

Linking and using Health and Social Care data in Scotland

charting a way forward

Report of a meeting held 22nd May 2014
Nine, The Bioquarter, Edinburgh



Contents

1. Executive summary	Page 4
2. Background	Page 6
3. Work already being done around Scotland	Page 7
4. Results of Group session 1: what do people want to use linked data for?	Page 9
5. Results of group work session 2: barriers and solutions	Page 11
6. Coordinating linkage and data use across Scotland	Page 13
7. Analysis of key themes arising from the day	Page 15
8. Conclusions and future work	Page 17

Meeting steering group

Iain Atherton: Lecturer in Nursing/ADRC Fellow, University of Stirling

Ellen Lynch: Statistician, Health Analytical Services Division, Scottish Government, Edinburgh

Andrew James Williams: Farr Institute Research Fellow, University of Edinburgh

Miles Witham: Clinical Reader in Ageing and Health, University of Dundee

Contact:

Dr Miles D Witham: m.witham@dundee.ac.uk

Ellen Lynch: Ellen.Lynch@scotland.gsi.gov.uk

Attendees

Mhairi Aitken:	Research Fellow in the public engagement stream of Farr Institute
Nick Bailey:	Senior Lecturer urban studies/health and wellbeing, Glasgow
Corri Black:	Senior Lecturer in Public Health, University of Aberdeen
Philip Brown:	Acting Senior Performance and Information Officer, Health and Social Care, City of Edinburgh Council
Duncan Buchanan:	ISD Scotland, Edinburgh
Dette Cowden:	Improvement Service
John Frank:	Director, SCPHRP, Edinburgh
Rachel Hardie:	Consultant in Public Health Medicine, NHS Lothian
Ellie Hothersall:	Consultant in Public Health, NHS Tayside
Nick Jones:	PhD intern, Scottish Government
Andrew Lee:	ISD Scotland, Edinburgh
Ellen Lynch:	Statistician, Scottish Government, Edinburgh
Martin Malcolm:	NHS Western Isles
Danny McAllion:	Senior Information & Research Analyst, Paisley, Renfrewshire
Alison McCallum:	Director of Public Health and Health Policy, NHS Lothian
Janice McGhee:	Senior Lecturer Social work, Edinburgh
Mark McGilchrist:	Senior Research Fellow in Health Informatics, University of Dundee
Christine McGregor:	Health Analytical Services, Scottish Government, Edinburgh
Andy McKeown:	Scottish Consortium for Learning Disabilities
Paul McNamee:	Professor of Health Economics, University of Aberdeen
Susan Murray:	Research Fellow, University of Stirling
Steve Pavis:	Programme Direction (Implementation), Farr Institute, Edinburgh
Julie Rintoul:	Statistician, Scottish Government, Edinburgh
Alasdair Rutherford:	Lecturer Quantitative Methods, Stirling
Anthea Springbett:	NHS National Services Scotland, Edinburgh
Gordon Stevenson:	Glasgow City Council
Fiona Strachan:	Trial Manager, Edinburgh Clinical Trials Unit

N.B. Throughout this document when discussing linked data, this is in reference to anonymised/de-identified data. However, there are a number of instances when using linked data operationally is discussed. In these instances link data is referring to identifiable data for use only by care professionals to coordinate/improve care.

1 Executive summary

Background

Integration of health and social care is the Scottish Government's ambitious programme of reform to improve services for people who use health and social care services. This will require the integration of data sources to support effective work practices and evaluation both locally and nationally. Furthermore, there is increasing recognition inside and outside academia that integrated health and social care data are required to inform service delivery and policy in order to improve public health and wellbeing.

A number of groups around Scotland have started to link health and social care data for academic or analytical purposes. On the 22nd May 2014, a meeting was held with representatives from some of these groups. The key aims of the meeting were as follows:

1. To share experiences from around Scotland on linking health and social care data
2. To better understand what stakeholders want to do with these linked data
3. To map out what needs to be done to allow use of linked data to achieve these ends

Examples

To begin with four brief talks outlined some of the projects – both at a local level and nationally – which are already taking place around Scotland. Rachel Hardie described work being undertaken in NHS Lothian and their partner local authorities, Mark McGilchrist outlined two Tayside

data linkage projects, Julie Rintoul provided an overview of the national social care – housing – health data linkage project, and Iain Atherton described the developing work of the ESRC Administrative Data Research Centre Scotland.

