SURE: Helping to get the most out of routinely collected data

Registry Special Interest Group Meeting

9 October 2015

Outline

1. The role of the Sax Institute
2. Why SURE was developed
3. How SURE is being used
4. Future developments
Sax Institute: Purpose built enabler

- **Mission**: To improve health and wellbeing by driving the use of research in policies, programs and services
- **Research enabler** that creates **assets** and solutions to support policy relevant research
  - Enables conduct of robust and high impact research
- **Research broker** that acts as a bridge between policy agencies and researchers to:
  - Generate new research findings that are more useful to policy agencies
  - Increase policy makers’ access to existing research findings

Sax Institute: Purpose built enabler

- Not-for-profit company limited by guarantee
- Has expert research and policy staff
- Independent from any one university or research group
- 45 member organisations - research groups (public health and health service) and universities
- Has assets and systems developed and tested over 13 years
- We have worked with more than 730 researchers
- Worked with more than 80 policy and program agencies
Our research assets

• **SURE** (Secure Unified Research Environment) - Australia’s first and only remote access platform to access and analyse large amounts of sensitive, routinely collected data without compromising privacy and security

• **45 and Up Study** - Australia’s largest long term cohort study

• **SEARCH** (Study of Environment on Aboriginal Resilience and Child Health) – long term cohort study on the health and wellbeing of urban Aboriginal children

What is SURE?

• **A powerful computing solution bringing researchers together to answer important health questions**

• Secure, purpose-built environment for the analysis of large and complex datasets
  • open new possibilities for research
  • facilitates linked data research

• Launched in 2012, developed by the Sax Institute with funding from the Australian and NSW governments as part of the **National Collaborative Research Infrastructure Strategy (NCRIS)**
Why do we need SURE?

1. Australia has a comprehensive range of routinely collected data across the health and human services sectors – valuable resource
   - to make best use of data, need to address data providers’ concerns over data confidentiality and information security
2. Research collaborations are increasingly multidisciplinary, often teams have researchers from multiple organisations in diverse locations across Australia and internationally.
3. Researchers are undertaking increasingly large and complicated analyses using multiple linked data sets but they need more advanced and powerful analytic environments

Safeguards in linked data

- researchers agree to abide by data custodian and ethics committee conditions:
  - data sets with only specific variables approved for each study
  - no attempts at re-identification
  - keep the data safe and restrict access to it
  - strong incentives for researchers to comply
  - technically easy to not comply
  - both researchers AND their computing environments are assumed to be trustworthy

However, there are a number of risks, particularly in computing environments
Consequences of data breaches

- Often, data breaches are not malicious but the result of accidental or unintended disclosure
- Consequences are still very serious
  - employee termination
  - restricted access to data inhibiting research and analysis
  - Example: Ministry of Health, British Columbia, Canada

How does SURE work?

- SURE provides project workspaces that are accessed remotely over encrypted internet connections (using Citrix)
- members of a research team have shared access to project workspaces
- each project workspace exists in its own security perimeter within SURE so that data from different projects cannot be combined
- all data storage and analysis happens on dedicated SURE servers housed in highly secure commercial data centres
- movement of data in and out of SURE is carefully controlled via specially designed software called the Curated Gateway
Curated Gateway

- provides the only way for files to move in or out of SURE (no Internet, no printing, cannot copy/paste, cannot copy to removable media)
- files to be used in SURE are uploaded to the Curated Gateway and are subject to review before being made available within SURE
- outbound files with research output and results files are also subject to review
- all files that pass through the Curated Gateway are logged and are subject to audit
Strong authentication:
- username
- passphrase
- physical token
- digital certificate
Who is using SURE?

- 50 active study workspaces involving 140 users
  - most use state/territory data, Commonwealth data or cohort data
  - Users based in NSW, ACT, South Australia, Queensland, West Australia, Victoria and overseas in academic, government, non government and private sector organisations
- Over 290 people have completed the SURE training program – mandatory for users

SURE growth

[Graph showing SURE usage from 2013 to 2015 with bars for workspaces and users]
Use of SURE

- Data custodians include the Australian Institute of Health and Welfare, health departments, clinical registries and the 45 and Up Study
  - More than 30 data collections used by SURE researchers
  - Most frequent use of SURE involves linkage of routinely collected data to our 45 and Up Study
- Provides a rich data set for analysis

Significance of 45 and Up Study

1) Australian-based data
2) Great tool to look at health and associated factors over time
3) Representative of the community sampled – 10% of NSW population, over-sampled in rural areas and among Aboriginal and Torres Strait Islander people
4) Self-reported data with consent for linkage to routinely collected data from a wide range of sources.
Partnerships and value-adding services

- Partner to AIHW Data Integration Services Centre
- ANZICS and ANZSCTS data registries have set up workspaces to allow researchers to access their data for linked data studies without compromising information security
- Hosting insurance data used in the Health Market Quality Program of CMCRC
- Only external access mechanism for the Pathways of Care longitudinal study managed by NSW Families and Community Services
- Facilitating access for approved researchers based overseas
  - Users in US, Netherlands and Scotland

User feedback

- "It's great having access to our study data even when I'm working offsite or at home. The added flexibility this brings will really help speed up our analyses and keep the project on track"
  
  Mr Michael Falster, Research Fellow, University of NSW
- “I think SURE is securing the future of linked health research in this country. It has been the beacon and shown the way forward. It is an absolute game changer.”
  
  Dr Anna Kemp-Casey, Research Fellow, University of Western Australia
Recognition of SURE

“SURE is just one excellent example of the type of advantage that Australia has been building through the development of research data infrastructure, which spans data management and reuse, storage, generation and tools for analysis and interpretation. Not only are developments of this kind of national importance, in many cases they are world-leading.”

Dr Ron Sandland, former CSIRO Deputy Chief Executive; Chair, Commonwealth Research Data Infrastructure Committee; Chair, Australian National Data Service

Research findings

• Examples of research findings made possible through the use of SURE and the 45 and Up Study:
  – Among people over 45, overweight and obesity counts for one in every six days spent in hospital and about one in every $6 spent on hospitalisation in Australia, adding up to nearly $4 billion a year (R Korda et al, ANU)
  – Management of diabetes through primary care results in a 30% drop in hospitalisations in a one year period (E Comino et al, UNSW)
  – More than 50% of women with breast cancer stop their treatment within 5 years, despite strong evidence that they should take the medication for much longer to reduce risk of cancer recurrence (Kemp-Casey et al, UWA)
Future developments

• Encourage a broader range of custodians and researchers to use SURE
  – Supporting ready access for researchers to a range of data sets to get the most out of data
• Sax Institute is keen to create more research assets to facilitate robust and high impact research
  – Biobanking – great potential for the use of blood samples collected on 45 and Up Study participants, particularly through linkage to 45 and Up cohort data and routinely collected data sources

Discussion and questions
CONTACT

Jo Khoo
02 9188 9559
jo.khoo@saxinstitute.org.au