

DATA AVAILABILITY AND USE

Submission to the Productivity
Commission Inquiry

July 2016

ABOUT RESEARCH AUSTRALIA

Research Australia is an alliance of 160 members and supporters advocating for health and medical research in Australia. Research Australia's activities are funded by its members, donors and supporters from leading research organisations, academic institutions, philanthropy, community special interest groups, peak industry bodies, biotechnology and pharmaceutical companies, small businesses and corporate Australia. It reflects the views of its diverse membership and represents the interests of the broader community.

Research Australia's mission is to make health and medical research a higher priority for the nation. We have four goals that support this mission:

- A society that is well informed and values the benefits of health and medical research.
- Greater investment in health and medical research from all sources.
- Ensure Australia captures the benefits of health and medical research.
- Promote Australia's global position in health and medical research.

Nadia Levin
CEO & Managing Director
02 9295 8547
nadia.levin@researchaustralia.org

Greg Mullins
Head of Policy
03 9662 9420
greg.mullins@researchaustralia.org

This document and the ideas and concepts set out in this document are subject to copyright. No part of this document, ideas or concepts are to be reproduced or used either in identical or modified form, without the express written consent of Research Australia Limited ABN 28 095 324 379.

DATA AVAILABILITY AND USE

SUBMISSION TO THE PRODUCTIVITY COMMISSION INQUIRY

INTRODUCTION

Research Australia welcomes the opportunity to make a submission to the Productivity Commission's Inquiry into Data Availability and Use.

Understanding health and the causes of ill health requires the examination and analysis of a range of different datasets in addition to health data, including, for example, birth and death records and information about educational attainment, criminal convictions and imprisonment, and contact with welfare and child protection. Research Australia's focus is on health and medical research (HMR) and this is reflected in our submission, and while our focus is HMR, we have an interest beyond health datasets.

This submission addresses issues relating to access to data by researchers and health practitioners, and the linkage of datasets within the Commonwealth Government and across jurisdictions. It identifies barriers to the availability of resources to facilitate access to datasets for research purposes and a lack of capability nationally in the preparation, manipulation and analysis of health and related data. It also identifies specific legal barriers to the better use of data for research purposes.

Research Australia contends that many public sector datasets are an underutilised national resource, and that the failure to make these datasets available to researchers prevents the full economic value of these datasets from being realised. In the case of HMR, the economic benefits from the better use of datasets are threefold:

- more efficient and cost effective research (much of which is funded by the Commonwealth, state and territory governments)
- greater efficiencies in the delivery of health care (again, much of which is funded by the Commonwealth, state and territory governments)
- the development of new and profitable health products and services.

Health and medical research is conducted in the public, higher education, private not for profit and corporate sectors and these groups are all represented in Research Australia's membership. This submission draws on input from across this membership and our annual national public opinion polling.

RESPONSES TO THE QUESTIONS POSED IN THE ISSUES PAPER

Research Australia has responded where applicable to the questions posed throughout the Issues Paper in the order in which the questions are presented. A full list of references is provided at the end of the submission.

QUESTIONS ON HIGH VALUE PUBLIC SECTOR DATA

What public sector datasets should be considered high-value data to the: business sector; research sector; academics; or the broader community?

High value data for the research sector is held in datasets that contain information about individuals' health and other characteristics.

At the Commonwealth Government level this includes: Pharmaceutical Benefits Scheme (PBS), Medical Benefits Scheme (MBS), Social Security and Veterans Affairs records, and more specific registers such as the Australian Childhood Immunisation Register. With the move to an opt-out model the My Health Record will also become an increasingly valuable dataset.

Within the jurisdiction of the Australian states and territories relevant datasets includes hospital admission records, Maternal and Child Health records, registers of births & deaths, primary care records where delivered by state agencies, educational attainment and school data, child protection, criminal convictions, and prison population data.

The private sector (both not for profit and for profit) also play a significant role in the delivery of healthcare in Australia. Relevant datasets from this sector include primary care data, hospital admission records, and data on childcare and aged care.

What characteristics define high-value datasets?

High value datasets are generally characterised by large volumes of reliable data on a large or complete population collected over an extended period. Sufficient and appropriate information to support reliable linkage with other datasets either through deterministic or probabilistic methods is also of value.

High value datasets should be collected and stored in a manner which enables the use and linkage of de-identified data for research purposes. Research should be recognised in the enabling legislation and other governing documents as a legitimate purpose for which information can be collected, used and disclosed; and access to data for appropriate research purposes should be facilitated by the owners/operators of data sets.

What benefits would the community derive from increasing the availability and use of public sector data?

Analysing existing data can help identify anomalies in treatments and processes which can be used to develop potential solutions. It can also assist with the validation of clinician's perspectives on the effectiveness or otherwise of particular treatments and the relative benefits of competing treatments.

Data is essential to the conduct of health and medical research. Data can be expensive and time consuming to collect, and so access to existing health data helps facilitate health and medical research. It can make research faster and more efficient, and delivering research outcomes earlier, leading to better health for the community.

In August 2013, the Menzies Foundation undertook a workshop on the topic of 'Public Support for Data-based Research to Improve Health'. A discussion paper based on the proceedings of the workshop includes a table listing 17 specific examples of research undertaken using data linkage which had been of public benefit. (Menzies Foundation, 2013 p. 25-26) The papers presented to this workshop and the report referred to above are available at the Menzies Foundation website.¹ The papers provide a useful exploration of many of the issues being addressed by the Productivity Commission's Inquiry.

Data helps us to understand disease, and whether treatments are working.

The number of newly diagnosed cancer cases in Australia increased from 66,393 in 1991 to 114,137 by 2009. Predictive modelling for the period 2007 to 2036 suggests that the impact of cancer is expected to increase dramatically, with approximately 110% in cumulative incidence of cancer in New South Wales alone. While such long-term predictions are tentative, it would be imprudent not to use routinely collected cancer data in large epidemiological and data linkage studies to identify priorities, to better plan and evaluate treatment strategies and screening programs, as well as care outcomes. Such large-scale research can also guide the development of more efficient and effective systems, and more effective evidence-based policies in a context where resources are limited. (Xathis, 2016)

For a further exploration of these benefits, and the relationship between public health research and public sector information, refer to Adams and Allen, 2014.

Research also has the potential to improve the productivity of our health sector. Analysis of hospital and MBS data can, for example, identify variations in care and inefficient and ineffective care practices. With total health expenditure in Australia of \$154.6 billion in 2013-14, even relatively small improvements in the quality, safety and efficiency of healthcare can lead to significant savings. With governments providing around two thirds of this expenditure, and with this representing around one quarter of taxation revenue, the potential savings to the Commonwealth, state and territory governments are significant, as the Productivity Commission itself has recognised Productivity Commission, 2015 p.18.² Research Australia supports this view.