Group sessions

Those present were then split into smaller groups for two sessions to firstly consider the intended uses of linked health and social care data and subsequently the barriers and solutions to achieve each of the uses. Some of the uses, barriers and solutions included:

Mapping trajectories in health and social care

- Barriers: Lack of analytical capacity and knowledge of data sources.
- Solutions: Investment in analytical capacity, improved metadata and communication.

Refining identification of at risk populations

- Barriers: Data quality, data availability (e.g. hospital prescribing)
- Solutions: Communication with data providers, awareness of data availability

Evaluation of interventions

- Barriers: Lack of incentives to evaluate properly, inadequate critical thinking
- Solutions: Guidelines for undertaking evaluations, peer review

Improving data quality in research

- Barriers: Knowledge of data sources, inconsistency of data
- Solutions: Validate against gold standards, communication with data providers

Providing evidence for the integration of services

- Barriers: Consent for data sharing, system compatibility
- Solutions: Public engagement, data standards, single procurement process

Co-ordinating linkage and data use across Scotland

Two brief presentations were given on some of the approaches to data linkage being undertaken in Scotland. Alison McCallum outlined some of the safeguards in place to protect patient/client privacy, whilst enabling valuable research to be undertaken. Steve Pavis provided an overview of the Scottish Informatics and Linkage Collaboration (SILC). SILC is the umbrella body which brings together the ADRC and the Farr Institute in Scotland.

Conclusion and next steps

Some progress has been made nationally and locally to link health and social care data. However there are still some significant challenges, particularly around understanding the data, having the analytical expertise, understanding stakeholder needs, communicating with data providers and the public.

In order to make progress with this agenda, four steps were suggested:

1. Disseminate this report to service providers in health and social care, government and the research community.
2. Establish a network to support those who use / would like to use linked health and social care data for statistical/research purposes.
3. Establish a website which contains resources to support the linking of health and social care, analysis and dissemination.
4. Provide / co-ordinate training to support those who use / would like to use linked health and social care data for statistical/research purposes.

The meeting steering group will progress these steps over the next few months.

2 Background

Integration of Health and Social care is a key public policy priority, both within Scotland and across the world. Such integration is likely to demand not only integration of finances, work processes, staff and cultures, but also integration of data sources to support effective work practices and evaluation, both at an operational level and at the level of service commissioning and public health delivery.

In addition, there is rapidly growing interest in linking health and social care data within the research community – there is a growing realisation that health problems cannot be properly researched without an understanding of social context and vice versa, and that evaluation of interventions in one domain (e.g. healthcare) must take into account effects on other domains (e.g. social care). Linked data are essential for such analyses to be conducted.

A number of groups around Scotland have started to link health and social care data, and initial contact between those groups revealed that a) groups were unaware of the work that other groups were undertaking, and b) similar challenges were faced by all groups trying to link and use health and social care data. This meeting was therefore convened to address these issues, as a first step towards building a community of expertise within Scotland to capitalise on the potential of linked health and social care data to underpin research and service delivery.

key aims

The key aims of the meeting were as follows:

1. To share experiences from around Scotland on linking health and social care data
2. To better understand what stakeholders want to do with these linked data
3. To map out what needs to be done to allow use of linked data to achieve these ends

In preparation for the meeting, attendees were asked to reflect on the following questions prior to the meeting:

- What challenges and barriers have people come up against?
- How have people overcome these challenges?
- What concerns do people have around linking these data sources?
- What questions do service providers, users and policymakers want answered?
- What structures need to be in place for successful linkage and use of linked data?
- What resources/expertise are available? What are the gaps?
- What skills are required in order that others can successfully use linked health and social care data for service planning/research?
- How can we best communicate about data linkage and linked data utilisation?
- How can we best collaborate to take forward this agenda?

