

The Senate

Select Committee on
Health

Sixth interim report

Big health data: Australia's big potential

May 2016

© Commonwealth of Australia 2016
ISBN 978-1-76010-367-5

This work is licensed under the Creative Commons Attribution-NonCommercial-NoDerivs 3.0 Australia License.



The details of this licence are available on the Creative Commons website: <http://creativecommons.org/licenses/by-nc-nd/3.0/au/>.

Printed by the Senate Printing Unit, Parliament House, Canberra.

Membership of the Committee

Members

Senator Deborah O'Neill (ALP, NSW) **Chair**

Senator Ricky Muir (AMEP, VIC) (from 25 June 2015) **Deputy Chair**

Senator Sean Edwards (LP, SA)

Senator Jenny McAllister (ALP, NSW) (from 14 May 2015)

Senator Claire Moore (ALP, QLD) (from 26 November 2015)

Senator Zed Seselja (LP, ACT)

Senator John Williams (NATS, NSW)

Substitute members

Senator the Sam Dastyari (ALP, NSW) 11 December 2015

Secretariat

Mr Stephen Palethorpe, Secretary

Miss Jedidiah Reardon, Principal Research Officer

Mr Josh See, Senior Research Officer

Mr Michael Kirby, Senior Research Officer

Mr Josh Wrest, Senior Research Officer

Ms Rosalind McMahon, Administrative Officer

Ms Sophie Clark, Administrative Officer

PO Box 6100

Parliament House

Canberra ACT 2600

Phone: 02 6277 3419

Email: health.sen@aph.gov.au

Website: www.aph.gov.au/select_health

Table of Contents

Membership of the Committee	iii
Terms of Reference	vii
Acronyms and abbreviations	ix
Glossary.....	xi
Executive Summary	xv
Recommendations	xix
Chapter 1.....	1
Introduction	1
Chapter 2.....	5
Big health data: Australia's big potential.....	5
Introduction	5
Australia's potential	9
Australian framework.....	13
Chapter 3.....	21
Improving health policy development	21
Introduction	21
The traditional approach.....	21
New opportunities for health policy development	23
Chapter 4.....	31
Restrictions on linking MBS and PBS data	31
Australian Information Commissioner's view	36
Proposed privatisation of Medicare payment systems	38
Chapter 5.....	41
Issues and barriers for researchers.....	41
Introduction	41
Custodianship issues.....	42
Approvals processes	47

Linkage	50
The desultory departmental approach	54
Appendix 1	61
Witnesses who appeared before the committee	61
Appendix 2	65
Submissions received by the committee in relation to data linkage	65
Appendix 3	67
Additional information and answers to questions on notice	67
Tabled Documents	67
Answers to Questions on Notice	68
Additional Information	69
Correspondence	69
Appendix 4	71
Data collections held by departments	71

Terms of Reference

That a select committee, to be known as the Select Committee on Health, be established to inquire into and report on health policy, administration and expenditure, with particular reference to:

- a. the impact of reduced Commonwealth funding for hospital and other health services provided by state and territory governments, in particular, the impact on elective surgery and emergency department waiting times, hospital bed numbers, other hospital related care and cost shifting;
- b. the impact of additional costs on access to affordable healthcare and the sustainability of Medicare;
- c. the impact of reduced Commonwealth funding for health promotion, prevention and early intervention;
- d. the interaction between elements of the health system, including between aged care and health care;
- e. improvements in the provision of health services, including Indigenous health and rural health;
- f. the better integration and coordination of Medicare services, including access to general practice, specialist medical practitioners, pharmaceuticals, optometry, diagnostic, dental and allied health services;
- g. health workforce planning; and
- h. any related matters.