¹ <http://menziesfoundation.org.au/partnerships/workshops/187-data-based-research-to-improve-health>

² Australian Institute of Health and Welfare 2015. Health expenditure Australia 2013–14. Health and

The National Strategic Framework for Chronic Conditions (the Framework) is being developed through the Australian Health Ministers' Advisory Council's Community Care and Population Health Principal Committee. The reason for the Framework is clear:

Chronic conditions are the leading cause of illness, disability and death in Australia. Tackling chronic conditions and their causes is the biggest challenge facing Australia's health system. Along with our ageing population, increasing consumer expectations and the high cost of pharmaceuticals and treatments, ever-increasing rates of chronic conditions are putting unprecedented strains upon individuals, communities and the health system.
AHMAC 2016 p.6

The Framework has a strong focus on prevention and identifies research as one of six key enablers. It emphasises the importance of sharing data, and provides a clear case for the benefits of doing so.

Strategic Priority Area 2.4: Information sharing

Quality information is a prerequisite for an efficient and effective health system that provides safe, evidence-based care. Data and research need to be connected at the system and population levels to facilitate better exchange of health information, improve collection and sharing of reliable and accurate data, and build the evidence base. Information collection and sharing is critical to building the evidence base to inform prevention, provide effective and appropriate care, and identify and better target priority populations.

Readily available, quality data and relevant research findings are also needed to inform continuous quality improvement processes, avoid duplication of effort and fragmentation in health information and data, and improve the safety and quality of care delivered.

Further, improved data capture, availability and coordination is needed to enable a clearer understanding of the health, social and economic impacts of chronic conditions, and, in turn, better support efforts to identify the most effective and efficient means of preventing and managing chronic conditions.

Quality information, as well as knowledge exchange and translation, is needed to grow the evidence base, support new and innovative solutions, and identify successful, clinically safe strategies for the prevention and management of chronic conditions at the individual, local, state and national levels. This requires more transparent health information and data, and quality information sharing and data linkage, to generate strong evidence and understand the bigger picture.

The Framework proposes the level of improvement in data linkage between existing health datasets as a means of measuring the success of the implementation of the Framework. AHMAC, 2016 p.33

International benchmarking

Australia is a member of the Organisation for Economic Cooperation and Development (OECD). A key advantage of membership is the opportunity to compare and benchmark Australia's national performance in a range of key areas against other OECD nations.

One of the areas in which the OECD collects data is Health Care Quality Indicators (HCQIs), enabling the comparison of a range of indicators of the quality of healthcare across OECD nations. The OECD's HCQI Project has been running since 2002, and the Australian Institute for Health and Welfare is the provider of data. For the most recent collection in 2015, Australia was able to provide data for only 26 of the 52 indicators.

Australia was not able to provide a number of HCQIs for the 2015 data collection for a range of reasons, largely relating to availability of data. Some of these reasons are given here by way of example:

- Data are not collected for some patient experience indicators.
- Indicators that aimed to measure mortality after discharge from hospital were not available, as **Australia does not routinely link hospital and mortality data.**
- **Data on suicides among admitted patients were not available as there is no agreed method for collecting and reporting these data in Australia.**
- **Indicators for readmissions to any hospital could not be prepared as individual patient records are not routinely linked in the national hospital data collections.**
- Some indicators required the use of patient-based data; however, the AIHW was unable to meet this requirement and so, instead, reported indicators based on separation-based data.

AIHW, 2016a

The inability to effectively benchmark Australian healthcare against the international community deprives us of an important means of understanding the relative performance and value for money of the Australian health system and the opportunity to identify potential target areas for improvement.

QUESTIONS ON COLLECTION AND RELEASE OF PUBLIC SECTOR DATA

What are the main factors currently stopping government agencies from making their data available?

The report of the Public Sector Data Management Project has effectively outlined the factors currently stopping the Commonwealth from making better use of data, and Research Australia submits that these are generally true of all Australian governments. (DMPC 2015)

Research Australia provides the following brief comments.

Culture and leadership

An emphasis by data custodians on privacy and security leads to a tendency to resist requests for access to data. The legitimate concerns of data custodians about data security and the protection of privacy need to be balanced by a recognition of the value of data and the importance of facilitating research because of the benefits it can provide to the community and governments in a contemporary society.

The provision of data for research purposes tends to be seen by data custodians as an additional impost rather than part of 'core business'. There is also evidence of a lack of certainty about who has

the authority to authorise the release of deidentified datasets for research purposes, or what steps are sufficient to deidentify data. (Much of the enabling legislation was written without recognising the value of data for research, or acknowledging research as a valid secondary purpose for collection of the data.) The 'mindset' needs to change to one of how to facilitate access to datasets while still mainlining security and individuals' privacy, and a risk management approach to facilitating access.

Resourcing

Making data available requires data custodians to prepare and provide datasets. It appears that many government departments and agencies are not adequately resourced to undertake this task, or at least to do so within a reasonable timeframe. Datasets are frequently held in databases that are old technology which can be difficult to interrogate and hinder the extraction of tailored datasets.

Capability

A shortage of individuals with the necessary skills to prepare and manipulate large datasets is an issue within the public sector and the Australian workforce more generally.

Should the collection, sharing and release of public sector data be standardised? What would be the benefits and costs of standardising?

Yes. The benefits of doing so are making the data more accessible and improving its reliability and accuracy. It would make research not only more effective but more efficient, reducing cost and saving time.

Collection of data

Data should be collected electronically wherever possible e.g. medical records and notes relating to hospital admissions. The greatest benefit of doing this is that it makes the data more available for the individual treatment of the patient (the record can easily be called up at a future time), but it also makes it more readily and cheaply available for research purposes.

Electronic collection of data also brings the benefit of standardised fields for the collection of data, and the potential for standardised terminology (choosing keywords from lists). The creation and use of a standard dictionary of terms for the description of diagnoses, conditions and treatments and the entry of this data in standardised fields would improve the reliability, usability and value of data.

For example, alcohol related harm is a key national issue. An attempt several years ago to use Queensland and NSW hospital admission records to identify alcohol related harm was hampered by the fact that unlike other Australian states, Queensland hospitals did not record whether a person was inebriated/intoxicated at admission.

Sharing of data

There are a number of existing Australian Government agencies and officially sanctioned organisations that can support the standardised sharing of data. These include the National Statistical Service, the Australian National Data Service, the Australian Institute of Health and Welfare, the Australian Bureau of Statistics, the Population Health Research Network and the Digital Transformation Office.