3 Work already being done around Scotland

**Four brief talks outlined some of the work already
taking place around Scotland**

A. **Rachel Hardie** described data linkage in Lothian, performed over the last several years as part of Integrated Resource Framework pilot work. Linkage was successful; linked data are now available for 2008-2011, with more linkage planned. Barriers to success included a lack of personnel to work on the linkage, and major challenges remain in understanding what the data means and how best to analyse it. Primary care health data are not currently linkable with the datasets, but ways of linking some primary care data (e.g. from the Quality and Outcomes Framework (QoF)) are being explored. Exemplar analyses around patient pathways, the cost of COPD (**Chronic Obstructive Pulmonary Disease**), and effects of a new intervention for hip fracture are being progressed.

B. **Mark McGilchrist** outlined two Tayside linkage projects; one funded by the Scottish Collaboration for Public Health Research & Policy (SCPHRP) linking health and social care data in Dundee on people aged 65 and over, the other linking health, social care and fire service data on vulnerable adults. Key learning was that building relationships with data providers takes time, that mechanisms to ensure that data providers retain a say in the use of their data was important in building trust, that a key barrier to extracting social care data was the lack of staff time available for organisations to commit to such work, but that deploying external researchers into the social care organisation could overcome this barrier. Major challenges exist in understanding how social care data are coded within social care organisations, and how categories of coding relate to similar categories between different social care organisations.

C. **Julie Rintoul** described work being undertaken by the Health Analytical Services (Scottish Government) and Information Services Division in collaboration with NHS Boards and local authorities. Yearly snapshots of social care data (including care home census, home care use, direct payments, mental health and learning disability services) from 12 social work departments are being linked with healthcare data held by ISD between 2010 and 2014; rollout to all social work departments in Scotland is now being offered, with initial linkage of data for 2014, followed by data from 2010-2013. Planned uses include modelling movement of older people into care homes, improved understanding of costs, and improved data on engagement with mental health services. Key challenges to be overcome include disseminating knowledge on these new datasets, increasing analytical capacity to use these datasets, and finding better ways to feedback results to data providers and service users. Further details can be found at:

<http://www.scotland.gov.uk/Topics/Statistics/Browse/Health/Datalinking/HealthSocialCareandHousin>

D. **Iain Atherton** described the developing work of the ESRC Administrative Data Research Centre-Scotland. His talk outlined how this new centre will provide a focal point for social science research in Scotland. Academics will be able to approach the new centre with a view to their facilitating data linkage projects. Staff will have expertise in administrative data and techniques for linking and analysing. These support staff will enable those with limited or no experience of this kind of work to move projects forward in a timely manner but also ensuring legal and ethical requirements are met. Academics from the centre will be moving forward exemplar studies that

- demonstrate the potential uses of administrative data in research,
- develop new techniques to better utilise data, and
- work with and involve data custodians to facilitate data access and address their needs and concerns.

Finally, the centre will provide a safe haven where anonymised linked data can be held and analysed for the duration of projects so as to ensure data security. Finally, Iain highlighted that the network of centres across the UK will have a website where more detail as to what will be provided and how to address services will be found (<http://www.adrn.ac.uk/>).

4 Results of Group session 1: what do people want to use linked data for?

Group work was undertaken to pool ideas as to what stakeholders wished to use linked data for. Group ideas were then pooled in a plenary session, and the following key themes emerged:

1.

Mapping trajectories in health and social care

E.g. trajectory of increasing dependency in individual older people over time, or longitudinal analyses of early life interventions, but also understanding how the characteristics and needs of groups of people are changing over time

2.

Refining identification of at-risk populations

E.g. people at risk of hospitalisation or care home admission

3.

Evaluation of interventions, especially public health intervention

E.g. data collection to allow evaluation of new service models, and predicting knock-on effects from changes in one area such as a new health treatment to other areas such as the need for social care. Such a use would facilitate 'natural experiments' in health and social care delivery

4.

Improving data quality in healthcare research

E.g. social care outcome measures in clinical trials; measuring currently unmeasured confounders in observational research

5.

Building integrated services

E.g. understanding the total need across both health and social care for individuals and groups

Other important issues that arose during this phase of the discussion included the ability to use linked data to plan for services and to monitor policy; to explore issues around inequality and the impact of this on health and social care; to drive integrated cost evaluations – critical if integrated budgets are to work in practice; using linked data to identify novel interventions for future testing (e.g. medications that might reduce the need for future social care support). The ability of linked data to drive convergence of health and social care culture was mentioned as another potential benefit; if linked data was available in real time at the operational level, this could significantly reduce duplication and redundancy in assessment of clients. It was also noted that many potential uses of linked data will become apparent only when data linkage is commonplace.