Acronyms and abbreviations

ABS	Australian Bureau of Statistics
ACIR	Australian Childhood Immunisation Register
ACS	Acute Coronary Syndrome
ACSQHC	Australian Commission on Safety and Quality in Health Care
AGIMO	Australian Government Information Management Office
AHES	Australian Health Economics Society
AIWH	Australian Institute of Health and Welfare
CAPHIA	Council of Academic Public Health Institutions Australia
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DHS	Department of Human Services
DSS	Department of Social Services
DVA	Department of Veterans' Affairs
GP	General Practitioner
GST	Goods and Services Tax
HIV	Human Immunodeficiency Virus
KPIs	Key Performance Indicators
MADIP	Multi-agency Data Integration Project
MBS	Medicare Benefits Schedule
MOU	Memorandum of Understanding
NACCHO	National Aboriginal Community Controlled Health Organisation
NCRIS	National Collaborative Research Infrastructure Strategy
NDIS	National Disability Insurance Scheme
NEHTA	National E-Health Transition Authority
NHMRC	National Health and Medical Research Council
NHPA	National Health Performance Authority
NT	Northern Territory
NSS	National Statistics Service
PAPT	Patient Admission Prediction Tool
PBS	Pharmaceutical Benefits Scheme
PHIDU	Public Health Information Development Unit

PHRN	Population Health Research Network
PIA	Privacy Impact Assessment
SURE	Secure Unified Research Environment
WA	Western Australia
WADLS	Western Australia Data Linkage System

Glossary

Aggregated data	means 'information about the population as a whole', such as the Census.
Big data	means 'high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight, decision making, and process optimization' ¹ An example would be an analysis of the 153 million records from six databases required to understand the unplanned hospital stays of Western Australian seniors.
Confidentialised data	See de-identified data .
Data linking	means 'the bringing together of two or more data sets to create a new, richer data set.' ² By bringing together sets of data that were previously isolated, researchers, clinicians and governments can deepen their understandings of the ways people actually use the health care system. This has the potential to inform government policy making and decisions about improving service delivery. ³
Data custodian	means 'agencies responsible for managing the use, disclosure and protection of source data used in a statistical data integration project. Data custodians collect and hold information on behalf of a data provider (defined as an individual, household, business or other organisation which supplies data either for statistical or administrative purposes). The role of data custodians may also extend to

1 Department of Finance and Deregulation, Australian Government Information Management Office, *The Australia Public Service Big Data Strategy*, August 2013, p. 8, www.finance.gov.au/sites/default/files/Big-Data-Strategy_0.pdf (accessed 8 January 2016).

2 The National Statistics Service is a network of Australian Government and State and Territory entities led by the Australian Bureau of Statistics that works together to improve Australia's statistics system. National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1, [www.nss.gov.au/nss/home.nsf/533222ebfd5ac03aca25711000044c9e/91242a5a14b12e26ca257ba8007b0819/\\$FILE/data%20linking%20w.pdf](http://www.nss.gov.au/nss/home.nsf/533222ebfd5ac03aca25711000044c9e/91242a5a14b12e26ca257ba8007b0819/$FILE/data%20linking%20w.pdf) (accessed 9 December 2015).

3 National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1.

producing source data, in addition to their role as a holder of datasets.'⁴

Data linkage key	means a data linkage key is a code that is constructed to replace identifying information, such as name, date of birth and address on a linked record in order to protect the privacy of the subjects of the study. By using a linkage key, researchers can link records that belong to the same person from multiple datasets without needing to know who the person is. ⁵
De-identified data	means 'a process by which a collection of data or information (for example, a dataset) is altered to remove or obscure personal identifiers and personal information (that is, information that would allow the identification of individuals who are the source or subject of the data or information).' ⁶
Enduring linkage	means that once links between datasets are created, the links are maintained and data, as it is progressively added to the dataset, is checked against the existing data to create an enduring research resource. ⁷
Ethics approval	means 'review of research by the Human Research Ethics Committee of the National Health and Medical Research Council or another body'. ⁸
Integrating Authority	means 'an authority responsible for the ongoing management of integrated data, ensuring it is kept secure, confidential and fit for the purposes of the approval process'. ⁹ The Australian Institute of Health and Welfare are an integrating authority.