What would standards that are ‘fit for purpose’ look like?

As much as possible we should look to adopt existing national and international standards and protocols. Significant work has already been undertaken in this regard, and provides a foundation for further advances.

In relation to health data, the Australian Institute of Health and Welfare (AIHW) is responsible for the METeOR (Metadata Online Registry), Australia's repository for national metadata standards for the health, community services and housing assistance sectors. It includes National Minimum Datasets and Dataset Specifications and Data Dictionaries.³ The Australian Collaborating Centre (ACC) at the AIHW supports the World Health Organization Family of International Classifications (WHO-FIC) relating to data collection.

The United States Food and Drug Administration (FDA) specifies standards which must be complied with for the provision of clinical trial and other research data. In doing so it draws on the work of organisations like CDISC.

‘CDISC is a global, open, multidisciplinary, non-profit organization that has established standards to support the acquisition, exchange, submission and archive of clinical research data and metadata. The CDISC mission is to develop and support global, platform-independent data standards that enable information system interoperability to improve medical research and related areas of healthcare. CDISC standards are vendor-neutral, platform-independent and freely available via the CDISC website.’⁴

The development of data standards for Australia need to take account of initiatives to develop global standards. These provide benefits in terms of the familiarity of software providers and data users with the standards and the scope for ready exchange of data with overseas researchers and improved interoperability.

What criteria and decision-making tools do government agencies use to decide which public sector data to make publicly available and how much processing to undertake before it is released?

Research Australia provides the following comments from the perspective of an organisation outside government whose members regularly engage with governments in seeking to utilise their data.

Role of data custodians

The decision-making power to authorise the disclosure and use of government health data for research stems from the statutes under which the specific information is collected. Most often this power is entrusted to the executive head of the particular agency or the relevant Minister.⁶¹ The decision-maker must take into account the restrictions and procedures required by the particular legislative provisions.

³ <http://meteor.aihw.gov.au/content/index.phtml/itemId/181162>

⁴ <http://www.cdisc.org/about/mission>

The interaction between the statutes empowering the collection of data, the applicable privacy legislation and the common law and equitable duties of confidentiality is complex and varies from one dataset to another. There is considerable variation in the statutes governing the collection of the data with some containing clear authority and procedures to release identified data for research, while others fail to clearly articulate the purposes and permitted uses of the data. ...

Data custodians, unlike HRECs (Human Research Ethics Committees), are not required to provide reasons for their decisions and so there is a lack of transparency with the decision-making process. Although researchers might seek informal internal review of a data custodian's decision by approaching, for example, the head of the agency for a second opinion, there is no obvious place to seek independent external review of a data custodian's decision.

Adams C and Allen J 2014, p.966

The default position should be that (deidentified) data is released.

The most important element of a pro-disclosure regime is the presumption of openness. Under FOI principles, access to information is defined as a right and the default position is that information must be released unless there is a strong public interest in protecting the information. In contrast, the default position for access to government health-related datasets for research is that the datasets should not be released in order to protect privacy. One can argue that privacy can be properly protected by imposing conditions for release and that the default position should, therefore, be reversed.

Adams C and Allen J 2014, p.968

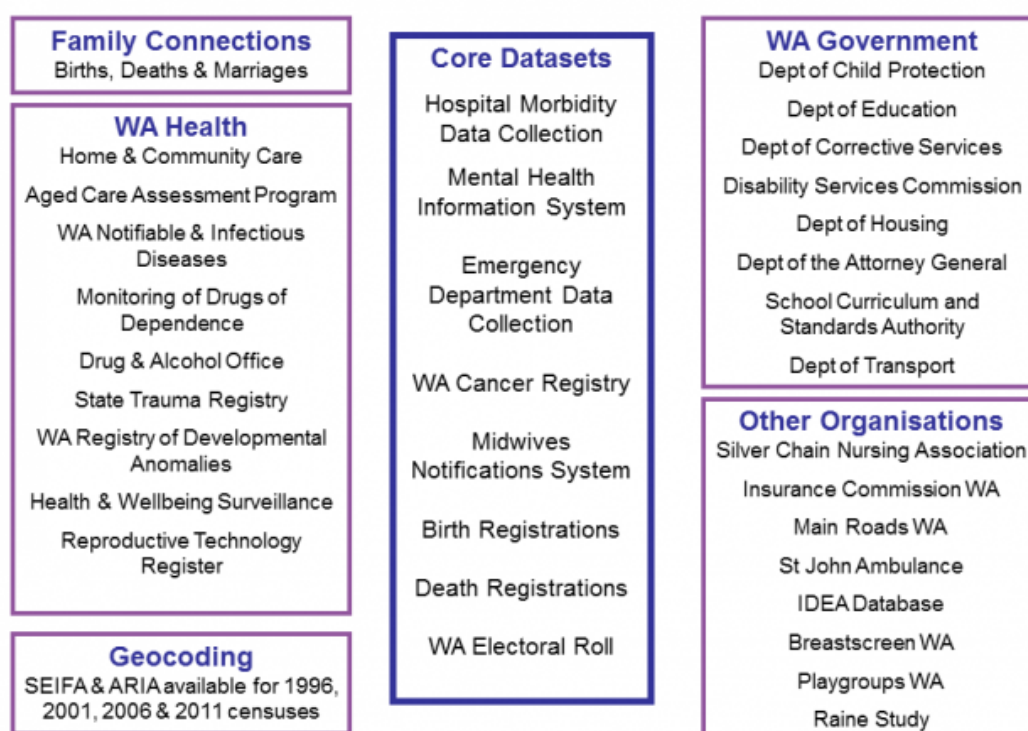
Beyond this default position, the public interest/public good should be a primary consideration. Proxies for an assessment of public good by the data custodian could be if/whether the research is being wholly or partially funded by a competitive funding program of a Commonwealth, State or Territory Government, or is being undertaken by a Government agency (e.g. CSIRO).

The legitimacy of commercial interests also needs to be recognised. Australia's healthcare system is complex and the private sector plays a significant part. For example, access to PBS, MBS and hospital data can assist pharmaceutical companies to monitor whether their medications are working as expected and the incidence of adverse events. It can also assist in identifying gaps in practice and the development of new products and services with the potential to generate commercial revenue. Commonwealth, state and territory governments offer a range of programs to support the commercialisation of research, recognising that it is in Australia's to do so; and there are several programs specifically supporting medical devices and pharmaceuticals. Providing access to data is another means of supporting commercial development in this area.

What specific government initiatives (whether Australian Government, state, territory or local government, or overseas jurisdictions) have been particularly effective in improving data access and use?

