Integrated Personalised Commissioning
SW Finance Event

Integrated Health and Social Care –
Patient Level Datasets and Data Analysis in Somerset

Kevin Hudson
Business Solutions & Innovation
LPF Bid Director
Kevin.hudson@swcsu.nhs.uk
07717 530 220
Why Integrated, Patient-level Datasets?

“If you look at the smarter CCGs, they are now taking on responsibility for the entire healthcare budgets in their areas, and there is a crucial bit of plumbing that needs to happen before we can really get behind the quality of whole person commissioning that we need – which is for CCGs to have proper per-patient costings so that they across primary care, secondary care, social care, Roy Lilley costed us £3,400 last year. Until we get that, we won’t actually get smart commissioning.”

Jeremy Hunt – ‘Health Chat’ with Roy Lilley
11th February 2015
Telling the story through data – BIG DATA

- Most data analysis (and commissioning) is ‘episodic’ - Episodes of care, amalgamated over time and categorised by provider

- Starting in 2012, We built a **Holistic data model** – where the patient is the ‘base unit’, not the episode

- 577,000 Somerset patients – activity mapped and costed
  - All their encounters with all aspects of Health & Social Care
  - Primary, Community, Acute, Mental, Social
  - £676M of health and social care spending

- Cross-cut against all their diagnostic conditions
  - mapped from 400 Episode Treatment Groups recorded through Primary and Secondary Care Coding
  - Distilled and analysed to 49 key conditions, and simplified further (to 8)

- Biggest Challenge: Not how to build it but how to understand it
What’s in the Dataset?

- Dataset for the whole of Somerset (577,000 patients): Fully pseudonymised; age, gender, ward of residence (LSOA), data available split by Trust and episode if required.

- What’s in
  - **Primary Care**: Number and cost of GP contacts; Prescribing activity and costs; Clinical conditions and diagnoses – *Sourced through RISC system*
  - **Community Services**: District nursing and health visiting; Ambulance service; Podiatry; Dietetics; Community diabetes service; Rehab; Community therapies; Tissue viability; Speech and language therapies; Continence; End of life: *Sourced from Somerset Partnership*
  - **Community hospitals**: Minor injuries; Outpatients; Interface clinics (Partnership only); Inpatient admissions; Length of stay; Tariff costs for minor injuries and outpatients; Bed-day cost for community hospitals – *Sourced through SUS / cost per bed day from Partnership*
  - **Mental health cluster costs**: Community and inpatient activity – *Sourced from Partnership*
  - **Acute Care activity and costs**: A&E; outpatients; elective inpatients; daycases; non-elective admissions); Clinical coding; Length of stay; Tariff costs – *Sourced through SUS*
  - **Social Care activity and costs**: residential care (net costs); home care; day care; direct payments; professional support; equipment; ‘on costs’; - *Sourced from Somerset CC*
  - **Continuing Health-care**: *Sourced from Somerset CCG / SWCSU*

- What’s not in.
  - **Non PBR Acute Services**: Acute prescribing, diagnostics.
  - **Actual ‘Provider costs’**: Currently it’s a ‘commissioner cost’ model.
Episode Treatment Groups

- A comprehensive condition classification tool that organises all relevant medical services across all sites of care (GP Practice, Inpatient, Outpatient, A&E) for over 400 specific medical conditions.
- Using routinely collected encounter data as input, the ETG software captures the relevant services (including prescriptions) provided during the course of a patient’s treatment, and organizes the encounter data into meaningful episodes of care.
- For Symphony, 49 subsets of ETGs were identified and included in the analysis.