4 National Statistics Service, 'Rights and responsibilities of data custodians', [www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/59fd060543b4e9e0ca257a4e001eacfe/\\$FILE/Rights,%20responsibilities%20and%20roles%20of%20data%20custodians_Dec2013.pdf](http://www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/59fd060543b4e9e0ca257a4e001eacfe/$FILE/Rights,%20responsibilities%20and%20roles%20of%20data%20custodians_Dec2013.pdf) (accessed 22 January 2015).

5 National Statistics Service, *Data Linking: Deterministic linking and linkage keys*, Information Sheet 3, p. 2.

6 Department of Finance, *Australian Public Service Better Practice Guide for Big Data*, January 2015, p. 30.

7 Dr Merran Smith, Chief Executive, PHRN, *Committee Hansard*, 11 December 2015, p. 23.

8 National Health and Medical Research Council, *National Statement on Ethical Conduct in Human Research (2007)*, May 2015, p. 88.

9 Cross Portfolio Statistical Integration Committee, *Data Integration Involving Commonwealth Data for Statistical and Research Purposes: Governance and Institutional Arrangements*, 6 October 2010, p. 5.

Metadata	is 'data about data'. ¹⁰ It can be of three types: definitional (provides definitions about aspect of the data), procedural operational (drives the collation of the data) or conceptual (dealing with concept sources and methods). ¹¹
Perturbed data	See confidentialised data
Public interest certificate	means an application made under the <i>Privacy Act 1988</i> for a certificate to engage in an act or practice that breaches, or may breach, an Australian Privacy Principle. ¹²
Unique identifier	means 'a number or code that uniquely identifies a person, business or organisation, such as a passport number or Australian Business Number'. ¹³

10 Prof Louisa Jorm, Director, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 13.

11 Mr Warren Richter, Head, Chief Information Officer Group, AIHW, *Committee Hansard*, 11 December 2015, p. 73.

12 See *Privacy Act 1988*, s. 72.

13 National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1.

Executive Summary

Australia has some of the world's highest quality and best value hospitals and primary care services. However, providing a world-class healthcare system is an expensive business. In 2013-14, the Commonwealth alone expended more than \$63 billion, the equivalent of 25 per cent of Australian Government tax revenue. Over the past decade overall health expenditure has grown at 5 per cent above the inflation rate.

In this context it is clear that new opportunities to evaluate current practices and deliver more effective and cost-efficient policies and programs should be vigorously pursued. 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.

Throughout this inquiry the committee heard of Australia's untapped potential to link health dataset such as births, childhood immunisation, Medicare Benefits Schedule (MBS), Pharmaceutical Benefits Scheme (PBS), hospital, aged care and deaths.

New opportunities for health policy development

The Department of Health provided the committee with a long list of significant benefits which data linkage can bring to the health system including:

- Better information to inform the government's policy decisions...
- A better understanding of what works, how well, for what cost, and in what circumstances...
- A more efficient health system, by supporting the most cost-effective treatments, strategies and interventions on broad-based independent evidence.

The committee has made several recommendations aimed at improving the development of innovative, evidence-based and more cost-effective health policy, which will lead to better patient outcomes and improve the standard of healthcare in Australia.

Linking Commonwealth and state datasets

Due to the shared responsibility for the development of health policy in Australia, significant quantities of health data is collected at both the Commonwealth and state levels. To fully realise the potential of linking health data, the Population Health Research Network explained the need to access de-identified Commonwealth and state data holdings:

Australia has a federated health system. The country also has high quality health data collections which can be used for planning and research.

However, because of the federated system, information about a person's lifetime health journey is collected and stored [by both the Commonwealth and the states]... In order to compare national trends and to evaluate the effectiveness of health policy for government and key policy decision makers it is necessary to be able to link this information together and use it in a timely fashion.

The committee has made several recommendations in this area aimed at streamlining data linkage across Commonwealth and state health datasets.

