

***Use of social care data for
impact analysis and risk
stratification***

Sunderland CCG

29 August 2014

Executive summary

Sunderland CCG currently gets access to secondary care and primary care data through their commissioning support unit. This arrangement gives them access to acute, community and mental health data.

Social care data does not fall under an existing section 251 and in order to share there must be a 'legal basis' for sharing.

Although other organisations have also built datasets for risk stratification or other modelling there do not appear to be any obvious lessons that Sunderland can take from these.

Some such as Torbay have not made use of social care data in their modelling. Others such as Somerset, where they have used social care data were put together before the restrictions on the use of patient level data.

Introduction

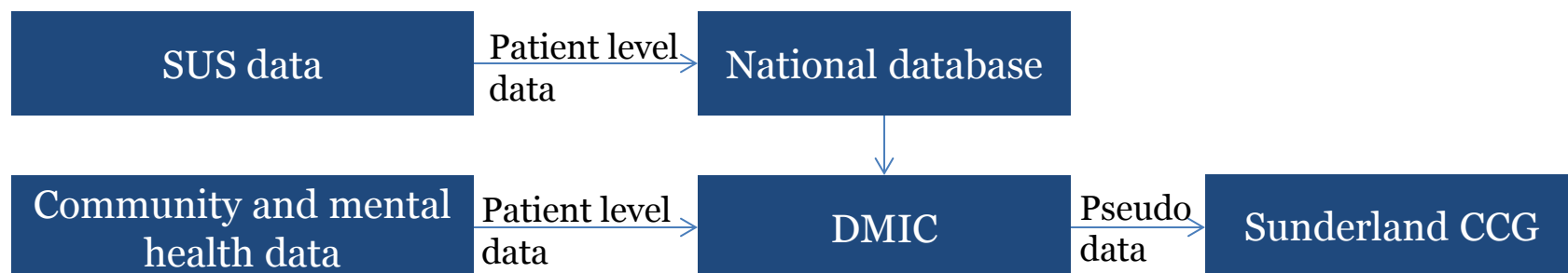
This report discusses experiences in the use of social care data in combination with health data for two purposes:

- Risk stratification of patients
- Impact analysis for proposed changes

It describes the work done by Sunderland CCG in this area, their successes and the issues that they have encountered. It also describes the experiences in some other areas that have also attempted to use the data in this way and lessons that can be learned from them.

Current use of health and social care data in Sunderland – primary data sets

Data management support is provided by the Commissioning Support Unit (DMIC) who process the data and provide pseudonymised data at patient level to the CCG. These staff are seconded to the Health and Social Care Information Centre (HSCIC) and operate in a Safe Haven environment and are thus allowed to process a range of patient identifiable information from a variety of allowed sources in pseudo form. The pseudonym is a unique value and can be used to join the datasets.



Secondary Uses data are processed by the DMIC and pseudonymised before being released to the CCG. Community and Mental Health data are also sent by Community and Mental Health Providers to the DMIC who then pseudonym the data and provide to the CCG.

Both of these data sets are covered by a Section 251 which allows provision of these pseudonymised data to the CCG.

Current use of health and social care data in Sunderland – additional datasets

Primary care data

Some data from primary care systems are currently used. These data are covered by a data sharing agreement between the CCG, General Practices, NECS and the HSCIC for a limited number of purposes which include patient Risk Stratification.

Data received does not include anything related to appointments in GPs. This issue is currently being explored but is closely related to the national issues such with Care.Data and Summary Care Record

Social care data

No patient level social care data are currently being used to supplement the health data. A request has been made for a similar arrangement through the DMIC for pseudonymised data that could then be linked to other datasets. This has not been possible due to data sharing law.

Summary of data sharing issues

Social care data are not covered by the existing section 251 so in order for the data to be shared there needs to be a 'legal basis'.

Suffolk County Council have also approached the DMIC with a similar request and created a data sharing agreement but this has been reviewed by HSCIC who believes it does not constitute a legal basis.