Western Australia

Western Australia is generally recognised as the Australian leader in the provision of public data and its utilisation. The Western Australian Data Linkage Branch links many data collections from the Department of Health WA (DOHWA) and other agencies. It maintains a set of core and infrastructure linkages that are updated on a regular basis, as illustrated below.



Source: <http://www.datalinkage-wa.org.au/data-collections>

The Western Australian *Whole of Government Open Data Policy* provides further information about the rationale and approach taken by the Western Australian Government. (Government of WA, 2015) The Chief Scientist of Western Australia (WA), Professor Peter Klinken, is currently chairing an expert advisory group that is examining the State's strengths and capabilities in data linkage, the barriers to data linkage, and the processes, governance and funding arrangements. It is due to report to the Premier by 31 August 2016. Research Australia suggests that the report of the Review could usefully inform the Productivity Commission's own current Inquiry.⁵

⁵ More information is available at <https://www.dpc.wa.gov.au/Consultation/Pages/Data-Linkage-Review.aspx>

Queensland

The Queensland Government has also taken steps to link several datasets.

‘There is also a Master Linkage File (MLF) containing permanently linked references to QHAPDC, QPDC, RG births, RG deaths, and EDIS. The use of this file saves a significant amount of RLG resources and results in a faster processing time, benefiting researchers. The MLF is now updated in near-real time such that data are extracted from all sources twice each month and linked with all other records included in the MLF. The currency of the data included in the MLF is limited by the currency of data that are available in the source data collections. See Appendix 1 for details of approximate lag times for submission of data for each data collection. Work is currently being undertaken to expand the MLF with data from further sources including the Notifiable Conditions System (NOCS), the Vaccination Information and Vaccination Administration System (VIVAS) and the Queensland Ambulance Service (QAS) to improve its coverage and therefore usefulness.

Current Master Linkage File coverage:

- QHAPDC Hospital Admissions
- Queensland Perinatal Data Collection
- Registrar General births
- Registrar General deaths
- Emergency Department Information System
- Elective Surgery Waiting List

Queensland Government, 2016

Denmark

The Danish National Patient Registry (DNPR) is often cited as an overseas example.

The DNPR provides nationwide longitudinal registration of detailed administrative and clinical data. Furthermore, the DNPR data are linkable at the patient level with data from other Danish administrative registries, clinical registries, randomized controlled trials, population surveys, and epidemiologic field studies.

For an overview of the register and the other components of the Danish system, refer to Schmidt, 2015.

QUESTIONS ON DATA LINKAGE

Which datasets, if linked or coordinated across public sector agencies, would be of high value to the community, and how would they be used?

Datasets that contain information about individuals' health and other characteristics and which are able to be linked to other data relating to the same individuals.

At the Commonwealth Government level this includes: Pharmaceutical Benefits Scheme (PBS), Medical Benefits Scheme (MBS), Social Security and Veterans Affairs records, and more specific registers such as the Australian Childhood Immunisation register. With the move to an opt-out model the My Health Record will also become an increasingly valuable dataset.

Within the jurisdiction of the Australian states and territories relevant datasets includes hospital admission records, Maternal and Child Health records, registers of births & deaths, primary care records where delivered by state agencies, educational attainment and school data, child protection, criminal convictions, and prison population data.

The private sector (both not for profit and for profit) also play a significant role in the delivery of healthcare in Australia. Relevant datasets from this sector include primary care data, hospital admission records, and data on childcare and aged care.

Case Study: Prospecion

Prospecion Pty Ltd is a health analytics company that works across the health sector with industry, research and government organisations using government, client and public data. The focus has been PBS data and medication utilisation, and it has recently been given access to a sample of linked PBS and MBS data. However, Prospecion reports that there are still significant gaps in research and industry needs. For example, linking hospital separations data and mortality data with PBS and MBS data would provide a more complete picture of the value of medications. This ‘real world evidence’ has benefits for payers, industry, researchers and consumers. Industry is able to better understand patient utilisation across the system and provide more accurate estimates for market estimates and market access. In post market surveillance, a more complete picture with a broader dataset can be utilized for adverse events reporting. Researchers would be better able to identify variation in use, potential healthcare efficiencies and analyse patient outcomes.

Which rules, regulations or policies create unnecessary or excessive barriers to linking datasets?

The legislative and regulatory barriers to linking datasets fall into two broad categories: legislation at the Commonwealth, State and Territory level dealing with privacy generally; and provisions in legislation dealing with specific databases and datasets.

Personally Controlled Electronic Health Records Legislation

There are situations in which current legislation prevents specific agencies and organisations from providing data in a manner which would facilitate data linking, even within Government. The My Health Record, regulated by the *Personally Controlled Electronic Health Records Act 2012* (the PCEHR Act) is one of these. The My Health system has the potential to be an important resource for health and medical research, and for this research to be of enormous benefit to the community.

Use of the data for research purposes is consistent with the objects of the PCEHR Act. In particular, research undertaken using the health information in the PCEHR system can help address the PCEHR Act’s objectives, which are to:

‘reduce the occurrence of adverse medical events and the duplication of treatment; and improve the coordination and quality of healthcare provided to consumers by different healthcare providers.’⁶

⁶ Section 3 of the PCEHR Act

The PCEHR Act also specifies that the provision of data for research or public health purposes is a function of the PCEHR System Operator (section 15), and makes provision for Rules to be made in relation to this function (section 109). However, these provisions relate solely to 'de-identified data'.

Research Australia acknowledges that researchers rarely need to know the identity of individuals, and that these exceptions typically relate to research activities such as clinical trials where the individual's consent to the provision of their health information can be obtained as part of their consent to participation in the trial. The real difficulty with the requirement that only de-identified data can be provided by the PCEHR operator, is the limitation this imposes on the ability to link data from the PCEHR system with other data.

Operating within the Commonwealth Government, The National Statistical Service provides a Statistical Data Integration Framework for combining information from different administrative and/or survey sources to provide new datasets for statistical and research purposes.⁷

There are Approved Integrating Authorities (Australian Bureau of Statistics, Australian Institute of Health and Welfare, Australian Institute of Family Studies) which are able to receive identified datasets for the purposes of data linkage. This service enables linked deidentified datasets to be created and provided to researchers. However, the Integrating Authorities can only do so if the legislation governing the data allows the data custodian to provide identified data. As shown above, the legislation governing the My Health record prevents the data custodian (the operator) from providing identified data to an approved Integrating Authority.

Even though it is relatively recent legislation, enacted in 2012, the PCEHR Act fails to take account of the importance of data linkage or the existence of the Statistical Data Integration Framework as a mechanism for providing researchers with access to linked deidentified data.