<table>
<thead>
<tr>
<th>Condition</th>
<th>Frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension</td>
<td>16903</td>
</tr>
<tr>
<td>Asthma</td>
<td>12208</td>
</tr>
<tr>
<td>Anxiety</td>
<td>7545</td>
</tr>
<tr>
<td>IBS</td>
<td>5491</td>
</tr>
<tr>
<td>Cancer</td>
<td>5466</td>
</tr>
<tr>
<td>Diabetes</td>
<td>5327</td>
</tr>
<tr>
<td>Skin infections</td>
<td>4741</td>
</tr>
<tr>
<td>Coronary Artery Disease</td>
<td>4440</td>
</tr>
<tr>
<td>Hypothyroidism</td>
<td>4095</td>
</tr>
<tr>
<td>Skin trauma</td>
<td>3449</td>
</tr>
<tr>
<td>Stroke</td>
<td>2455</td>
</tr>
<tr>
<td>Tendinitis</td>
<td>2402</td>
</tr>
<tr>
<td>Gastric signs symptoms</td>
<td>2036</td>
</tr>
<tr>
<td>Breast disorders</td>
<td>1838</td>
</tr>
<tr>
<td>COPD</td>
<td>1719</td>
</tr>
<tr>
<td>Epilepsy</td>
<td>1401</td>
</tr>
<tr>
<td>Fractures</td>
<td>1345</td>
</tr>
<tr>
<td>Appendicitis</td>
<td>1187</td>
</tr>
<tr>
<td>Gastritis</td>
<td>1146</td>
</tr>
<tr>
<td>Mental health</td>
<td>1143</td>
</tr>
<tr>
<td>Hemorrhoids</td>
<td>1142</td>
</tr>
<tr>
<td>Gynecological infections</td>
<td>1069</td>
</tr>
<tr>
<td>Hernias</td>
<td>1021</td>
</tr>
<tr>
<td>Heart Failure</td>
<td>984</td>
</tr>
<tr>
<td>Cataract</td>
<td>954</td>
</tr>
<tr>
<td>Pneumonias</td>
<td>808</td>
</tr>
<tr>
<td>Adult Rheumat Arthritis</td>
<td>693</td>
</tr>
<tr>
<td>Acute bronchitis</td>
<td>620</td>
</tr>
<tr>
<td>Open wound</td>
<td>618</td>
</tr>
<tr>
<td>Late effects compl</td>
<td>618</td>
</tr>
<tr>
<td>CKD/Renal failure</td>
<td>596</td>
</tr>
<tr>
<td>Late effects compl</td>
<td>596</td>
</tr>
<tr>
<td>Glaukoma</td>
<td>536</td>
</tr>
<tr>
<td>Pancreatitis</td>
<td>473</td>
</tr>
<tr>
<td>Dementia</td>
<td>465</td>
</tr>
<tr>
<td>Obesity</td>
<td>401</td>
</tr>
<tr>
<td>Enviromental trauma</td>
<td>380</td>
</tr>
<tr>
<td>Nutr Deficiency</td>
<td>316</td>
</tr>
<tr>
<td>Macular degeneration</td>
<td>280</td>
</tr>
<tr>
<td>Pois eff drugs</td>
<td>278</td>
</tr>
<tr>
<td>Burns</td>
<td>260</td>
</tr>
<tr>
<td>Incontinence</td>
<td>232</td>
</tr>
<tr>
<td>Alcohol dependece</td>
<td>160</td>
</tr>
<tr>
<td>Kidney stones</td>
<td>145</td>
</tr>
<tr>
<td>Eating disorders</td>
<td>106</td>
</tr>
<tr>
<td>Cirrhosis</td>
<td>95</td>
</tr>
<tr>
<td>Diabetic retinopathy</td>
<td>89</td>
</tr>
<tr>
<td>Ulcer</td>
<td>78</td>
</tr>
<tr>
<td>Occup Pulmonary disease</td>
<td>38</td>
</tr>
<tr>
<td>Hepatitis</td>
<td>26</td>
</tr>
</tbody>
</table>
Data Sharing & Information Governance

To **build the dataset you need identifiable data** and to link through the patient. Our IG protocols are based around the following tenets:

- In a moving governance environment, be at the forefront of all national accreditations (e.g. risk stratification)
- For information that you will need add to that dataset, pseudonymise at source. (plus all sources need to use NHS number – potential challenge for Local Authorities)
- Those individuals handling and building the data are national secondees to HSCIC (so are legally authorised) but work out of Bridgwater.
- Assume patients opting out of Summary Care Record also opt out here. (not material)

To **use the data for commissioning, you do not need identifiable data.**

- Relatively simple data-sharing agreement has allowed ‘many eyes’ on the outputs
- ‘Eyes’: CCG / CSU / local providers / Public Health, York University, SWAHSN, Berings / We Predict (SBRI), GE Finnamore / PWC / Oliver Wyman / Mckinseys (for Monitor)

Also used ‘identifiably’ for **direct patient care** in test and learn group
Centre for Health Economics at York University provided support to help interpretation and modelling.

Led by Andrew Street, Professor of Health Economics, and involved in the DH’s ‘Year of Care’ work.

York have described Symphony as the richest data set they’ve seen combining health and social care.

Lesley Gallier, SWCS senior information support manager worked very closely with Panos Kasteridis, research fellow at York.

Series of weekly conference calls as the interpretation was developed and refined.
Starting point – A ‘frail-elderly’ project
An opportunity to analyse by condition?
**What drives cost – age or conditions?**

<table>
<thead>
<tr>
<th>Regression variables</th>
<th>Age</th>
<th>Number of conditions</th>
<th>Age, Number of conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Variation explained</td>
<td>3.36%</td>
<td>18.76%</td>
<td>19.30%</td>
</tr>
</tbody>
</table>
...and this can lead to a ‘capitated formula’
(with appropriate risk pooling arrangements)

Cost of person $i$

$$\text{cost}_i = \alpha_1 + \beta_1 \text{age}_i + \varepsilon_i$$

Statistical error

Number of conditions

$$\text{cost}_i = \alpha_1 + \beta_2 \text{ETG}s_i + \varepsilon_i$$
### Patients with one and multiple ETGs

<table>
<thead>
<tr>
<th>Condition</th>
<th>Number</th>
<th>Single</th>
<th>Plus 1</th>
<th>Plus 2</th>
<th>Plus 3 or More</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypertension (17,777)</td>
<td>4955</td>
<td>5116</td>
<td>3659</td>
<td>4047</td>
<td></td>
</tr>
<tr>
<td>Asthma (12,769)</td>
<td>6209</td>
<td>3215</td>
<td>1665</td>
<td>1680</td>
<td></td>
</tr>
<tr>
<td>Anxiety (7,962)</td>
<td>2627</td>
<td>1451</td>
<td>1522</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cancer (5,932)</td>
<td>1078</td>
<td>1409</td>
<td>1825</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IBS (5,688)</td>
<td>1865</td>
<td>1700</td>
<td>1073</td>
<td>1050</td>
<td></td>
</tr>
<tr>
<td>Diabetes (5,676)</td>
<td>829</td>
<td>1362</td>
<td>1926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Skin infections (5,086)</td>
<td>1945</td>
<td>1294</td>
<td>772</td>
<td>1075</td>
<td></td>
</tr>
<tr>
<td>CAD (4,695)</td>
<td>706</td>
<td>1084</td>
<td>1904</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hypothyroidism (4,275)</td>
<td>981</td>
<td>1179</td>
<td>954</td>
<td>1161</td>
<td></td>
</tr>
<tr>
<td>Skin trauma (3,611)</td>
<td>1334</td>
<td>873</td>
<td>584</td>
<td>820</td>
<td></td>
</tr>
<tr>
<td>Stroke (2,665)</td>
<td>269</td>
<td>641</td>
<td>1160</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tendinitis (2,578)</td>
<td>739</td>
<td>432</td>
<td>585</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gastr signs symptoms (2,144)</td>
<td>724</td>
<td>378</td>
<td>494</td>
<td></td>
<td></td>
</tr>
<tr>
<td>COPD (1,989)</td>
<td>206</td>
<td>450</td>
<td>901</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Breast disorders (1,951)</td>
<td>259</td>
<td>565</td>
<td>636</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fractures (1,546)</td>
<td>511</td>
<td>382</td>
<td>380</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Percentage and number of patients

