Adelaide Primary Health Network and Australian Healthcare & Hospitals Association (AHHA)

Data Collaboration Network meeting

Tuesday, 28 July 2015
SA Health Medical Research Institute (SAHMRI)

Chris Radbone, Associate Director, SA NT DataLink

The Australian Government has provided financial support to SA NT DataLink through the National Collaborative Research Infrastructure Strategy (NCRIS).
Who is SA NT DataLink?
The Australian Government has provided financial support to the PHRN and its members through the National Collaborative Research Infrastructure Strategy (NCRIS).
What is Data Linkage? SA NT DataLink Animation
What is Data Linkage – Record Linkage?

- Secondary use of linked administrative data is often referred to as ‘data linkage,’ ‘record linkage,’ or ‘linked data.’
- This is typically population based longitudinal data that has originally been collected for another purpose, but can include integrating surveys.
- Linkage may take place across data sets in a single domain (i.e. health) or across domains (i.e. health, education, social services, early childhood.)

Why is Data Linkage Important?

- Lack of ‘unique identifier’ for each member of the population
- Data linkage is an invaluable tool for enabling population wide analysis and research. It provides a completely unbiased picture of the entire population, is cost-effective relative to other data collection mechanisms, and enables studies to be done that could not otherwise be performed.
- Enables smaller targeted populations, study cohorts (eg consenting studies) to be evaluated and enhanced by comparing and integrating data.
What do Data Linkers do?

ie Members of the PHRN

SA NT DataLink along with other PHRN member organisations, facilitates the ability to perform data analysis and research using integrated data by initially linking demographic data across multiple government agencies and organisations, including across jurisdictions, to support:

• Policy analysis
• Program / Operational Evaluation
• Academic Research

SA NT DataLink are establishing a secure ‘Data Repository’ similar to the leading international Data Linkage facilities; particularly Ontario, Manitoba in Canada, and Scotland and Wales in the UK.
What does SA NT DataLink do?

"Good work ........but I think we need just a little more detail right here!"
What does SA NT DataLink do?

Two key roles of SA NT DataLink

1. Client Services

- Consult on study design; facilitates data access and/or approvals for researchers between Data Custodians and/or HREC
- Negotiate access to agency datasets for research projects
- Seek required Privacy Committee approvals for access to data that can be used for specific research projects
What does SA NT DataLink do?

Two key roles of SA NT DataLink

1. Client Services

- Consult on study design; facilitates data access and/or approvals for researchers between Data Custodians and/or HREC
- Negotiate access to agency datasets for research projects
- Seek required Privacy Committee approvals for access to data that can be used for specific research projects
What does SA NT DataLink do?

Two key roles of SA NT DataLink

2. The Data Linkage Unit

- Receives individual level demographic data from agencies
- Provides the matched demographic data and the Project Specific Linkage Keys* to Data Custodians for linking to their agency service data
- Provides feedback to Data Custodians on data quality issues in the data they have provided as part of this matching process
- Collates and holds the identifiable demographic data and their PSLKs as part of the Master Linkage File

* Randomly generated alphanumerical keys associated with the same matched records of an individual from across the same or different agencies. These later replace the personal identifier in the data provided by Data Custodians to approved researchers.
Data and Research Governance
Security and protection of privacy

SA NT DataLink legal obligations :

- The *Commonwealth Privacy Act 1988*, Section 95 ethical use of data.
- The *Northern Territory Information Act 2002*.
- Only SA Health employed staff employed who are then governed by the *Public Sector Act 2009* can access the data held in the Data Linkage Unit.

