

Productivity Commission Inquiry into Data Availability and Use

Robert Wells
Deputy CEO
29 July 2016

The Sax Institute welcomes the opportunity to make a submission to the Productivity Commission in relation to its current inquiry on Data Availability and Use.

As an organisation that assists policy agencies and service providers to use research evidence and expertise, the Sax Institute strongly supports greater access to big data holdings, both public and private, and believes that there are many benefits that will arise from the greater use of these data.

About the Sax Institute

The Sax Institute is a national and international leader in promoting the use of research evidence in health policy and program work. The Sax Institute has developed expertise and systems to support decision makers use the best evidence about what works when they are designing and evaluating policies and services to improve health. We partner with a large network of public health and health services experts to bridge scientific rigour and real-world policy practice. Examples of our work are available on the Sax Institute website www.saxinstitute.org.au

The Sax Institute has a sophisticated understanding of the Australian health system with 13 years' experience in supporting the use of research and data in decision making. We work with over 80 agencies at the federal and state levels (e.g. Commonwealth Department of Health, Australian Institute of Health and Welfare, Therapeutic Goods Administration, Cancer Australia, Australian Commission on Safety and Quality in Health Care, NSW Ministry of Health, NSW Treasury). The Sax Institute is an independent not-for-profit organisation with extensive networks among policymakers, practitioners and researchers. We currently have 46 member organisations comprised of public health and health services research groups and universities. We work closely with our members to draw on their expertise in diverse health and related disciplines.

One of the key objectives of the Sax Institute is to build and maintain sustainable research assets. These assets provide a foundation for innovative research that inform policy and practice and make better use of existing data and resources. These assets are designed for broad access by the research and policymaking community, rather than tools developed solely for the use by one research group or university. The Secure Unified Research Environment is one of our key research assets which will be described further in the section below. Other key assets of the Sax Institute include the 45 and Up Study, the largest ongoing study of healthy ageing in the Southern hemisphere with more than a quarter of a million participants, SEARCH, the Study of Environment on Aboriginal Resilience and Child Health, which is a unique resource for understanding the causes of ill health in urban Aboriginal children, and for developing and implementing strategies to improve their health.

About the Secure Unified Research Environment

The Productivity Commission approached the Sax Institute in June 2016 to speak about the Secure Unified Research Environment (SURE). Detailed information was provided during this discussion about the design, features and usage of SURE and this section provided a brief summary as well as a response to further questions requested by the Commission. SURE has been purpose-built to facilitate access to large, sensitive, human datasets to support better use of our national knowledge base.

Funded through the National Collaborative Research Infrastructure Strategy as part of the Population Health Research Network, SURE has been providing services to researchers and data custodians nationally since 2012 and has rapidly expanded its number of users, research projects, data collections and fields of research supported.

Due to privacy concerns relating to the types of datasets that users access, SURE has not been established as a data repository. Specific workspaces are set up to accommodate discrete projects or programs of work with users able to access and analyse the data without it been stored on their local computing environment. Only data specific to those projects are stored in SURE. SURE provides a high-security environment, that isn't dependent on a researcher's workplace, which facilitates the use of data from different custodians and also the collaboration of researchers working across multiple institutions, including overseas-based researchers. SURE takes a multi-faceted approach to managing data security including controls related to people, technology and processes including:

- **Curated Gateway** – the control point for the movement of data into or out of the facility. Copies of all files must be approved before they are made available within, or outside the facility. Periodic audits are taken of a sample of these files.
- **User authentication** – multiple authentication measures are in place including unique username, password, token and client digital certificate.
- **Secure data storage** – data are stored in servers located in a high security commercial data centre in which there are strict access controls and continuously staffed surveillance.
- **Breach and incident management** – all users are required to undertake mandatory training on information security and sign agreements that comply with Australian research standards and the Australian Privacy Principles.

SURE currently hosts data for over 60 projects and has more than 300 researchers involved in projects using the facility. Data extracts analysed in SURE come from a wide range of sources including state, territory and national government departments and non-government data sources:

- Hospital admissions including emergency departments
- Mortality data
- Births registration
- Electoral roll data
- Cancer registries
- Communicable disease registries
- Perinatal data
- Immunisation registries
- Cohort studies – 45 and Up Study, Australian Longitudinal Study on Women's Health, Pathways of Care Longitudinal Study
- Clinical registries – Australian and New Zealand Intensive Care Society CORE Data, Australian & New Zealand Society of Cardiac & Thoracic Surgeons
- Medicare Benefits Schedule
- Pharmaceutical Benefits Scheme
- Australian Early Development Census

Collection and release of public sector data

Improving policy and service delivery in our health and social sectors requires ready access and analysis of a wide range of data sources. While great strides have been made since the establishment of the Population Health Research Network, we need to ensure that access to data in the broader human services fields is also part of these developments.