Restrictions on linking MBS and PBS data

Many submitters noted that significant health policy developments and medical research advances could be made if linked de-identified MBS and PBS data were more readily accessible. For example Professor Sallie-Anne Pearson, the Head of the Medicines Policy Research Unit at the Centre for Big Data Research in Health told the committee:

The linkage of PBS, MBS and other Commonwealth collections, such as those held by the Department of Social Services, can expand our opportunities to explore value, real-world use and pivotal issues such as equity of access... When I talk to consumers, they are surprised to learn that comprehensive postmarket surveillance research does not occur routinely in Australia. Why is this the case? Activity of this kind actually requires Commonwealth and state based data holdings to be linked... The currently fragmented data systems in Australia make it difficult, if not impossible, to systematically capture these impacts.

Linked MBS and PBS data is the fourth most requested data from the Australian Government. However, presently there are legislative restrictions and binding privacy guidelines that strictly constrain the linkage of de-identified MBS and PBS data.

The current privacy guidelines were made by the Privacy Commissioner in 2008 and provide that data from the PBS and MBS databases may only be linked:

- if it is necessary to comply with law;
- to determine eligibility for a benefit under one program, where eligibility depends upon services provided by the other program;
- where Medicare reasonably believes that doing so would prevent or lessen a serious and imminent threat to life or health; or
- for release where a person has provided their consent.

If linkage is undertaken for medical research purposes, the claims data can only be released where an individual has consented to having their data released and where the researcher undertakes to destroy the claims information provided to them at the conclusion of the research.

Changes in technology relevant for MBS-PBS linkage

One of the pioneers of Australian data linkage, Emeritus Professor D'Arcy Holman, has detailed the remarkable transformation technology has brought to data linkage techniques and the resultant privacy benefits:

...one might query if [data linkage] represents a significant invasion of privacy. To the contrary, the effects of data linkage on privacy have been exactly the opposite, with a profound privacy benefit compared with the way we did research before.

[Raw patient administrative data comprise] lots of documents, and now computer screens, liberally plastered with the patient's name and address. During the [1970s and 80s], I waded through countless thousands of [personal patient] records... Data linkage has turned this approach on its head...so that during the last 20 years, what I've worked with has looked like this: No names and addresses, age rather than date of birth...and just a number is used to represent each person... Nevertheless, use of the same number for the same anonymous person in each project, illuminates the crucial connections within and between different data collections, so that the outcomes can be measured.

The committee heard from representatives of the Department of Health who acknowledged the 'very strong concerns about privacy' which historically dominated departmental assessments of data requests. However officials noted the paradigm shift that has occurred:

...what has happened fairly recently is that there has been a significant cultural shift in the way data is regarded. It is regarded as an asset; it is regarded as a key tool in informing policy development and research. I think we are shifting from a culture of protecting data at all costs to one of protecting data but also identifying ways we can use it.

The need for review

The Acting Australian Information Commissioner and Privacy Commissioner, outlined his position that sometimes legislation needs to be revisited in light of technological changes:

Something that we find with a number of the laws that I deal with is that there is a need to review some of those because the situations change quite dramatically in terms of technologies you can use to bring together information sets and how they can be dispersed...

in 2011—I said that we were certainly open to having [the MBS-PBS legislative restrictions] looked at because it was, for want of a better description, an old piece of legislation that was developed at a different time when there were different community expectations and different mechanisms to simply store the information.

So what I am saying is yes—I think it is entirely appropriate to have that piece of legislation reviewed, to look for other mechanisms which may be able to make more efficient use of that information in terms of...freeing up data for good social policy purposes. But at the same time I would then say that if we are going to do that, what can we build in to ensure there is the right level of protection about that information in a newer environment of how it is going to be used?

In December 2015, the *Public Sector Data Management Report* called the current privacy arrangements 'over-cautious and cumbersome'.

The committee has recommended that the government undertake legislative and regulatory reviews of privacy arrangements with the aim of improving access to de-identified MBS and PBS datasets whilst maintaining robust and appropriate privacy protections. The committee's intention is to strengthen health policy evaluation and development as well as medical research undertaken in the public interest whilst ensuring effective protections for sensitive personal information.