RECOMMEDATION

The PCEHR legislation needs to be amended and/or some other legislative mechanism used to enable the PCEHR System Operator to provide identified data to appropriately accredited agencies for data linking purposes. This may require more formal legal recognition of the Statistical Data Integration Framework.

⁷<http://www.nss.gov.au/nss/home.NSF/pages/Data+Integration+Landing%20Page?OpenDocument>

Case study: Population Health Research Network

The Population Health Research Network (PHRN) has been established by the Australian Government to build a nationwide data linkage infrastructure capable of securely and safely managing health information from around Australia. The PHRN is developing and testing leading-edge technology to ensure the safe and secure linking of data collections whilst working to protect individuals' identity and privacy. It is also developing mechanisms for the secure exchange of linkable data between data custodians and the researchers who receive approval to analyse the linked data. A key aspect of these arrangements is that the linked datasets provided to researchers do not contain the identifying information of the individuals whose data is included in the dataset, thus protecting their privacy.

Addressing the limitations that currently prevent organisations like the PHRN from reaching their full potential is critical.

The Productivity Commission has itself put the case for reform in this area:

There are ways to protect privacy and confidentiality while still allowing data to be used by researchers. Strategies include using anonymous identifier numbers to link records across datasets without identifying individuals, suppressing variables where there are too few observations, and making data less precise by changing the level of detail (PC 2013). Where there are risks of misuse, access to data can be restricted to authorised researchers for pre-approved purposes. Such approaches for protecting privacy are already reflected in the Australian Government's principles for linking government datasets (Australian Government 2010). State-based efforts in using and linking administrative datasets for research have shown that significant benefits can be realised with low risks, manageable costs and the protection of people's privacy (PC 2013).

Political will is often needed to address privacy concerns in a way that allows data to be released. Policy makers need to make tradeoffs between a high level of confidentiality and the consequences of not making data available. Concealing data can mean that patients receive ineffective (or even harmful) care, adverse effects of drugs go undetected, or significant money is spent on interventions that do not improve health outcomes (rather than on interventions that do). It can also make it difficult to hold health care providers to account for their performance.

Productivity Commission, 2015 p.85-86

Research Australia has conducted polling of the Australian public in relation to health and medical research and health more broadly since 2003. In recent years we have investigated the public's attitudes to personal health information being used for research. Research Australia polling conducted in 2013 found that 79% of the Australian population supported the use of patients' medical records for research purposes. Research Australia, 2013. In 2014, we asked questions about tissue samples, which found strong support for samples that were collected for medical procedures being stored and used for subsequent research. Research Australia, 2014. Polling in 2015 again found very strong support for personal health information being used for research purposes. Individuals were motivated by a desire to advance medical research, improve patient care and assist public health officials to track diseases and disability and understand the causes. Research Australia, 2015

This polling provides evidence of strong levels of support for the use of personal health information for research purposes.

Privacy Legislation

Commonwealth, State and Territory governments all have legislation governing how personal information collected by public and private sector organisations can be collected, stored and used. Many also have special provisions relating to health information. The jurisdictions overlap (for example, a private hospital in Victoria can be subject to requirements under the Victorian and Commonwealth legislation). The concepts and approaches are broadly similar and the same terminology tends to have similar but subtly different meanings.

Researchers trying to navigate these different requirements to negotiate access to data find the existing framework bewildering and frustrating, as do many of the agencies and ethics committees they work with. Research projects are typically funded for a fixed period (for example 3 years) and it can take months and years to negotiate access to data. This can prevent the research being undertaken at all in some cases, and significantly adds to the cost of conducting research in all cases. Both the cost and the length of time required to achieve access to datasets may result in the research team's inability to achieve key aspects of the research, therefore effectively wasting valuable public funding.

This complexity is illustrated the table provided in the Appendix, and some further examples of specific legislative provisions which follow the table.

Use of Government Identifiers

Privacy legislation also imposes specific restrictions on the use of some Government identifiers which appear to prevent them being used as a data item for linkage of datasets, even where they are removed from the final dataset provided to researchers. The Medicare number is a case in point.

Australian Privacy Principle 9 restricts the adoption, use and disclosure of government related identifiers by organisations. A government related identifier is an identifier that has been assigned by an agency, a State or Territory authority, an agent of an agency or authority, or a contracted service provider for a Commonwealth or State contract. A Medicare number is such an identifier. An organisation must not use or disclose a government related identifier of an individual, unless an exception applies. The restriction on using a government related identifier prevents an organisation (including a Commonwealth Government Department or agency) from using a Medicare number to link records in two or more datasets. Research is not one of the exceptions provided under the Australian Privacy Principles.

Allowing Medicare numbers to be used by Approved Integrating Authorities to link datasets would be an important step forward. Other organisations that hold Medicare numbers should also be allowed to use the Medicare number to link separate datasets within their own field of operations. (This does not involve the provision of Medicare numbers to researchers.)

Identification and consent

Issues of consent and the need to use identifiable data are often raised as barriers to linking of data, and raise issues about the types of consent that are required, how data can be linked and who can

perform this task. Public expectations about privacy are often provided as a reason for not linking datasets.

Despite this, we know that the public supports the use of health information for research purposes. Attitudes to the use of personal information for data linkage are more complex, and raise concerns about privacy and ‘who knows what’. However, explaining the way in which data linkage is conducted can assuage these fears and generally makes people more supportive, reducing the number of people who withhold consent and/or opt out.

With some exposure to the features of data linkage, lay people have the capacity to understand the processes sufficiently in order to consider ethical issues associated with consent preferences. Shifts in views reveal the complexity of such decisions. While privacy protection remained an important consideration for most participants, adequate protection measures adopted in best practice data linkage were viewed by most as protection enough for data linkage to proceed without specific individual consent.
Xafis, 2015

The role of consent also needs to be examined. In health and medical research, consent models have arisen in respect of participation in experimental procedures, as part of informed consent to treatment. This approach has tended to carry over to research activities using individual’s personal information, notwithstanding that the circumstances and risks associated with the two activities are very different. In particular, the opportunity to obtain consent presented by, for example, a clinical trial with individual enrolment and an epidemiological study using existing data are very different. Despite specific provisions for research in the Australian Privacy Principles and in the National Statement on Ethical Conduct in Human Research, this difference, and the role of secure, third part linking agencies is not always recognised by HRECs and data custodians.