Gains in the health sector have been facilitated by the existence of specific guidelines under the Privacy Act relating to the use of personal information in health and medical research [1, 2]. In other areas, where such guidelines do not exist, there is work to be done in developing a clear set of standards or guidelines that exist for the application, assessment and approval of access to data for the purposes of research and analysis, both within the public sector and beyond it.

It has been our experience, that the assessment of potential benefit of the data access is not given the same weight as the assessment of risk. This may be due to the historical lack of interaction between researchers and policymakers that results in poor alignment between government policy and research aims but there is great opportunity for this gap to be closed in the future with greater data availability and research that would then feed back into government policy and services. Research into the acceptability of data linkage in the area of vaccine surveillance research, found that a majority of parents supported data linkage for vaccine safety surveillance (94%) and trusted its privacy protections (84%) [3]. The increasing collection and use of personal information in social media and information technology, more broadly, means that community views on this issue are evolving. A strong system of data governance and security in research using public sector health data has been developed and, when clearly communicated, this should have an impact of the acceptability on the greater availability and use of this information for research.

Access to private sector data

While there is a great wealth of public sector data that could be used to research important policy questions, there is also benefit that could be derived from improving availability and use of private data sets.

In the health sector, which we are most familiar with, this includes private health insurance data, electronic medical records (such as those held by general practice), pharmacy and pathology data.

Data linkage

Australia leads the world in the linkage and analysis of population health data; this provides an opportunity to really understand key drivers in the health system and to improve care and efficiencies. Government is increasingly interested in using these data for research to inform its work; indeed the Senate Select Committee on Health recently published its finding on big data noting that:

A recent estimate by Lateral Economics suggests that Australian government held health-specific data alone could contribute up to \$5.9 billion per annum across the economy. Data linking is the bringing together of two or more de-identified datasets to create a new, richer dataset. Using data linkage techniques, researchers, clinicians and government administrators can deepen their understandings of the ways people use the healthcare system while maintaining patient privacy. This has the potential to inform government policy making and decisions about improving health service delivery [4]

The quote above can be extended to many areas of public policy making where data linkage is required to fully understand and analyse the problem. This is of course true in the health sector, with the different Commonwealth and State and Territory responsibility for health administration, funding and service delivery but is also true of many other areas including family and children services, education, housing, transport and the environment.

Resource costs of access

Greater availability and access to data will have resource implications for both public and private sectors. This resource cost includes the need to build capability to curate data and make it available. Processes such as the development of data dictionaries and preparing publicly available samples of data take time and effort but ensure that high quality research can be completed – minimising the risk of researchers misinterpreting data, which has been previously raised as a concern in releasing data. There is a need for greater collaboration between the data custodians and those using the data so that the needs of each group can be better understood and skills shared where possible. The United Kingdom's Administrative Data Research Network provides a good model of how this can be implemented (www.adrn.ac.uk).

Skilled labour is also an issue in the analysis of these data sets. While there can be great advances made in the health and social scientists by improved collaboration with skilled computer and data scientists, this does not remove the need for researchers that specialise in health and social science as they have the knowledge and expertise in how to pose research questions, interpret findings and understand many of the nuances in the way historical policy processes affects the collection of data.

References

1. National Health and Medical Research Council (2014). Guidelines under Section 95 of the Privacy Act 1988. Accessed 13 July 2016 at: <https://www.nhmrc.gov.au/guidelines-publications/pr1>
2. National Health and Medical Research Council (2014). Guidelines under Section 95A of the Privacy Act 1988. Accessed 13 July 2016 at: <https://www.nhmrc.gov.au/guidelines-publications/pr2>
3. Berry, J. G., Ryan, P., Duszynski, K. M., Braunack-Mayer, A. J., Carlson, J., Xafis, V., & Gold, M. S. (2013). Parent perspectives on consent for the linkage of data to evaluate vaccine safety: a randomised trial of opt-in and opt-out consent. *Clinical Trials*, 10(3), 483-494.
4. Commonwealth of Australia. Senate Select Committee on Health: Big health data: Australia's big potential.