- Single
- Plus 1
- Plus 2
- Plus 3 or more
Average costs for patients with diabetes by setting and number of other ETGs

- Diabetes: N= 829
- plus 1: N=1,529
- plus 2: N=1,362
- plus 3: N= 898
- plus 4: N= 527
- plus 5: N= 279
- plus 6 or more: N= 222

Number of ETGs
But not all follow this pattern…(Dementia)
For multi-morbidity, number \approx type
For multi-morbidity, number ≈ type
For multi-morbidity, number ≈ type
For multi-morbidity, number $\approx$ type

**Average costs for Dementia plus 4 or more comorbidities**

What to change?

Multi-morbidity, not disease type
...Which led to the creation of the 3+ (of 8) cohort in Somerset:

2013/14 Year-of-Care costs by condition

<table>
<thead>
<tr>
<th>Condition</th>
<th>Total Cohort N=7468</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cancer</td>
<td>N=3715</td>
</tr>
<tr>
<td>Heart Disease</td>
<td>N=5335</td>
</tr>
<tr>
<td>COPD</td>
<td>N=2330</td>
</tr>
<tr>
<td>CKD</td>
<td>N=1780</td>
</tr>
<tr>
<td>Dementia</td>
<td>N=1327</td>
</tr>
<tr>
<td>Diabetes</td>
<td>N=4237</td>
</tr>
<tr>
<td>Depression</td>
<td>N=2366</td>
</tr>
<tr>
<td>Stroke</td>
<td>N=3307</td>
</tr>
</tbody>
</table>
...although for this cohort, 'multi-morbidity is the norm'
And for the LIG’s (Local Implementation Groups) we map where the patient’s are...
...their cost...
...and the setting of care ('at home' or 'in a home')
Mosaic Public Sector classifies all households and postcodes to comprehensively describe their social, economic and cultural behaviour.
Mosaic Social Indicators can help inform scope and engagement:
Type M56 – Older people living on social housing estates with limited budgets

Key Features:
• State pensioners
• Low use of credit
• Enjoy reading
• Small housing
• Basic education
• Shop locally
• Traditional
• Lifelong council tenants
• Face to face contact

Communication Preferences.
Access Information
• Local Papers and Face to Face
• Not Internet, Telephone, Magazines, SMS Text

Service Channels
• Face to Face
• Not Internet, Mobile Phone or Telephone

2% Population, 8% of full cohort, 12% of high cost patients
Family Lifestage (2011)

- The population has also been analysed in terms of patient household.

Family Lifestage (2011) classifies households into the following groups:

<table>
<thead>
<tr>
<th>Code</th>
<th>Family Lifestage 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>00</td>
<td>Young singles/homesharers</td>
</tr>
<tr>
<td>01</td>
<td>Young family no children</td>
</tr>
<tr>
<td>02</td>
<td>Young family with children</td>
</tr>
<tr>
<td>03</td>
<td>Young household with children</td>
</tr>
<tr>
<td>04</td>
<td>Mature singles/homesharers</td>
</tr>
<tr>
<td>05</td>
<td>Mature family no children</td>
</tr>
<tr>
<td>06</td>
<td>Mature family with children</td>
</tr>
<tr>
<td>07</td>
<td>Mature household with children</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Code</th>
<th>Family Lifestage 2011</th>
</tr>
</thead>
<tbody>
<tr>
<td>08</td>
<td>Older single</td>
</tr>
<tr>
<td>09</td>
<td>Older family no children</td>
</tr>
<tr>
<td>10</td>
<td>Older family/household with children</td>
</tr>
<tr>
<td>11</td>
<td>Elderly single</td>
</tr>
<tr>
<td>12</td>
<td>Elderly family no children</td>
</tr>
<tr>
<td>U</td>
<td>Unclassified</td>
</tr>
</tbody>
</table>
What is the split in care costs by Family Lifestage (2011) within the South Somerset Population?