5 Results of group work session 2: barriers and solutions

A second round of group work was undertaken to elicit barriers to linking and utilising health and social care data to address the five key areas identified in the first round of group work. Each group was given a single area to address; group ideas were then pooled in a plenary session, and the following key themes emerged:

Barriers

Solutions

1. Mapping trajectories in health and social care

Lack of capacity for complex modelling
Inconsistency in coding of data
Lack of understanding of what data means

Investment encouraging widening capacity, including in academia
Feedback to end users
Develop common coding 'minimum datasets'
Clarity – written metadata used in analysis and by data providers.

2. Refining identification of at-risk populations

Insufficient data quality; inaccurate, missing or unavailable data
Some prescribing data (esp. hospital prescribing) not available electronically

Feedback from users to data providers to improve data quality
Work required to better understand data availability and limitations
Electronic prescribing

3. Evaluation of interventions

Lack of incentives to evaluate properly
Inadequate critical thinking within (national and local) government, NHS Scotland and the private and third sectors

Body with teeth required to enforce need for evaluation
Guidelines for evaluation, resources
Commissioned and independently reviewed evaluations
Engagement of academia – work requires peer review and recognition of academic impact
Better communication of findings

4. Improving data quality in healthcare research

Understanding what the data really codes for
Inconsistency of data

Validate data against gold standards
Build relationships with data providers and users to improve data quality
Database to feedback issues; metadata to describe data

5. Building integrated services

Consent for data sharing across services
System compatibility between services

Education – public engagement
Standards for data collection and coding
Single procurement process

6 Coordinating linkage and data use across Scotland

Two short presentations were given on some of the broader high-level activities outlining the approach to data linkage being taken in Scotland.

A. Alison McCallum began by giving some examples from her experiences in Scotland and Northern Europe of the valuable insights which can be gained from linked data. However, Alison then highlighted that patient rights need to be maintained, and posed the following question: ‘How do you protect personal data, and the lives that it describes, while being unable to measure quality or improve care without it?’ Given the small population of Scotland and the increasing capacity to link data, Alison listed that physical, technological and human safeguards were required to address her question indicating the need for safe havens.

The key features of safe havens were given as:

- Indexing ensuring that personal identifiers were never attached to the data of interest, while also enabling linkage.
- All linked data held in a secure environment
- Access only to named researchers
- All outputs approved by disclosure committee before release

Alison illustrated this approach using the examples of the Scottish Health and Ethnicity Linkage Studies and the Scottish Imaging Network (SINAPSE). She concluded by highlighting that governance and ethics are recognised as crucial to Scotland’s research strategy and thus public and patient engagement is paramount.

B. Steve Pavis built on many of the elements of Alison’s presentation by introducing the Scottish Informatics and Linkage Collaboration (SILC). SILC is the umbrella body above the big data initiatives such as the ADRC and Farr Institute in Scotland.

Steve described the Scottish approach to data linkage research as follows:

- Facilitating research that is in the public interest whilst protecting individuals’ privacy
- Avoiding large data warehouses but ensuring data can be brought together efficiently to answer important research questions
- Creating partnerships and networks across sectors (academia, public and commercial sectors)
- Sharing resources and expertise to create efficient public services

While the legal and ethical responsibilities for the data remain with the data controllers, SILC will provide the facilities and expertise for secure data linkage and use. Steve then illustrated the secure indexing and linking process established in Scotland. The electronic Data Research and Innovation Service (eDRIS) was introduced as the one-stop-shop for those wishing to access linked data. Steve then concluded by listing the aims of SILC:

- To support a Scotland wide strategic approach to informatics and administrative data use.
- To create networks and partnerships across disciplines and sectors.
- To create infrastructure that is coordinated, efficient and sustainable.

7 Analysis of key themes arising from the day

The level of interest in the meeting demonstrated that there is a lot of activity already underway or planned, using linked health and social care data. Furthermore, this interest is evident across sectors (e.g. government, NHS Scotland, academia). The meeting highlighted that there is a lot of expertise across Scotland which can be better utilised to facilitate:

Working relationships across sectors with people who have a variety of skills: In order to understand and maximise the linked health and social care data we will need to communicate with a range of people from multiple sectors and organisations, for example, social workers, health professionals, IT professionals, analysts, data protection officers/solicitors, service planners.