Although information sharing is permitted for the purposes of care for an individual, when the data are shared for the benefit of the community the expectation is that only anonymised information can be shared. This allows analysis of summary data in its own right but does not allow linking of data sets for more detailed analysis.

NHS England approves the sharing of data from GPs for the purposes of patient risk stratification (www.england.nhs.uk/ourwork/tsd/ig/risk-stratification/) where this is carried out by an approved organisation but again this approval does not extend to social care where the guidance is *“the relevant parties need to assure themselves there is a legal basis for the disclosure and linkage for this purpose. This can be achieved either by using a third party and pseudonymised data, or with consent.”*

Summary of data sharing issues

HSCIC provides guidance on how to perform linkages for the purpose of social care (www.hscic.gov.uk/media/12822/Guide-to-confidentiality-in-health-and-social-care/pdf/HSCIC-guide-to-confidentiality.pdf). Linkages must be done within a trusted environment with controls which may include:

- signed contracts or agreements which stipulate how the information will be used
- demonstration of meeting the required standards of security
- an independent auditor's opinion of security and privacy measures.

These data are only considered anonymised within these controls and not if shared outside of them.

Use of social care data on impact analysis

Based on desktop research, there has only been a limited amount of detailed impact modelling done using social care data. In many areas the main focus has been on health impact and in particular secondary care impact as this is where the majority of the cost reductions are being targeted.

Where modelling has taken place this has been done on a 'top down' basis by making assumptions about changes in the number of people in care homes and scaling costs accordingly.

One method for doing this kind of modelling is the LGA tool, which can provide an indication of the potential size of opportunities by using the value cases included within the toolkit. This does only provide indicative modelling and contains simplistic assumptions around the cost impact of changes in activity levels.

As part of BCF submissions a number of health economies have forecast social care metrics but this seems to have been done using aggregated and publicly available data rather than detailed patient / client level data.

Work from other health economies

Many health economies have used a variety of data sets either for risk stratification or impact modelling. Below are a few examples that illustrate some lessons from other areas:

- 1) The 'Torbay' model (Devon Predictive Model)
- 2) Symphony model at South Somerset
- 3) Kent and Medway Risk Stratification project
- 4) Nuffield Trust research study
- 5) Social care modelling for understanding the impact of the Care Act

The Torbay model

Predictive modelling in Devon began with the PARR and PARR++ model. In 2010/11, the Torbay model, a predictive model, was designed to forecast emergency admissions in Devon. It uses variables which are significant and uses local datasets (registration duration, deprivation, ambulance, etc.) to capture more risk.

The Torbay model is currently not based on any social care data and their experience when building it was that social care variables in the predictive model don't help to explain significant variation in healthcare outcomes so it was not a high priority.

The modellers have since attempted to get hold of more social care data in order to test further hypotheses around social care but have not been able to get the data required due to anxieties over Caldicott 2.

Work from other health economies (cont.) – The Symphony Project

The Symphony Project at South Somerset

The purpose of this modelling project is to show the importance of multi-morbidity to explain utilisation and costs across health and social care settings.

The Symphony Project has built a large dataset which links acute, primary care, community, mental health and social care data for each individual in the South Somerset population. The data are derived from various primary sources all of which cover twelve months from April 2012 to March 2013. The dataset contains anonymised individual-level data about what care has been received and at what cost across all organisations. The dataset has three key features:

1. The Symphony Project dataset spans primary, community, acute, mental health and social care. This includes activity, costs, clinical conditions, age, sex and ward of residence for the entire population of 114,874 people in 2012.
2. Costs have been calculated for each individual according to the type of care they have received in each setting. These calculations generally reflect the costs to the commissioner of procuring care of a particular type.
3. Demographic characteristics are available for each individual, including age, gender, socioeconomic measures, and indicators of morbidity.

Each individual's morbidity profile was constructed using United Health's RISC tool used locally to predict unplanned hospital admissions. 49 chronic conditions were identified to construct the morbidity profile of each individual in the population. The Symphony project was created before organisation structural changes and new rules around information governance and so there was no data issues in using patient level data.