Family Lifestage (2011)
0 - Young singles/homesharers
1 - Young family no children
2 - Young family with children
3 - Young household with children
4 - Mature singles/homesharers
5 - Mature family no children
6 - Mature family with children
7 - Mature household with children
8 - Older single
9 - Older family no children
10 - Older family/household with children
11 - Elderly Single
12 - Elderly family no children

Average cost for 11 (elderly single) is nearly £500 higher than 12 (elderly family, no children)

Average cost for 9 & 10 (older with family) are lower than for 8 (older single)
KAH comment for discussion: We think across all mosaic groupings, single patients have higher costs that those in families. Also note higher social costs and mental health costs for groups single patients in groups 58/59.

York are calculating the effects of these social indicators on cost in the same way as they did for age and conditions.
### Complex maths is needed to model the future

<table>
<thead>
<tr>
<th>3+ group</th>
<th>% with zero costs</th>
<th>Av. Cost for users</th>
<th>R²</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>0</td>
<td>8,152</td>
<td>28%</td>
</tr>
<tr>
<td>Primary care</td>
<td>0</td>
<td>892</td>
<td>17%</td>
</tr>
<tr>
<td>Inpatient</td>
<td>35</td>
<td>4,842</td>
<td>17%</td>
</tr>
<tr>
<td>Outpatient</td>
<td>28</td>
<td>567</td>
<td>16%</td>
</tr>
<tr>
<td>A&amp;E</td>
<td>51</td>
<td>239</td>
<td>15%</td>
</tr>
<tr>
<td>Mental health</td>
<td>88</td>
<td>3,922</td>
<td>15%</td>
</tr>
<tr>
<td>Community</td>
<td>91</td>
<td>7,476</td>
<td>20%</td>
</tr>
<tr>
<td>Social care</td>
<td>69</td>
<td>5,622</td>
<td>14%</td>
</tr>
<tr>
<td>Continuing</td>
<td>91</td>
<td>7,773</td>
<td>27%</td>
</tr>
</tbody>
</table>

‘The starting point of the modelling’: Extract from York’s presentation to iHEA 2014 – High level summary of each setting of care showing a high level of ‘zero cost’ for patients in some settings (but not all) and a strong correlation (R²) between how well recorded patient characteristics can predict cost in those respective settings.

(iHEA 2014 – World Congress of Health Economics, Dublin, July 2014)
Total costs – 3+ group

<table>
<thead>
<tr>
<th>Variable</th>
<th>Estimate</th>
<th>Marginal Effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>0.007**</td>
<td>£68</td>
</tr>
<tr>
<td>No comorbidities</td>
<td>0.260***</td>
<td>£2,501</td>
</tr>
<tr>
<td>Died</td>
<td>0.802***</td>
<td>£7,709</td>
</tr>
<tr>
<td>Dementia</td>
<td>0.669***</td>
<td>£7,569</td>
</tr>
<tr>
<td>CKD/Renal</td>
<td>0.520***</td>
<td>£5,635</td>
</tr>
</tbody>
</table>

‘The next step – calculate marginal effects’: ‘The multivariate regression will calculate the marginal effect of changes in the patient characteristics. For instance, an increase in age (1 year) does not add much to the patient cost (£68), an increase of 1 additional co-morbidity adds £2,501 and patient death increases cost in that year by £7,709 etc…

…and from this model the expected changes in cost arising from changes either in the characteristics, location or overall volume of care…. 