How can (national and local) government, NHS Scotland and other care providers build links with academia?

Good practice for sharing, linking and analysing data to maintain people's privacy: It is pivotal that those involved in sharing, linking and analysing health and social care data follow good practice in order to maintain the trust of service users and the public.

Improvements in our understanding of:

- The projects being undertaken - this should save duplication and effort.
- Data availability,
- Data quality,
- Coding and definitions of data,
- Areas for improvement – for example data gaps such as hospital prescribing, community health.

The identification of where to direct resources (e.g. training, IT, research):

For example, if organisations are expecting to utilise linked health and social care data to improve evidence around service delivery and research, should there be a training programme for analysts? Who should be involved in setting this up? Who should be involved in providing the training? How can we encourage knowledge exchange?

Do we have the necessary infrastructure in place to support the sharing, linking and analysing of linked health and social care data? Can analysts from different sectors get access to the linked data? Do analysts have appropriate software to do their work? How should we prioritise and fund research, especially in times of severe budget constraints?

Further data linkage, for example Scotland-wide coverage**Improvement of our health and social care knowledge more generally:**

There are very few analysts who have in depth knowledge of health and social care. In order to identify what analysis would be useful for stakeholders, analysts will need to broaden their subject knowledge and/or know who to contact for further specific support.

Development of our analytical skills:

Analysts from national and local government, NHS Scotland, and other organisations involved with social care are being asked to work with – for the first time in many cases - multiple datasets, from multiple sectors, and undertake complex analysis. Analysts will need appropriate training and infrastructure.

In academia there are not many researchers who have used linked health and social care data, this is an area where there are opportunities to develop knowledge and expertise.

Analysts from national and local government, NHS Scotland, other organisations involved with social care and academia all have a role to play in knowledge exchange to facilitate each other's work.

Enhanced research potential:

Making linked health and social care data accessible for researchers can greatly enhance Scotland's research potential. This will hopefully lead to improving treatments and services for service users, for example, understanding social care outcomes measures in clinical trials, risk prediction tools.

Improved evidence for service planning, evaluation and policy development:

Analysis based on linked health and social care data should greatly improve our evidence for service planning, for example, understanding patient/client pathways, impacts of interventions, ensuring staff have the right skills to deal with patient/client needs, informing the spending of public resources. Furthermore, NHS Boards and local authorities will need to use health and social care data in their strategic plans to support integration.

The dissemination of statistics/research to a wider audience (including the public and service users) in a range of suitable formats:

A wide range of stakeholders will be interested in using the statistics/research produced from the health and social care data. Consideration should be given to the wide range of stakeholders and that dissemination might need to take a variety of formats. Dissemination should include data providers as this often helps improve data quality if people can see the consequences of their time and effort.

Service users and the public should also be included in dissemination plans. It is important that analysts/researchers are transparent in what they use data for, and how it is for the wider public good.

10 Conclusions and future work

Although some progress has been made regarding the technical ability to extract and link health and social care data, significant challenges remain. These include the need to ensure that the uses to which the data will be put are well defined, that the provenance of the data are understood, that the quality and limitations of the data are defined. In addition, appropriate analytical methods (which may need to include data mining of free text) need to be developed and validated, and that appropriate mechanisms are put in place for sharing methodological expertise, disseminating results, and feeding back current gaps in data understanding and quality to end users and data providers.

In order to make progress with this agenda, four steps were suggested:

1. that the results of this symposium will be disseminated to service providers in health and social care, to government and to the research community;
2. that those in attendance at the symposium form the nucleus of an interest group to allow pooling of knowledge and skills in this area, and as a contact

point for those interested in using linked health and social care data;

3. that a website and resource centre be constructed to provide a focus for information and future efforts in this area; possibly hosted by the Administrative Data Research Centre (ADRC)
4. that training events, facilitated by ADRC and the Scottish Government, on emerging datasets (e.g. the linked national health & social care dataset) could be organised to make the research and service communities aware of what data resources are becoming available.

The meeting steering group undertook to progress these steps over the next few months, and it is likely that a further meeting in approximately 1 year would be useful to review progress and set new directions for activity.