

Productivity Commission Inquiry into Data Availability and Use

Submission by
BetterStart Child Health and Development Research Group
School of Public Health, The University of Adelaide
July 2016

Scope of our submission: Early Years Data Systems and a Proof-of Concept Example

1. Our submission is concerned with **early years** (under 10) data systems and how they can be used to better inform policy, practice and intervention innovation.
2. To achieve this we provide a **concrete ‘proof-of-concept’ example** of a functioning public good data resource comprising a large number of government administrative sources built and used in partnership with the SA Government.

We are a research group in the School of Public Health at the University of Adelaide led by Professor John Lynch with interests in early childhood health and development. The group has backgrounds in epidemiology, public health, nutrition, paediatrics, biostatistics and psychology. The *BetterStart* group has many years of experience in using public sector data in data linkage projects.

SA Early Childhood Data Project: A Functioning Academic-Government Research Data Platform

The South Australian Early Childhood Data Project (SA ECDP) has been built over the past 6-7 years into one of the most comprehensive de-identified linked data resources in Australia. It captures the experience of more than 280,000 children encompassing more than 30 different administrative datasets by linking development, education, child protection, health, and welfare services information. It contains information on children born from 1999 onwards and follows these children through government systems from before birth to adolescence. Since 2009 we have used competitive grant funding, largely from the NHMRC, to support building the ECDP with all the necessary research ethics approvals. The ‘joining’ up of these data across government agencies offers new opportunities for examining a broad range of child health and development outcomes.

Our Goal

The intention of the ECDP has always been for the data to be used as a **‘public good’ resource** to ensure the community benefits from use of their data. We believe public sector data should, under the correct conditions, be available to be used for public good purposes in ‘safe-haven’ environments that are operated through **trusted academic-government partnerships** to inform policy and practice relevant research to improve child outcomes in the community.

BetterStart is a functioning example of a safe-haven that works collaboratively with government agencies including the SA Department of Premier and Cabinet, the Child and Family Health Service, the Women’s and Children’s Health Network, and the Department for Education and Child Development to use the ECDP to inform policy and service provision.

We are now also engaging with community groups including Together SA, to inform what key child health and development indicators are monitored at a local community level. The ECDP is also a key platform contributing to priorities within the NHMRC EMPOWER Centre of Research Excellence - a unique interdisciplinary collaboration focussing on early life prevention of problems that disrupt optimal trajectories of child health and development for disadvantaged children and families.

Potential

A business intelligence system for ECD investment

There is ample evidence on the effects of early life experiences on health, development, and education, and later effects on labour market experience, employment, income and living conditions. The policy and practice question is how do we better support children and families who experience early life disadvantage in its myriad forms?

An early life business intelligence system could help inform Commonwealth and jurisdictional investments.

Due to our Federal system and the funding arrangements that exist in Australia, state and federal governments often have unconnected priorities in the early childhood space. This can mean that federally funded programs may not be coordinated with state systems. This then misses out on the chance of delivering coordinated programs that are designed to leverage off existing state resources and investments. A whole-of-Australia business intelligence system focussing on early childhood could inform a coordinated approach to targeting policies and programs to children most in need and potentially evaluate program effects on child outcomes across the country.

The importance of Commonwealth data

There are some high-value Commonwealth datasets that are extremely difficult to gain access to. This includes federal government welfare and healthcare data (Centrelink, MBS, PBS). The value of this data is that it directly informs costs of negative experiences or development trajectories in early life to the community and the government in later life. This type of information is integral to building a business case for early intervention to break what can be intergenerational cycles of disadvantage. However, it must be paired with jurisdictional data to leverage maximum value – both levels of government invest in supportive services for early child development, so it is the joining of jurisdictional and Commonwealth data that offers the greatest potential for a national investment business intelligence system.

Academic - Government Partnerships

Bringing researchers and government together in **true partnerships** enables linked administrative data to be used in a way that is most likely to benefit the broader community. There are such models in Wales and Scotland where a suitably qualified university partners with Government to hold that data- i.e. it is held in the university environment. They are then adequately resourced to interrogate data for government and other relevant parties. Government provides the service and policy expertise that directs avenues of investigation, while researchers provide high levels of analytic expertise in how to use data to answer relevant questions. Such analytic capacity is not necessarily available in government. The use of the ECDP in SA enriches the resources government can access to evaluate or target programs as it integrates government information which normally exists in isolation, and is difficult to link internally. Ultimately, it enables whole-of-government buy-in to obtain evidence to support effective programs and policy development to improve investment in early childhood- and most importantly, outcomes for children.

The importance of private sector and non-government data

Private sector and non-government data including private hospitals, independent schools and not-for-profit organisations, should also be available to be used for public good purposes. These data have significant potential to improve the ability for evaluative research to inform program funding, design and targeting. A great deal of service delivery in early childhood is done by the non-

government sector, but the governments who fund such work have little idea of the value of that public spending because non-government agencies have little capacity for effective evaluation.

One of the common benefits cited in regards to using linked government data is that it is 'whole-of-population'. However, this is not true when using data from health and education systems. The private and public divide in systems and data availability means there is a 'black-hole' for ~ 30% of the population. This creates significant gaps for research attempting to understand the precursors of poor health and development outcomes in children. There needs to be more government advocacy and support for integrating private and non-government sector data into data linkage systems if we really want to generate a 'whole-of-population' view.

Challenges

Culture – moving from “why to why not?”

Making government data more freely available is not a new idea. However, the current culture of some parts of government may not be entirely comfortable with such an idea. Within government systems there are 'gatekeepers' –in some cases literally single individuals - who provide their own interpretation of data requests without broader consultation with managers and executives, and control data release in a risk-averse way that impedes the use of government data to inform intervention research and evaluation of government funded services. Extensive ethics approvals, project governance structures and consultations are not sufficient to alter risk-averse approaches. This can result in concerns that obstruct public sector data access, with the most often cited reason concerning re-identification of individuals. There needs to be reassurance from government departments that a more open data policy means that if there is a public good purpose for sharing the data with trusted researchers - the attitude to releasing data for public good use should move from 'why' to 'why not'?

Resourcing of data departments within government also negatively affects the culture of data sharing for public good. There is currently a significant burden on data providers, and resourcing needs to be improved to encourage a more open attitude to data sharing. An alternative option to reduce the burden on data providers is a safe haven repository so that data is obtained once from custodians and updated regularly. This data can then be made available to relevant parties.

Data linkage agencies: moving from supply side to demand side concerns

Probabilistic data linkage is central to public good data use. It needs to be properly funded and coordinated. However, crucial to the success of the data linkage process is to have the data curated by people who actually know the data and how it might be used for public good. Currently, linkage agencies are perhaps understandably supply side oriented – they cater mainly to the perceived needs of data custodians. In the future they will need to be actively supporting demand-side activities – actual use by academic-government partnerships to improve service delivery for public good.

These linkage agencies care less about and are much less experienced in how the data actually get used for public good – and that, after all is the point – to use data for public good, not just collect data. In the end, it is how public **data resources are put to public good uses** that will decide how we as a society want our data used. Data linkage agencies and others who work in this space at a national level may not be the best arbiters of public use good. This can result in a risk averse, overly bureaucratic approach to facilitating the release of data to researchers, negatively impacting on the ability of researchers to conduct rigorous and meaningful analyses with limited data access. In the end, researchers can conclude it is just not worth the enormous investment of time and effort.