The important role that explicit informed consent plays in the conduct of research cannot be denied. Inhumane medical research has prompted over many decades the articulation of guidelines, legislation, and codes to ensure that research participants are protected from the harms inherent in some forms of research. However, there are now certain kinds of research, such as large epidemiological studies or data linkage studies, which offer potentially great benefits for whole populations but which, at the same time pose minimal, if any, harms to those included. These forms of research should not be required to adhere to the traditional informed consent requirements for the reasons articulated in this paper. The paper focuses on consent options for cancer population studies and examines the ethical issues associated with each model.
Xafis, 2016

RECOMMENDATION

Australia needs a national, uniform and consistent framework for the provision of linked datasets to researchers. This framework needs to recognise the importance of data linkage and endorse the role of agencies that are authorised by Governments to receive datasets with identified data for the purpose of creating linked datasets for research purposes. These linking agencies need to be jointly authorised by State, Territory and Commonwealth Governments to receive and link datasets across state boundaries to create national linked datasets using data collected by Commonwealth, State and Territory Governments. Legislative reform is required to remove existing barriers that prevent data custodians from providing datasets for linking and to provide further exceptions to the use of government identifiers by linking agencies for linkage purposes.

The Council of Australian Governments has played an active part in achieving national, consistent approaches in specific health related areas in the past. The National Core Maternity Indicators Reports, which originated from an Inquiry in Western Australia, is one example of what is possible. AIHW, 2016b Research Australia proposes that COAG play a lead role in the development of this framework.

How can Australia's government agencies improve their sharing and linking of public sector data? What lessons or examples from overseas should be considered?

This question was considered by the Public Sector Data Management Project auspiced by the Department of Prime Minister and Cabinet in 2015. The Project's report identified a 'Roadmap to deliver transformation'. Research Australia endorses the actions identified in this Roadmap (and acknowledges that this current Inquiry is one of the outcomes of the Roadmap).

In addition, we propose that there be explicit recognition of the role that researchers in academia and publicly funded research organisations can play in using data to help improve the delivery of Government services. In particular, Research Australia emphasises the public benefit that can accrue from providing health and medical researchers with greater and more effective access to linked datasets.

The Public Sector Data Management Project has a focus on Commonwealth Government datasets. Research Australia proposes that the real value for the health of Australians can only be unlocked where data linkage occurs across Commonwealth, State and Territory jurisdictions and includes data collected by government and private sector service providers (the latter including not for profit and for profit operators and service providers). The submission by the Population Health Research Network to the Productivity Commission's current Inquiry provides a clear description of how the *separation principle* can be used to link data across datasets, including across jurisdictions while addressing privacy and security concerns.

RECOMMENDATION

Research Australia recommends the development of a whole of governments approach to improving the sharing and linking of public sector data along the lines of the Public Sector Data Management Project's 'Roadmap to deliver transformation'. This should incorporate the national, uniform and consistent framework for the provision of linked datasets to researchers recommended above.

Such a Framework would not only increase the volume and utility of research conducted but improve the efficiency of the research process by streamlining existing processes which can impose significant delays and costs to research projects as well as the government departments and agencies dealing with the requests. The VALiD project outlined below is a stark example of why such reforms are needed and the benefits of doing so.

Casestudy: The VALiD project- six years to secure approvals to link data

The Vaccine Assessment using Linked Data Safety Study (VALiD) is a research project funded by the Australian Research Council's Linkage Program with additional financial and/or in-kind support provided by state governments and state based agencies. Its central objective is to identify and assess adverse events following immunisation by linking the Australian Childhood Immunisation Register (ACIR) with other datasets, to explore if a relationship exists between vaccination and admission to hospital or death. In addition, the study examined the ethical and legal acceptability of data linkage for vaccine safety surveillance. It also examined the effectiveness of methods to obtain consent for data linkage and convened a Citizens' Jury to weigh the evidence and make recommendations about the use of data linkage for public health surveillance.

The public benefit of VALiD is clear- Australia is keen to increase childhood immunisation rates, provide evidence on the safety of vaccines and to counter spurious claims that immunisation is responsible for conditions such as autism. VALiD can provide the data to do this, as well as detecting problems or issues that arise, for example from a defective vaccine batch.

In addition to the ACIR, the original plan was to link with the National Death Index (NDI) and hospital admission and emergency department data from South Australia, NSW, Victoria, Queensland and Western Australia. The project commenced discussions with the Commonwealth Government in August 2008.

At the Commonwealth level the project was required to liaise with six separate agencies or sections to attain the necessary individual approvals for access to the ACIR and the NDI. These approvals were not provided until 2012, four years after the project commenced. A further two-and-an-half years passed before the integrated data were available for the researchers to commence their analyses.

The State Governments would not, on the whole, consider the application to link data until the Commonwealth approvals had been secured. The project has now secured a commitment from the state governments with the exception of two datasets in each of Western Australia and Victoria. While access has now been obtained to the other state datasets, none of the data has yet been linked. This task will be undertaken by the AIHW as the approved integrating authority, and it is anticipated that this will occur in October 2016, some eight years after the project commenced.

The project originally only had funding for three years. It has been able to continue until now by seeking approval to carry forward unexpended funding, project staff transitioning into student roles as well as submitting further funding applications to continue the work.

The Canadian Institute for Health Information is an international example of effective data collection and linkage within a federal system.

Casestudy: Canadian Institute for Health Information

The Canadian Institute for Health Information (CIHI) was established in 1994 by agreement between the federal, provincial and territorial governments, to improve the “deplorable state of health information” in Canada. Working in partnership with federal, provincial and territorial ministries of health, the CHI has become the leading source of high-quality, unbiased, comparable information about the delivery of health care in Canada, the performance of our health systems and the factors that affect Canadians’ health. It identifies its core strengths as:

- Identifying health information needs and priorities;
- Coordinating and promoting standards and data quality;
- Developing and managing health system databases and registries;
- Developing comparable measures of health system performance;
- Conducting analyses in the areas of population health and health services; and
- Building capacity and conducting education sessions.⁸

⁸ Canadian Institute of Health Information, *Strategic Plan 2016-2021* p.8 <https://www.cihi.ca/en/strategicplan2016-2021-enweb.pdf>

QUESTIONS ON HIGH VALUE PRIVATE SECTOR DATA

What private sector datasets should be considered high-value data to: public policy; researchers and academics; other private sector entities; or the broader community? In each case cited, what characteristics define such datasets?

Primary care patient records – private practice

Primary care is a critical component of Australia's health system. It has the greatest engagement with the Australian population and with other elements of the health system. There is enormous scope for primary care records to aid research, bit as a sole source of data and as a part of the broader picture of patient care in Australia when combined with other records (PBS, hospital admission, births and deaths etc.)

