some similarities to the way things are calculated in the SCT publication tables and also from the clinical profiles.

**Potentially Preventable Admissions indicators**

These adopt the same criteria for CIS selection on dates and admission types as the above indicators. However where they differ is that this indicator looks to provide analysis at a specialty level. It adopts the same approach to the counting CIS as the SCT IPDC table, which counts all the specialties in the stay and assigns CIS to each. When presenting the total stays for all specialties, avoids double counting by only counting the admission will only be counted once. For further information and examples see http://www.isdscotland.org/Products-and-Services/Discovery/Metadata/PotentiallyPreventableAdmissions.pdf.

**SMR04 Continuous Inpatient Stays**

The term CIS is used not only in SMR01, but in SMR04, the mental health inpatient dataset too. Though the same term is used and the broad concept of what it is the same, due to the nature of the mental health data in the two settings it is calculated differently. The nature of the SMR04 data means there are various cases where episodes in SMR04 can overlap as patients can be released ‘on pass’ rather than transferred if they have to get care in an acute setting and so forth. The methodology used to calculate this accounts for the fact there are overlapping episodes of care and tries to assign them to a CIS accordingly. For reference, please see syntax in Appendix 2.

**RECOMMENDATIONS ON STANDARD USE WITHIN PHI**

These approaches should be understood and internalised by analysts and used when undertaking CIS analysis, thus providing a series of standard approaches.

It is recommended that all ISD and HPS analysts undertaking CIS based analysis adhere to the general principles and guidance outlined here. However the following variety of different methods for calculating CIS for different purposes detailed in this paper, for QI alone, it is not practical to apply a finite approach to the calculation and counting of CIS across ISD as a whole. Where possible the methodologies described and corresponding standard syntaxes for each approach could be kept on the PHI Git Hub PHI SPSS git hub, or R Git hub. As developments occur and there is a need for these methods to be updated or altered this can be documented using the version control capabilities within this tool. This would ensure transparency of method and consistency at and organisational level. These snippets of code could either be exported from Git hub, or run calling the programmes from SPSS, or in the future R.
## APPENDICES

### 1. METHODOLOGIES

<table>
<thead>
<tr>
<th>Title of indicator</th>
<th>Description</th>
<th>Reporting Level</th>
<th>Frequency</th>
<th>Analysis Level</th>
<th>Operational Definition</th>
<th>Further Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>CIS by Specialty in Acute Activity Annual publication</td>
<td>Count of the number of times a CIS contains a particular specialty</td>
<td>NHS Board level and specialty level, age, gender, deprivation</td>
<td>Annually</td>
<td>Stay level</td>
<td></td>
<td>[Methodology section of report]</td>
</tr>
<tr>
<td>Number of Emergency Admissions and multiple Emergency Admissions Tables in Acute Activity Annual publication</td>
<td>Count of the number of times a CISs where the stay starts with and emergency episode.</td>
<td>Age, gender, deprivation</td>
<td>Annually</td>
<td>Stay level</td>
<td></td>
<td>[Methodology section of report]</td>
</tr>
<tr>
<td>Title of indicator</td>
<td>Description</td>
<td>Reporting Level</td>
<td>Frequency</td>
<td>Analysis Level</td>
<td>Operational Definition</td>
<td>Further Info</td>
</tr>
<tr>
<td>--------------------</td>
<td>-------------</td>
<td>-----------------</td>
<td>-----------</td>
<td>----------------</td>
<td>-------------------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Hospital Standardised Mortality Ratio (HSMR)</td>
<td>The Scottish HSMR is a risk adjusted mortality measure. The HSMR is calculated as: [ \text{HSMR} = \frac{\text{Observed Deaths}}{\text{Predicted Deaths}} ]</td>
<td>Hospital of Treatment</td>
<td>Quarterly</td>
<td>Patient level (to prevent double counting of deaths) Hospital, date selection, and variables used for case mix adjustment are extracted from the first episode of the patient’s last stay.</td>
<td>Numerator: Observed deaths within 30-days of admission. This is calculated for each patient using the admission date of the first episode of the last stay and the date of death. If the patient is seen in more than one hospital within a stay the outcome is counted against only the first hospital in the stay. Patients with admissions in different quarters will be counted in each quarter. If a patient was admitted in one quarter but died in the subsequent quarter, any admissions in this latter quarter are excluded. This ensures that the analysis is patient-based, within quarter, and that deaths are counted only once. Denominator: Predicted number of deaths within 30-days of admission. This is calculated from a case-mix adjusted model based on the patient’s primary diagnosis; specialty (medical or surgical); age; sex; where the patient was admitted from; the number and severity of prior morbidities in the previous (i) 12 months (ii) 5-years; the severity of co-morbidities; the number of emergency admissions in the previous 12 months; whether admitted as an inpatient or day case; type of admission (elective/ non-elective); and deprivation.</td>
<td>See <a href="#">Technical Specification</a> for more information.</td>
</tr>
</tbody>
</table>
Clinical Profiles (using Sepsis as an example)