By contrast, the committee believes that the government's planned privatisation of the Medicare payments system risks the unintended disclosure of sensitive MBS and PBS data and that this proposal should not proceed.

Improving access to health datasets

The committee was deeply concerned by the lack of transparency regarding the data held by government departments. If Australia is to maximise the value of its data resources, researchers need to know what data sets are available and how to access them.

During the course of this inquiry it became obvious that some departments were uncertain about what datasets they held.

It ought to be clear to both researchers and departments what data each department collects and what part of the department is responsible for data custodianship. The *Public Data Policy Statement*, the *Public Sector Data Management Report* and a chorus of witnesses all agreed that non-sensitive data should be publicly available as a right and that sensitive data should be accessible subject to appropriate privacy and security constraints. Accordingly, the committee has recommended that departments publicly list their dataset holdings on their websites and on data.gov.au. Departments should also to publish a statement that clearly explains their dataset approvals process.

The committee has also recommended that:

- by default, de-identified datasets should be released on an enduring basis;
- the government review the cost of data access and linkage;
- consideration be given to accrediting State data linkage units to link Commonwealth data with State data collections, subject to comprehensive privacy and security protocols;
- the government take a whole-of-government approach to streamlining the ethics approval process and the authorising environment; and
- departments set and publicly report on data linkage performance benchmarks.

Conclusion

Australia has the potential to significantly benefit from improved data linkage arrangements. This would lead to strengthened evidence-based policy making and advances in healthcare service delivery. The committee urges the government to seize this valuable opportunity to improve individual patient outcomes and raise the overall standard of healthcare in Australia.

Recommendations

Recommendation 1

2.36 The committee recommends that Australia forms partnerships with other countries engaged in data linking to ensure that Australian data access and linkage policies and regulations are developed to world's best practice.

Recommendation 2

3.37 The committee recommends that the Department of Health, as a high priority, actively explore and then implement measures to advance cost-effective, evidence-based policy development through the use of data linkage.

Recommendation 3

3.38 The committee recommends that relevant government departments should include information in their annual reports which describes the processes and projects being undertaken to establish evidence-based policy based on data linkage as well as strategies they have adopted to contribute to the government's public data policy.

Recommendation 4

4.40 The committee recommends that given the changes in technology, and mindful of the capacity and moral obligation for governments to hold and strongly secure personal data and privacy, the government review the operation of section 135AA of the *National Health Act 1953*, with the aim of improving access to de-identified MBS and PBS data for the purpose of health policy evaluation and development as well as research undertaken in the public interest.

Recommendation 5

4.41 The committee recommends that the Australian Information Commissioner, in consultation with privacy advocates, data custodians, academics and healthcare consumers, review the *Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programs* in order to ensure that the government:

- retains ownership and management of Australian MBS and PBS data and improves technological capacity to ensure the privacy of all Australians health data; and

- **develops a strategy to improve access to de-identified MBS and PBS data for the purpose of health policy evaluation and development as well as research undertaken in the public interest, in ways that don't decrease privacy.**

Recommendation 6

5.75 The committee recommends that each Australian Government agency develop and maintain on its website a list of datasets held by the agency along with the contact details of the data custodian. This list should be updated at least twice annually.

Recommendation 7

5.76 The committee recommends that all datasets held by the Commonwealth be listed on www.data.gov.au, identifying which agency is the data custodian.

Recommendation 8

5.77 The committee recommends that each Australian Government agency that is a data custodian develop and publish on its website guidance for researchers detailing its process for data requests and approvals.

Recommendation 9

5.80 The committee recommends that the government take a whole-of-government approach to streamlining the ethics approval process and the authorising environment in consultation with the Privacy Commissioner, privacy advocates, the NHMRC, data custodians, academics, consumers and the States and Territories. The government should also work with the States and Territories to establish a national accreditation system so that ethics approvals from accredited jurisdictions are recognised by the Commonwealth.

Recommendation 10

5.87 The committee recommends that relevant government agencies give greater priority to, and adequately resource, their data custodians.