SA NT DataLink’s security protocols :

- South Australian ‘Information Security Management Framework’
Security and Privacy Protection

The Separation Principle

1. Data Custodians
   - Datasets
   - Only data custodians see the dataset as a whole

2. SA NT DATALINK
   - Performs Linkage Process
   - Output: Project Specific Linkage Keys (PSLK)
   - PSLK + Record ID

3. RESEARCHERS
   - Merges and cleans the data in preparation for analysis.

4. Review and Feedback Process
Available datasets

SA NT DataLink can facilitate access to growing number of datasets, these include:

- SA Hospital Separations (ISAACS)
- SA Emergency Department
- SA Cancer Registry
- SA Dental Service (Titanium)
- SA Perinatal
- SA Women’s and Children’s Health Network - Child Health Checks¹
  - Midwife Network Universal Contact Visit
  - Family Home Visiting for Mothers and Infants at Risk
  - Now mother’s Day Service
  - Blue Book Child Health Checks
  - Childhood Immunisation registry
- SA Electoral Roll
- SA Child Protection
- SA Public School Enrolments Census
- SA NAPLAN
- Years 1-3 Reading Assessment
- English as a second Language
- Housing SA
- SA Birth Registry
- SA Death Registry
- Australian Early Development Census
- NT Government Client Master Index
- NT Death Registry
- NT Birth Registry
- NT Perinatal (Trends)

SA NT DataLink is not ‘currently’ a Data Repository – but provides the linkage to these datasets.
Available Datasets

Inventory of Linked Data - As at 9th February 2015

Data Capture (yrs)


NT Client Master Index
NT Death Registry
NT Birth Registry
NT Perinatal (Trends)
NT School Enrolments
NT Immunisations
NT Inpatient Activity
NT Child Protection
SA Hospital Separations
SA Emergency Department
SA Cancer Registry
SA Dental Service (Titanium) ¹
SA Perinatal
SA Perinatal (by Mother)
SA Child Health Check
SA Cervical Screening
SA Electoral Roll ²
SA Child Protection
SA Public School Enrolments Census
SA NAPLAN
Housing SA
SA Birth Registry
SA Death Registry
Australian Early Development Census
ANZData ⁴

Data received for Project Specific linkage (external to MLF):
- SA Pathology (IMVS)
- SA Private Pathology, including Healthscope Pathology SA (Gribbles), ClinPath, Gramps Pathology, Adelaide Pathology Partners
- Whole of a population collection
- NT School Enrolments; NT Immunisations

To be provided by Data Custodian
- Linked
- Currently linking
- Received, lined up for linkage
- Linked - Project Specific

¹ Data linked only for 1999-2005 birth cohort
² Data linked is only for SA and NT
³ Cross sectional extract containing currently enrolled residents
⁴ Data access agreements for inclusion to the MLF, being finalised
I didn't have any accurate numbers so I just made up this one.

Studies have shown that accurate numbers aren't any more useful than the ones you make up.

How many studies showed that?

Eighty-seven.

Source: Dilbert.com  8th May 2008
Metadata and Data Quality

Data Quality

- Quality of Linkage – *for Researcher’s benefit in publishing papers*
- Quality of Linked Data – *working with Data Custodians*

Metadata

- Responding to the demand from Researchers and Analysts for information on researchable variables
- Use of online Metadata tools ie ANDS, METeOR
- Currency and completeness – *Data Quality Statements*
Key Challenges and Future Directions

Key challenges

- Timeliness of Data Custodian extraction and Ethical approvals
- Capacity and resourcing from all parties to respond to the demand from Researchers and Government for linked data
- Increased workload, reduced resources for Data Custodians and their agencies – *data support being seen non core work?*
- Improving the Efficiency of providing de-identified linked data

Change / evolution of the SA NT DataLink system

- Establishing a ‘Custodian Controlled data repository’ including maintaining *separation and security*
- Secure Access - online or physical access to de-identified linked data through SA NT DataLink. Data Integration Unit, SURE.
- South Australian legislation to support Data Custodians releasing Personal: Medical and other sensitive data
Data Linkage Project Examples

South Australia: Colorectal Cancer – *Early detection, clinical management & survival outcomes at a population level*

Australian: Cross Jurisdictional – *In-hospital and post-discharge mortality: learning about quality of care using data linkages from Australian States (Proof of Concept 1)*

International: UK  *Routes from diagnosis – a detailed map of cancer survivorship*
Data Linkage Projects

South Australia: Colorectal Cancer – Early detection, clinical management & survival outcomes at a population level

The research project linked information from the South Australian Cancer Registry, hospital separations, radiotherapy services, hospital-based cancer registry systems and data on the socio-demographic, clinical and treatment characteristics for 4641 CRC patients, aged 50–79 years, diagnosed from 2003 to 2008.