<table>
<thead>
<tr>
<th>Title of indicator</th>
<th>Description</th>
<th>Reporting Level</th>
<th>Frequency</th>
<th>Analysis Level</th>
<th>Operational Definition</th>
<th>Further Info</th>
</tr>
</thead>
<tbody>
<tr>
<td>Average and Median Length of Stay</td>
<td>Clinical profiles have been designed as management information systems to be interpreted and used locally by NHS boards and clinicians. They present a range of data about medical and surgical activity and outcomes in Scotland specifically Length of Stay, Emergency Readmissions and Mortality.</td>
<td>Hospital of Treatment</td>
<td>Quarterly</td>
<td>Stay Level</td>
<td>Length of stay is only calculated for patients initially admitted as an ‘inpatient’ and identified with an ‘index’ episode. The ‘index’ episode is the first SMR01 in the stay that meets the selection criteria (e.g. diagnosis of sepsis). Average Numerator = Total length of stay following an admission to hospital in the reported time period, in which the patient was diagnosed with sepsis. Denominator = The number of hospital stays, following initial admission to hospital within reported time period, in which the patient was diagnosed with sepsis. Median This is the middle value when the data are ranked in order.</td>
<td>See Clinical Profiles for further information.</td>
</tr>
</tbody>
</table>

Quarterly Stay

The length of stay is measured from the date of initial admission to the continuous inpatient stay through to the ultimate date of discharge for that stay and potentially encompasses in-hospital and out-of-hospital/board transfers.

Length of stay is only calculated for patients initially admitted as an ‘inpatient’ and identified with an ‘index’ episode. The ‘index’ episode is the first SMR01 in the stay that meets the selection criteria (e.g. diagnosis of sepsis).

Average Numerator = Total length of stay following an admission to hospital in the reported time period, in which the patient was diagnosed with sepsis.

Denominator = The number of hospital stays, following initial admission to hospital within reported time period, in which the patient was diagnosed with sepsis.

Median This is the middle value when the data are ranked in order.
<table>
<thead>
<tr>
<th>Metric</th>
<th>Reporting Level</th>
<th>Date Selection</th>
<th>Variables Used for Standardisation</th>
<th>Numerator</th>
<th>Denominator</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Mortality Rate within 30-days of admission</strong></td>
<td>Patient Level</td>
<td>Quarterly</td>
<td>(age, sex and SIMD)</td>
<td>Number of patients that died within 30 days of an admission to hospital in which the patient was diagnosed with sepsis.</td>
<td>Number of patients that had a diagnosis of sepsis recorded in any position during their hospital stay.</td>
</tr>
<tr>
<td><strong>Emergency Readmissions within 30-days of ultimate hospital discharge</strong></td>
<td>Stay Level</td>
<td>Quarterly</td>
<td>(age, sex and SIMD)</td>
<td>Number of emergency readmissions to hospital within 30 days of</td>
<td>See Clinical Profiles for further information.</td>
</tr>
</tbody>
</table>
ultimate hospital discharge following a diagnosis of sepsis within the reported time period.

Denominator = The number of hospital stays, following initial admission to hospital within reported time period, in which the patient was diagnosed with sepsis.
2. SPSS SYNTAX

This is SPSS syntax, once syntax in R has been developed and tested this will be added to the paper.

Acute activity publication – Stays by Specialty

********************************************************************************
*******************************
*Part 1 - Create initial stays file to count up.
********************************************************************************

get file=!output + 'initial_data.zsav'.

*sort to get in correct order for creating stays.
sort cases by link_no cis_marker admission_date discharge_date.

*** Add on variables to every record before aggregating.
aggregate outfile=*
mode=addvariables
  /break=LINK_NO CIS_MARKER
  /hbtrt_first=first(hbtrtname)
  /hbres_first=first(hbresname)
  /council_first=first(council)
  /finyear_last=last(finyear)
  /period_last=last(period).
execute.