Recommendation 11

5.88 The committee recommends that relevant government agencies provide guidance to data custodians to assist them in their decision-making, with a view to making more de-identified data available on an enduring basis.

Recommendation 12

5.89 The committee recommends that the government adopt the Productivity Commission's proposed principle that open access to de-identified datasets should be the default position.

Recommendation 13

5.90 The committee recommends that the government should direct relevant agencies to release de-identified datasets on an enduring basis as the default position.

Recommendation 14

5.91 The committee recommends that departments that have data custodianship responsibilities must establish and publish realistic Key Performance Indicators for the timely consideration and approval of datasets requests. These departments must publicly report on their KPIs in their annual reports.

If after 5 years departments continue to delay the release of datasets, then the committee recommends that the government establish binding timeframes for processing applications for data. Failure to comply with the timeframe should trigger appeal rights similar to those found in other information access regimes.

Recommendation 15

5.93 The committee recommends that Government encourage collaboration on data linkage projects between government agencies, as well as academia and industry to provide for evidence-based policy development and facilitate research that is undertaken in the public interest.

Recommendation 16

5.98 The committee recommends that government consider accrediting State data linkage units to link Commonwealth data with State data collections, subject to comprehensive privacy and security protocols.

Recommendation 17

5.99 The committee recommends that the Government review the cost of data access and linkage work undertaken by Commonwealth entities with a view to facilitating research and innovation in the national interest.

Chapter 1

Introduction

1.1 On 25 June 2014, the Senate established the Senate Select Committee on Health.¹ The final reporting date for the committee is 20 June 2016. The committee's resolution allows the committee to make interim reports such as this one which focusses on creating new opportunities for public good medical research and policy design from the government's vast holdings of health data.

Public hearings

1.2 The committee has completed 48 public hearings to date. A list of hearings which focused on big linkage is at Appendix 1.²

1.3 Through its extensive program of public hearings, the committee has taken evidence from many health experts, researchers, practitioners, consumers and communities.

1.4 Throughout the committee's inquiry, witnesses and submitters have raised issues with the committee relating to the collection, access, linkage and use of medical data for research purposes. In order to examine these issues in more detail, the committee held three hearings focussing specifically on the use and linkage of medical related data:

- 11 December 2015, Sydney; and
- 2 and 3 February 2016, Canberra.

Submissions

1.5 The committee has received 203 submissions since the beginning of its inquiry. In relation to data linkage issues, the committee has received 29 submissions. A list of submissions relating to data linkage is at Appendix 2.³

1.6 The committee's terms of reference are wide-ranging. It is the committee's intention to explore various issues in depth over the course of its inquiry. While the committee is still accepting general submissions, it is the committee's intention to seek submissions on specific topics as the need arises over the course of the inquiry.

1.7 Additional information, tabled documents, correspondence and answers to questions on notice received by the committee which relate to data linkage are listed at Appendix 3.⁴

1 *Journals of the Senate*, 25 June 2014, pp 996–998.

2 Public hearing details can also be accessed via the committee's website: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Public_Hearings.

3 The submissions received by the committee can be accessed via the committee's website: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Submissions.

Health Committee's first interim report

1.8 The committee's first interim report was tabled on 2 December 2014.⁵ That report detailed the committee's findings and conclusions at that time, focussing on issues raised during the committee's hearings and through submissions. Key areas of focus in the first report were:

- the government's proposed patient co-payments, cuts to hospital funding and the abolition of Australian National Preventative Health Agency;
- the government's plan to close the 61 Medicare Locals and replace them with 30 Primary Health Networks; and
- the merger of the Organ and Tissue Authority and the National Blood Authority.

Second interim report

1.9 The committee's second interim report was tabled on 24 June 2015.⁶ That report encompassed the committee's findings regarding the government's primary healthcare and general practice policies. In particular the report was a record of the government's frequent changes of policy since the 2014-15 Budget. The second interim report focused specifically on:

- the vital importance of general practice and primary healthcare and the threat posed by the government's numerous policy changes since the 2014-15 Budget;
- the responses of GPs and the primary healthcare sector to the government's various primary healthcare policies; and
- an examination of the 2015-16 Budget's health measures and commentary from stakeholders.