Work from other health economies (cont.) – Risk stratification at Kent and Medway PCT

The risk stratification work led by Kent and Medway Public Health Observatory involved designing and developing a baseline profile of population utilisation of health and social care services for commissioners, with special emphasis on the impact of multiple morbidities.

The work involved bringing together information and intelligence teams from various provider organisations particularly the Kent & Medway Health Informatics Service Business Intelligence. To begin with a historical registered K&M population list as at April 2011 was obtained. The risk stratification tool was then applied, generating a risk score for each citizen in the population based on 2 years' worth of hospital activity data between 09/10 to 10/11. This was then risk stratified into 4 risk groups or bands (based on the Kaiser Permanente Pyramid model of care) using the locally developed tool, a variant of King's Fund Combined Predictive model. Datasets from all other services such as community health, mental health, social care, primary care, continuing health care and mortality were linked via pseudonymisation of NHS numbers, the common patient identifier.

Data Linkage of adult social care

The first major dataset that was linked to the risk stratified population list was the adult social care dataset. Kent County Adult Social Care utilisation data is collected on the County Directorate SWIFT database which currently holds activity data for more than 44,000 clients. Batch loading of NHS numbers was carried out in early in 2012 with a 70% matching rate on social care clients - 31,260 clients. Information governance approval was sought by the respective Caldicott Guardians to link both datasets together via a common pseudonymisation encryption program after applying the risk stratification tool on the hospital activity data. Out of the 31,260 clients, 24,320 of them were linked to the risk-stratified list of patients. Therefore approximately 55% of the clients in the social care database were matched to the risk stratification list of Kent patients.

Other Datasets used (2009/10 to 2011/12):

Hospital activity (Kent & Medway), Social Care Activity (Kent), Mental Health Contacts (Kent & Medway), Mental Health admissions (Kent & Medway), Community Health contacts (Kent), Continuing Care (West Kent and DGS CCGs)

Feasibility of predictive modelling for health and social care

Nuffield Trust Research Study

The Nuffield Trust carried out a study to see if predictive models for health and social care are feasible. The main questions the study was trying to understand were:-

- Do the data exist in local systems?
- Can individual, person-level data be extracted?
- Can a wide range of health and social care data at the person level be linked?
- Are the data accurate and complete enough to use?
- Can a valid statistical model be built?

The duration of the project was 18 months and the Trust worked with 5 sites to extract health and social care datasets at individual patient level. Data from GP records, GP register, Hospital and social care were linked together (though not completely) and exploratory analyses were undertaken and a range of models created to test the impacts of different datasets on the models.

Nuffield faced a few problems when linking the data across health and social care as there is no NHS number in social care and data was missing from some datasets. Nuffield also highlighted that there were important issues surrounding patient confidentiality and approval to use data to be dealt with at the onset of any predictive modelling. Nuffield could access user level social care data but due to different systems there were issues with standardisation/coding and the absence of standard, structured coding schemes.

The overall conclusion was valid models can be built though they need data usage approval and the refinement of data systems to facilitate the modelling.

Use of social care data for risk stratification and impact analysis

Work from other health economies (cont.) – Models related to the Care Act

Models related to the Care Act

Birmingham Care Bill Model

Birmingham City Council have been looking to address the financial implications of the Care Act and have created a financial model to address this.

- The model calculates the impact of both the Care costs Cap and the increase to the means tested threshold accurately for both Older and Working age adults.
- It also models residential clients and clients in the community separately.

Lincolnshire Social Care Impact model

This model is designed to understand the cost impact on local authorities following increased demand for information and services. It builds on work done previously done by Lincolnshire County Council.

Current national work is focused on ensuring costs are estimated accurately; distributional impacts are correctly addressed within allocation formulae; and councils are supported to effectively manage demand for additional assessments and deliver carers' support.

The Lincolnshire model is used to estimate 2015/16 costs to ensure costs stay within the financial envelope of £470 million. However, the initial focus of this model is the over 65s.