**aggregate to get stays.
aggregate outfile=*
  mode=replace
  /break link_no cis_marker specname spec1
  /DISCHARGE_DATE=max(DISCHARGE_DATE)
  /ADMISSION_DATE=min(ADMISSION_DATE)
  /finyear=last(finyear_last)
  /quarter=last(period_last)
  /adm_type=first(adm_type)
  /ipdc_first=max(ipdc_first)
  /hbres_first=first(hbres_first)
  /council_first=first(council_first)
  /hbtrt_first=first(hbtrt_first).
execute.

*create variables to aggregate by.
*Note: the specialty total should only count each person once for each specialty.

*create a counter to identify episodes.
compute epnum = 1.
execute.

*if same person and same stay, add 1 for every additional episode they have.
if (link_no = lag(link_no) and cis_marker = lag(cis_marker)) epnum = lag(epnum)+1.
execute.

*****All Specialties*****.
*if the first episode in the stay, count in the all specialty total.
string allspec (a50).
do if epnum=1.
compute allspec='All Specialties'.
end if.
execute.

*all admission types.
string alladm (a20).
compute alladm='All Admission Types'.
execute.

*scotland - including 'other'.
string allscotland (a50).
compute allscotland='All Scottish residents and those outwith'.
execute.

*save file for reference.
save outfile=!output+'cis_basefile.zsav'.

********************************************************************************
****************************
*Part 2 - Create macro to perform aggregations
************************************************************************************

define !macro (location = !tokens(1) /admtype= !tokens(1) /data=!tokens(1) /spec=!tokens(1) /time_period=!tokens(1)).
dataset activate initial.

*count total stays at required level.
aggregate outfile=* 
  /break !time_period !location !admtype !spec 
  /stays=n.
execute.

*rename variables ahead of merge.
rename variables (!location=location). 
rename variables (!admtype=admtype).
rename variables (!spec=spec).
rename variables (!time_period=time_period).

*temporarily save dataset.
dataset name !data.

!enddefine.

***********************************************************************
***********************************************************************
*Part 3 - Aggregate for Health Board of Residence combinations - financial years
***********************************************************************
***********************************************************************

*read in basefile.
get file=!output+'cis_basefile.zsav'.

*create additional hbres variables to aggregate by.
*scotland - excluding other/outwith scotland.
string nhsscotland (a50).
do if hbres_first ne 'Other'.
  compute nhsscotland='NHS Scotland'.
end if.
execute.

*temporarily save dataset.
dataset name initial.
*hbres combinations.

!macro location=hbres_first admtype=adm_type spec=spec1 time_period=finyear data=temp1.

!macro location=hbres_first admtype=alladm spec=spec1 time_period=finyear data=temp2.

!macro location=hbres_first admtype=adm_type spec=allspec time_period=finyear data=temp3.

!macro location=hbres_first admtype=alladm spec=allspec time_period=finyear data=temp4.

*NHS Scotland combinations.

!macro location=nhsscotland admtype=adm_type spec=spec1 time_period=finyear data=temp5.

!macro location=nhsscotland admtype=alladm spec=spec1 time_period=finyear data=temp6.

!macro location=nhsscotland admtype=adm_type spec=allspec time_period=finyear data=temp7.

!macro location=nhsscotland admtype=alladm spec=allspec time_period=finyear data=temp8.

*All Scotland combinations.

!macro location=allscotland admtype=adm_type spec=spec1 time_period=finyear data=temp9.

!macro location=allscotland admtype=alladm spec=spec1 time_period=finyear data=temp10.

!macro location=allscotland admtype=adm_type spec=allspec time_period=finyear data=temp11.

!macro location=allscotland admtype=alladm spec=allspec time_period=finyear data=temp12.

*add temp datasets together.
add files file=temp1
   /file=temp2
execute.

dataset name combined.

*close temp datasets.
dataset close temp1.
dataset close temp2.
dataset close temp3.
dataset close temp4.
dataset close temp5.
dataset close temp6.
dataset close temp7.
dataset close temp8.
dataset close temp9.
dataset close temp10.
dataset close temp11.
dataset close temp12.

dataset close initial.
dataset activate combined.

save outfile=!output+'cis_hbres_finyear.zsav'.