The study demonstrated the utility of data linkage for population-wide assessment of CRC management and highlighted limitations for measuring detailed aspects of clinical care.

This was possibly the first Australian study to examine broad patterns of CRC management (surgery, radiotherapy and chemotherapy) according to site and stage, with entire statewide coverage of both public and private health care services.

About the methodology, it concluded that linkage studies are likely to become important tools for conducting population-wide evaluations of cancer care. Establishing methods to determine accuracy of data linkage and the impact of linkage errors on study findings also requires further development.

Treatment patterns among colorectal cancer patients in South Australia: a demonstration of the utility of population-based data linkage - Kerri R. Beckmann MPH, Alice Bennett FRACP, Graeme P. Young FRACP MD and David M. Roder MPH DDSc
Cancer Data Linkage Projects

Australian: Cross Jurisdictional — *In-hospital and post-discharge mortality: learning about quality of care using data linkages from Australian States (Proof of Concept 1 - PoC#1)*

The project was designed to:

- develop and test data acquisition and linkage processes needed to support ongoing cross jurisdictional research; and

- demonstrate the value of these linkages in informing a public health issue of national importance; in this case *Hospital Group Standardised Mortality Ratio* (HGSMR) estimation, which is an important indicator of hospital performance.
Cancer Data Linkage Projects

Australian: Cross Jurisdictional — *In-hospital and post-discharge mortality: learning about quality of care using data linkages from Australian States (Proof of Concept 1 - PoC#1)*

PoC#1 successfully linked de-identified WA, NSW and QLD and SA public and private hospital separations data from 2004 to 2008, and also linked mortality data for the 12-month period following hospitalisation over these years.

Key preliminary findings are:

• significant issues about data variability and access need to be addressed.

• that to include deaths 30 days post hospital separation in the HGSMR estimation, tended to reduce the variability in hospital mortality rates between hospital peer groups and across jurisdictions; and

• the reduction in HGSMR variability was mostly attributable to within state data linkage as the proportion of cross-jurisdictional patients was relatively low. Cross-jurisdictional data linkage showed a further small reduction in variability.

The PHRN is expected to release the findings in the near future.
Data Linkage Projects

International: UK  *Routes from diagnosis – a detailed map of cancer survivorship*

Linked data from almost 85,000 patients using the UK National Cancer Data Repository

Mapped patient ‘journeys’ over a 7 year period from diagnosis using set of clinically defined outcomes

Complete national database of patients with tumours diagnosed in 2004, tracking their inpatient hospital activity over the seven-year period.

The cancers included breast, lung and prostate cancers, in addition to rarer brain and central nervous system tumours about which less was known.

Demonstrated that the technology and data is already available and it is possible to understand cancer journeys at an unprecedented level of detail.

By linking routinely collected clinical data it provided an evidence-based framework to apply a systemic approach to cancer care services, system design and evaluation, policy formulation and to inform the direction of research.

[http://www.macmillan.org.uk/Aboutus/Ouresearchandevaluation/Programmesofwork/Routesfromdiagnosis.aspx](http://www.macmillan.org.uk/Aboutus/Ouresearchandevaluation/Programmesofwork/Routesfromdiagnosis.aspx)
National Data Linkage Infrastructure: *including Online Application process*

National remote access secure data analysis platform


Statistical Data Integration involving Commonwealth Data


SA NT DataLink – ‘For Researchers’ section of the website [www.santdatalink.org.au](http://www.santdatalink.org.au)

SA NT DataLink Access and Pricing Policy
Chris Radbone
SA NT DataLink, Associate Director
(08) 8302 2777
chris.radbone@unisa.edu.au

Client Services:
santdatalink@unisa.edu.au
Adelaide: (08) 8302 1604
Darwin: (08) 8985 8011

www.santdatalink.org.au