But a ‘currency converter’ is being developed.
Current Utilisation of the Datasets

- Patient lists for the Test and Learn pilots & practice summaries
- Rehab services review
- Dermatology Services Review
- 1+cohort analysis and patient vignettes
- Close working with Integrated Personalised Commissioning
- Capitated Budgets – Analysis of Provider costs, including a split of Community Services costs.
Further Improvement of our Understanding of the Datasets

- Further research commissioned from York University
  - Are the observations sustained? Repeat and refine the analyses
  - Do social conditions drive cost, if so by how much?
  - Is there a data-driven evaluation methodology for integrated care?
  - Econometric Calculator / ‘Currency Convertor’ between settings of care

- Wide sharing with ‘advisories’ supporting CCG, Trusts and national organisations (e.g. Monitor) for modelling & research

- Link with Small Business Research Initiative (SBRI) through Academic Health Science Network
  - Bering Research Ltd (Academically led bioinformatics)
  - We Predict Ltd (GP led with links to automotive & manufacturing)
Preparing for Outcomes Based Commissioning & Capitated Budgets
2014 Health & Social Care for the population of Somerset

<table>
<thead>
<tr>
<th>Number of Patients</th>
<th>3 or more morbidities (of 8)</th>
<th>1 or more morbidities (of 8)</th>
<th>Adults 18 and over</th>
<th>All Somerset Patients</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>7,468</td>
<td>113,981</td>
<td>464,227</td>
<td>577,982</td>
</tr>
<tr>
<td>%age of total</td>
<td>1.3%</td>
<td>19.7%</td>
<td>80.3%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Proportion of group received care in:</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GP Practice</td>
<td>98.5%</td>
<td>93.8%</td>
<td>74.5%</td>
<td>71.3%</td>
</tr>
<tr>
<td>Acute Admissions (Elective &amp; Non)</td>
<td>63.6%</td>
<td>34.0%</td>
<td>16.5%</td>
<td>15.4%</td>
</tr>
<tr>
<td>Mental Health</td>
<td>14.8%</td>
<td>6.0%</td>
<td>2.3%</td>
<td>1.8%</td>
</tr>
<tr>
<td>Community Services</td>
<td>60.0%</td>
<td>27.6%</td>
<td>13.9%</td>
<td>14.2%</td>
</tr>
<tr>
<td>Social Care</td>
<td>22.6%</td>
<td>6.4%</td>
<td>2.1%</td>
<td>1.7%</td>
</tr>
<tr>
<td>Cost of Care</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total cost of Health &amp; Social Care (£)</td>
<td>67,628,034</td>
<td>374,282,380</td>
<td>605,673,233</td>
<td>646,441,735</td>
</tr>
<tr>
<td>%age of total</td>
<td>10.5%</td>
<td>57.9%</td>
<td>93.7%</td>
<td>100.0%</td>
</tr>
<tr>
<td>Average cost per patient (£)</td>
<td>9,056</td>
<td>3,284</td>
<td>1,305</td>
<td>1,118</td>
</tr>
</tbody>
</table>
Where are the patients? (by Local Implementation Group - LIG)

Symphony Project

Somerset patients by cohort and LIG area
Where are the patients? (by hospital isochrones)
“It is of interest to note that there is very little relationship between A&E visits and A&E admissions (despite the seeming linear relationship for the larger values)”

“The Somerset Disease Network

“Our network inference algorithm determines causative relationships between diseases using the Symphony dataset. The algorithm is completely data-driven and does not involve any user input. In this network, each circle represents a disease, whilst an arrow between two circles represents a mathematically-inferred causative relationship”
What’s on the wish list

• **Fast Followers** – others are struggling to repeat (mainly due to IG)

• **Patient Activation Measures** – at scale and liked to H&SC cost

• **More Frequent Data** - Are our local partners able and willing to supply data more quickly than annually

• **How will we measure change** – Getting beyond ‘a baseline’ - will we be able to observe, evaluate and predict when change occurs?

• **Bring it back to ‘episodes’** - what can this data tell us about why episodes occur and how we might prevent them?

• **Better Socialisation of ‘Big Data’** - get away from ‘large datasets’ and back to ‘it being about the patient’
Bringing it all back to the patient

• Commissioned by the CCG to support Outcomes Based Commissioning

• Deeper analysis and patient vignettes to understand ‘who makes up these groups and what is the best way to help them’