**Acute activity publication – Emergency Admissions**

GET DATA
/TYPE=ODBC

/CONNECT='DSN=SMRA;UID=\'+!username+';PWD=\'+!password+';SRVR=SMRA.nss.scot.nhs.uk'

/SQL='SELECT "DR_POSTCODE", "URI", "SEX", "DATAZONE_2011", "ADMISSION", "DISCHARGE" "ADMISSION_DATE", "ADMISSION_TYPE", "DISCHARGE_DATE", "HBTREAT_CURRENTDATE","DATE_RECORD_INSERTED","AGE_IN_YEARS","OLD_SMR1_TADM_CODE","DOB","LENGTH_OF_STAY", "LOCATION", "COUNCIL_AREA", "HBRES_CURRENTDATE", "LINK_NO", '+' "CIS_MARKER" FROM "ANALYSIS"."SMR01_PI_AUG2017" WHERE ("DISCHARGE_DATE" >= "20110401") ORDER BY "LINK_NO" ASC, '+'
"ADMISSION_DATE" ASC, "DISCHARGE_DATE" ASC, "ADMISSION" ASC, "DISCHARGE" ASC, "URI" ASC
/ASSUMEDSTRWIDTH=255.
CACHE.
EXECUTE.

*Sort file before aggregating to CIS level.
sort cases by link_no(a) cis_marker(a) admission_date(a) discharge_date(a).

*Aggregate to create analysis by CIS.

aggregate outfile =*
/break LINK_NO CIS_MARKER
/admission_date=MIN(admission_date)
/discharge_date=MAX(discharge_date)
/sex=first(sex)
/hbres_currentdate=first(hbres_currentdate)
/council_area=first(council_area)
/dr_postcode=first(dr_postcode)
/datazone_2011=first(datazone_2011)
/admission_type=first(admission_type)
/old_smr1_tadm_code=first(old_smr1_tadm_code)
/dob=first(dob)
/length_of_stay=sum(length_of_stay)
/age_in_years=first(age_in_years).
execute.

*Financial year is coded as the year with the majority of the months. Need to update this with the most recent financial year.
*The p is for provisional so needs too be added to the most recent financial year only.

string discharge_finyr (A8).
if (discharge_date >= date.dmy(01,04,2012) & discharge_date < date.dmy(01,04,2013))
   discharge_finyr ='2012/13'.
if (discharge_date >= date.dmy(01,04,2013) & discharge_date < date.dmy(01,04,2014))
   discharge_finyr ='2013/14'.
if (discharge_date >= date.dmy(01,04,2014) & discharge_date < date.dmy(01,04,2015))
   discharge_finyr ='2014/15'.
if (discharge_date >= date.dmy(01,04,2015) & discharge_date < date.dmy(01,04,2016))
   discharge_finyr ='2015/16'.
if (discharge_date >= date.dmy(01,04,2016) & discharge_date < date.dmy(01,04,2017))
   discharge_finyr ='2016/17p'.
execute.
*take out years not of interest.
select if discharge_finyr ne "'.
execute.

******************************************************************************
******
*Recode postcode and match on deprivation quintile. *SIMD. (1-most deprived 5-least deprived).
*Check at next update for new SIMD postcode file.
******************************************************************************
******

*rename and sort to enable match on.
rename variables dr_postcode=pc7.
sort cases by pc7.

*match on simd quintiles for each datazone.
match files file="
  /table = '/conf/linkage/output/lookups/deprivation/postcode_2017_1_simd2016.sav'
  /by pc7.
execute.

string SIMD(a20).
if simd2016_sc_quintile eq 1 simd eq '1 - Most Deprived'.
if simd2016_sc_quintile eq 2 simd eq '2'.
if simd2016_sc_quintile eq 3 simd eq '3'.
if simd2016_sc_quintile eq 4 simd eq '4'.
if simd2016_sc_quintile eq 5 simd eq '5 - Least Deprived'.
execute.

*delete variables which aren't needed.
delete variables simd2016rank to simd2016_crime_rank.

******************************************************************************
******
*Only select emergency admissions.
*Urgent admissions are included as these are mapped to old TADM 8
*This may be reviewed in future??
******************************************************************************
******

select if any(admission_type,'30','31','32','33','34','35','36','37','38','39','20','21','22').
execute.

Clinical Profiles

CLINICAL PROFILES - EMERGENCY READING SAMPLE SYNTAX

******************************************************************************
******
**Calculate Readmission Rates (not standardised).
**Assumes file has been read and sorted as appropriate first.**

*Calculate episode numbers.*
compute epinum = 1.
if (lno = lag(lno) & cis = lag(cis)) epinum = lag(epinum) + 1.
execute.