Medical practitioners are increasingly using clinical software to manage patient records in their practices and in some instances, this data is already being used for research purposes. See for example Carrington et al, 2015, which analysed recorded blood glucose levels in 467,955 deidentified patients. The participating GP practices use the same software and participate in the General Practice Research Network (GPRN). 'Practices of the GPRN are provided with a data collection software tool which encrypts and de-identifies all data prior to it been sent on a weekly basis to the HCN.'

Private hospitals

Similarly, private hospitals, (both for profit and non-profit) are essential components of Australia's health system and are repositories of significant amounts of patient data that can be useful to researchers. The extent to which these hospitals are engaged with the Australian health and medical research community varies; some are actively engaged with research and have their own research institutes and/or formal arrangements with universities and independent Medical Research Institutes. Others are involved on a more sporadic basis or not at all.

Private health insurers

Private hospitals are also repositories for large amounts of information about treatments provided within Australia, and complements data held by the MBS. Together this data provides a more complete picture of who is receiving treatment, where, by whom, and for which conditions.

Reciprocity

Reciprocity is an important consideration. The private sector participants in the health system frequently seek access to public data; the provision of private sector data for research purposes provides a more complete understanding of health and the healthcare system, and patients are ultimately the beneficiaries.

QUESTIONS ON PRIVACY PROTECTION

How could coordination across the different jurisdictions in regard to privacy protection and legislation be improved?

Research Australia recommends the development of a whole of governments approach to improving the sharing and linking of public sector data along the lines of the Public Sector Data Management Project's 'Roadmap to deliver transformation'.

This should incorporate a national, uniform and consistent framework for the provision of linked datasets to researchers. This framework would formally recognise the importance of data linkage and endorse the role of agencies that are authorised by governments to receive datasets with identified data for the purpose of creating linked datasets for research purposes. Linking agencies need to be jointly authorised by State, Territory and Commonwealth governments to receive and link datasets across state boundaries to create national linked datasets using data collected by Commonwealth, State and Territory Governments.

Legislative reform is required to remove existing barriers that prevent data custodians from providing datasets for linking and to provided further exceptions to the use of government identifiers by linking agencies for linkage purposes.

How effective are existing approaches to confidentialisation and data security in facilitating data sharing while protecting privacy?

There are very effective existing approaches. Organisations such as the approved integrating authorities, the Sax Institute, the Population Health Research Network and the state based data linking agencies have developed secure facilities for the manipulation, analysis and storage of datasets.

Techniques for the protection of privacy while facilitating the utility of data continue to evolve. The application of such measures to patient records held by GPs gives a good indication of the measures that can be employed to protect identity while still providing detailed information about where patients are residing. (Mazumdar et al 2014) This information can be important understanding the sociodemographic context and to the planning of services.

CONCLUSION

Research Australia has appreciated the opportunity to make this submission. It has drawn on input from across our membership and Research Australia acknowledges the contributions made by many individuals and organisations.

As the quantities of data collected about all of us grows, so does our capacity to utilise this data for the benefit of all. There is evidence that community attitudes to data and questions of privacy and security are changing, and legislation and government practices need to change to both reflect this change and enable us to harness this data as a valuable resource. Now is the time to act if we are to make the most of the opportunities public data provides to improve the delivery of government services and the health of our community, and to create the new economic opportunities and jobs of tomorrow.

Appendix: Comparison of Commonwealth, Victorian, NSW and ACT Privacy Frameworks

The following table provides some idea of the complexity faced by researchers seeking to utilise datasets in more than one jurisdiction

Australian Privacy Principles, Health Privacy Principles and Privacy Principles

Commonwealth	Victoria	New South Wales	Australian Capital Territory
13 APPs Privacy Amendment (Enhancing Privacy Protection) Act 2012, which amends the Privacy Act 1988 (Cth)	11 HPPs Health Records Act 2001 (Vic)	15 HPPs (NSW) Health Records and Information Privacy Act 2002	12 PP's Health Records (Privacy and Access) Act 1997 (ACT) -
<i>apply to organisations, and Australian Government (and Norfolk Island Government) agencies</i>	<i>apply to public and private sectors</i>	<i>apply to public sector agencies and private sector organisations</i>	<i>public and private sector health records</i>
1 Open and transparent management of personal information	1. Collection	Collection	1. Manner and purpose of collection of personal health information
2 Anonymity and pseudonymity	2. Use and disclosure	1. Lawful	2. Purpose of collection of personal health information to be made known
3 Collection of solicited personal information	3. Data quality	2. Relevant	3. Solicitation of personal health information generally
4 Dealing with unsolicited personal information	4. Data security and data retention	3. Direct	4.1. Storage, security and destruction of personal health information—safekeeping requirement
5 Notification of the collection of personal information	5. Openness	4. Open	4.2. Storage, security and destruction of personal health information—register of destroyed or transferred records
6 Use or disclosure of personal information	6. Access and correction	Storage	4.3. Storage, security and destruction of personal health information—destruction of health information

7 Direct marketing	7. Identifiers	5. Secure	5. Information relating to records kept by record keeper
8 Cross-border disclosure of personal information	8. Anonymity	Access and accuracy	6. Access to health records by people other than the consumer
9 Adoption, use or disclosure of government related identifiers	9. Transborder data flows	6. Transparent	7. Alteration of health records
10 Quality of personal information	10. Transfer or closure of the practice of a health service provider	7. Accessible	8. Record keeper to check accuracy etc. of personal health information before use etc.
11 Security of personal information	11. Making information available to another health service provider	8. Correct	9. Limits on use of personal health information
12 Access to personal information		9. Accurate	10. Limits on disclosure of personal health information
13 Correction of personal information		Use	11. Relocation and closure of health service practice
		10. Limited	
		Disclosure	12.1. Consumer moves to another health service provider
		11. Limited	12.2. Health service provider moves to another health service practice
		Identifiers and anonymity	
		12. Not identified	
		13. Anonymous	
		Transferrals and linkage	
		14. Controlled	
		15. Authorised	

This table has been adapted by Dr Vicki Xafis, updated from the table originally published in Xafis, 2011. It is not intended to be comprehensive but rather intends to highlight the variations between privacy principles relating to health in different jurisdictions.

Appendix: Comparison of Commonwealth, Victorian, NSW and ACT Privacy Frameworks

The following comments provide further examples of the complications posed for researchers by the privacy Frameworks. These are provided as examples only and not intended as a comprehensive or complete list of provisions which provide barriers to research.

New South Wales Health Privacy Principles

Health Privacy Principles 14 and 15 deal with 'Transferrals and Linkage'. HPP 15 deals specifically with authorisation for linkage: 'Only use health records linkage systems if the person has provided or expressed their consent.'