*Calculate cis date of admission and discharge.*
aggregate outfile = *
mode = addvariables
/break lno cis
/cis_doadate = min(fulldoa)
/cis_dodisdate = max(dodis).

*Identify episodes with relevant diagnosis (as agreed with customer/terminology team) for indicator (e.g. for sepsis profile we identify any diagnoses of A40/A41).*
compute sepsis = 99999.
if (any(diag1_3,"A40","A41") or any(diag2_3,"A40","A41") or any(diag3_3,"A40","A41")
or any(diag4_3,"A40","A41") or any(diag5_3,"A40","A41") or any(diag6_3,"A40","A41")) sepsis = 1.
execute.

*Where diagnosis appears in more than one episode of a single CIS, flag the first episode as the index event.*
compute index = 99999.
if sepsis = 1 index = epinum.
execute.

aggregate outfile = *
mode = addvariables
/break lno cis
/index_marker = min(index).
compute index_event = 0.
if (index_marker = epinum) index_event = 1.
ex.
delete variables index index_marker.

*Identify type of admission - using new type of admission variable as recommended by SAG.*
if (any(admtype,18,20,21,22) | (admtype>=30 & admtype<=39) | (admtype>=40 & admtype<=48)) admgrp = 2.
if any(admtype,10,11,12,19) admgrp = 1.
value labels admgrp
1 "Elective"
2 "Non-Elective".

*Identify date of emergency readmissions i.e cis_doadate of next stay.*
sort cases by lno(a) cis(d) epinum(d).
if lno = lag(lno) & cis ne lag(cis) & lag(adgrp) = 2 emerg_read_doadate = lag(cis_doadate).
exe.
formats emerg_read_doadate (DATE).

*Calculate time to readmission.
aggregate outfile = *
mode = addvariables
/break lno cis
/cis_emerg_read_doa = max(emerg_read_doadate).

sort cases by lno(a) cis(a) epinum(a).

compute emerg_readdays = datediff(cis_emerg_read_doa, cis_dodisdate,"days").
recode emerg_readdays (sysmis = 99999).

*Flag if the emergency readmission episode is <= 30 days from the cis date of discharge.
compute emergread = 0.
if emerg_readdays >=0 and emerg_readdays <=30 emergread = 1.
execute.

**Identify exclusions**.
*ie cis' where the patient was discharged dead or died within the readmission period.

*Calculate death_inhosp marker.
sort cases by lno (a) cis (a) fulldoa (a) dodis (a).
compute death_inhosp = 0.
if ((distype >= 40 & distype <= 49)) death_inhosp = 1.
execute.

aggregate outfile=*
mode=addvariables
/break=lno cis
/death_inhosp_cis=MAX(death_inhosp).
execute.

*Calculate days between discharge and death.
compute dthdays_dis = datediff(ddeath, cis_dodisdate,"days").
recode dthdays_dis (sysmis = 99999).

compute excluderead= 0.
if ((dthdays_dis ne 99999 and (dthdays_dis = 0 or (dthdays_dis >0 and dthdays_dis <=30 and dthdays_dis < emerg_readdays)) or death_inhosp_cis =1) excluderead=1.
execute.

*Select index episodes (excluding those that have been flagged to be excluded).
select if index_event = 1 & excluderead = 0.
execute.

*Aggregate by location to total number of emergency readmissions and number of stays.
aggregate outfile=* 
/break quarter Locname 
/read=sum(emergread) 
/stays=n.
execute.

** SMR04 SPSS - CIS Syntax **

** Sort by admission date before discharge date. **
sort cases by link_no admission_date discharge_date.

** Rewrite admission and discharge dates as number of days. **
compute admitday = ctime.days(admission_date).
compute disday = ctime.days(discharge_date).

** Initialize end of stay with discharge date. **
compute cis_end = disday.
compute cis_marker = 1.

** Main CIS criteria. **
do if linkno = lag(linkno).
   do if admitday <= lag(cis_end) or (admitday = lag(cis_end)+1 and
   (any(lag(dis_type),'1','4','8','12','13','14')) or missing(lag(cis_end)).
   compute cis_marker = lag(cis_marker).
   if not missing(cis_end) cis_end = max(cis_end,lag(cis_end)).
   if (missing(lag(cis_end)) and (any(dis_type,'1','4','8','12','13','14'))) cis_end=lag(cis_end).
else.
   compute cis_marker = lag(cis_marker)+1.
end if.
execute.

delete variables admitday disday cis_end.
alter type cis_marker (f8).