The NSW Handbook to Privacy (p.48) seeks to clarify the application of HPP 15:

Health privacy principle 15

See Schedule 1 of the *Health Records and Information Privacy Act 2002 (NSW)*

When does this health privacy principle apply?

HPP 15 does not apply to all electronic storage and linkage of health records. It does not affect the ability of organisations to store or link health information electronically.

HPP 15 is designed to deal with electronic health records systems that link health records at a state or national level, for example:

- The Health e-Link electronic health records scheme in NSW
- The Health *Connect* electronic health records scheme at a federal level

It requires that such schemes must be "opt-in", meaning that a patient must give express consent to participate.

The explanation regarding the NSW HPP 15 provided in the *Handbook to Privacy* remains silent on its application to data linkage in the research context. The lack of clarity may intend that (or lead to the understanding that) data linkage research/projects also require consent. If this were the case, it would not be possible to conduct large data linkage projects. The explanation fails to make any reference or concession to the identification of data that has been linked.

VICTORIA

At least two pieces of Victorian legislation are relevant to the linkage of data in Victoria.

The Privacy and Data Protection Act 2014 (Vic)

The *Privacy and Data Protection Act 2014 (PDPA)* replaced the *Information Privacy Act 2000* and the *Commissioner for Law Enforcement Data Security Act 2005*. The PDPA governs the collection and handling of personal information (excluding health information) in the Victorian public sector and, uniquely, provides for the establishment of a protective data security regime for the Victorian public sector.

The Health Records Act 2001 (Vic)

Although information pertaining to an individual's health is considered to be that individual's personal information, health information is expressly excluded from the PDPA.

The PDPA is an additional piece of privacy legislation that some researchers would also need to be familiar with if they were linking health related data to non-health related data. This is only one example of such legislation.

Interaction of legislation between jurisdictions

Despite legislative changes to privacy legislation in recent years, it remains complex, variable between States and Territories, and very challenging for researchers as they continue to need to develop legal expertise in order to conduct their research or seek those with legal training. For researchers not overly familiar with legislation, the explanation regarding definitions in the above excerpt would be confusing.

In addition, the various pieces of health privacy legislation describe their scope of application differently with definitions for key terms found in various parts of the legislation. For example, the *Privacy Amendment (Enhancing Privacy Protection) Act 2012*, which amends the Privacy Act 1988 (Cth) applies to organisations, and Australian Government (and Norfolk Island Government) agencies. In contrast, the *Health Records Act 2001 (Vic)* applies to public and private sectors. On the other hand, the *NSW Health Records and Information Privacy Act 2002* applies to public sector agencies and private sector organisations. Furthermore, key terms are defined differently in different pieces of legislation. Examples of different definitions for the same terms are:

In the *Privacy Amendment (Enhancing Privacy Protection) Act 2012 (Cth)*, organisation means:

- (a) an individual; or
- (b) a body corporate; or
- (c) a partnership; or
- (d) any other unincorporated association; or
- (e) a trust;

However, in the *NSW Health Records and Information Privacy Act 2002*, 'organisation' is defined as: *a public sector agency or a private sector person*.

Despite a significant review of the Commonwealth legislation and a number of amendments over several years, no simplification of privacy legislation has been achieved with revisions to the Commonwealth privacy legislation apart from removing one set of privacy principles. The articulation of the new APPs is convoluted and labour intensive which complicates their use for lay people. There is certainly still not a framework that applies consistently throughout Australia and this will be difficult to achieve with States and Territories maintaining separate pieces of legislation (accompanied by their own sets of privacy principles).

REFERENCES

Adams C and Allen J 2014, Government databases and public health research: Facilitating access in the public interest, *Journal of Law and Medicine*, June 2014, 957-972

AHMAC 2016, Australian Health Ministers' Advisory Council, 2016 National Strategic Framework for Chronic Conditions, Second Draft, Online Consultation May 2016

Australian Institute of Health and Welfare 2016a. OECD health-care quality indicators for Australia 2015. Cat. no. PHE 209. Canberra: AIHW.

Australian Institute of Health and Welfare 2016b. National core maternity indicators—stage 3 and 4: results from 2010–2013. Cat. no. PER 84. Canberra: AIHW.

Carrington MJ, Stewart S, Gerber T and Cohen N 2015 Bittersweet findings of blood glucose levels in 467,955 patients in primary care. Baker IDI Heart and Disease Institute, Melbourne Australia

Department of Prime Minister and Cabinet 2015, Public Sector Data Management, Commonwealth of Australia

Government of Western Australia 2015, The Western Australian Whole of Government Open Data Policy Version 1.1 April 2015

Queensland Government, 2016 Queensland Data Linkage Framework (Queensland Health), January 2016, p.8

Menzies Foundation, 2013 Public Support for Data-based Research To Improve Health: A discussion paper based on the proceedings of a Menzies Foundation Workshop 16th August, 2013 available at <http://menziesfoundation.org.au/partnerships/workshops/187-data-based-research-to-improve-health>

Productivity Commission 2015, Productivity Commission 2015, *Efficiency in Health*, Commission Research Paper, Canberra. P.

Research Australia, 2013 What do Australians think about Health and Medical Research?

Research Australia, 2014 Australia Speaks! 2014 Opinion Poll- views of 1000 Australians

Research Australia, 2015 Australia Speaks! Research Australia Opinion Polling 2015

Schmidt M, Schmidt SAJ, Sandegaard JL, Ehrenstein V, Pedersen L, Sørensen HT. The Danish National Patient Registry: a review of content, data quality, and research potential. *Clinical Epidemiology*. 2015; 7:449-490. doi:10.2147/CLEP.S91125.

Soumya Mazumdar, Paul Konings, Michael Hewett, Nasser Bagheri, Ian McRae, Peter Del Fante 2014 Protecting the privacy of individual general practice patient electronic records for geospatial epidemiology research, Aust NZ J Public Health. 2014; 38:548-52; doi: 10.1111/1753-6405.12262

Xafis V, Thomson C, Braunack-Mayer AJ, Duszynski KM and Gold S, 2011 Legal impediments to data linkage, Journal of Law and Medicine (2011) 19 300-315

Xafis, V. 2015 The acceptability of conducting data linkage research without obtaining consent: lay people's views and justifications BMC Medical Ethics (2015) 16:79 DOI 10.1186/s12910-015-0070-4

Xafis V. 2016, What Consent Model Is Ethically Justifiable In Cancer Population Research?, Cancer Forum Volume 40 Number 2 July 2016

RESEARCH AUSTRALIA LIMITED

384 Victoria Street Darlinghurst NSW 2010

T: +61 2 9295 8546 | ABN: 28 095 324 379

www.researchaustralia.org