Third interim report

1.10 The committee's third interim report was tabled on 17 September 2015.⁷ That report examined the government's proposed privatisation of Australian Hearing and the National Acoustics Laboratories.

1.11 The report outlined the evidence taken at the 10 July 2015 public hearing and the related written submissions made by witnesses. It also examined:

- the impacts privatisation would have on users of the Australian Hearing services; and

4 The submissions received by the committee can be accessed via the committee's website: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Additional_Documents.

5 *Journals of the Senate*, 2 December 2014, p. 1948. The report can be accessed at: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/First_Interim_Report.

6 *Journals of the Senate*, 24 June 2015, p. 2809. The report can be accessed at: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Second_Interim_Report.

7 *Journals of the Senate*, 17 September 2015, p. 3158. The report can be accessed at: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Third_Interim_Report.

-
- the relationship between the National Disability Insurance Scheme (NDIS) and Australian Hearing.

Fourth interim report

1.12 The committee's fourth interim report was presented out of session on 8 October and tabled in the Senate on 12 October 2015.⁸ That report examined mental health issues raised by witnesses and submitters in the context of the National Mental Health Commission's *National Review of Mental Health Programs and Services Report*.

1.13 The fourth interim report specifically focussed on:

- the findings of the Commission's report;
- the Government's response to the Commission's findings;
- the importance of ensuring that there is a smooth transition of mental health programs into the NDIS.

Fifth interim report

1.14 The committee's fifth interim report was presented on 28 April 2016.⁹ That report examined the recent re-emergence of the debilitating lung disease Coal Workers' Pneumoconiosis (also known as black lung disease) in former Queensland coal miners. The report highlighted issues around safe exposure levels, dust monitoring regimes and worker screening practices.

Structure of this report

1.15 This sixth interim report examines the 'big data' and data linkage issues witnesses and submitters have raised with the committee. In addition to this introductory chapter, the report includes three chapters:

- background—setting the broad context of what 'big data' is and how it could be used in Australia (Chapter 2);
- Chapter 3 examines the new opportunities for governments to use big data in health policy development;
- Chapter 4 considers the constraints applying to linking data from the Medicare Benefits Schedule and Pharmaceutical Benefits Scheme; and
- Chapter 5 examines issues relating to access to data including ethics approvals, issues faced by data custodians, the 'linkage logjam' and its relationship to the Australian Government departments.

8 *Journals of the Senate*, 12 October 2015, p. 3175. The report can be accessed at: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Fourth_Interim_Report.

9 *Journals of the Senate*, 2 May 2016, p. 4165. The report can be accessed at: www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Fifth_Interim_Report.

Notes on references

1.16 References to submissions in this report are to individual submissions received by the committee and published on the committee's website. References to the committee Hansards are to the proof transcripts.¹⁰

Acknowledgements

1.17 The committee thanks the many organisations and individuals who participated in the big data and data linkage hearings on 11 December 2015 and 2 and 3 February 2016 as well as those that made written submissions. The committee also acknowledges the contribution of all those who have raised data issues in the committee's previous hearings.

10 Committee Hansards can be accessed via the committee's website:
www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Public_Hearings.

Chapter 2

Big health data: Australia's big potential

Introduction

2.1 Big data has the potential to create big opportunities for Australia. A recent estimate by Lateral Economics suggests that open government data could contribute up to \$25 billion per annum across the economy.¹ This analysis also suggests that Australian government held health-specific data alone could account for an increase of \$5.9 billion per annum.²

2.2 Big data also creates opportunities for considerable savings to the Australian health care system. Professor Fiona Stanley, Patron and former director of the Telethon Kids Institute told the committee that significant gains could be made with the health budget if government appropriately harnessed linked health data. Professor Stanley suggested that linked data could be used to reduce costly but ineffective clinical interventions, detect and prevent harmful health outcomes through early intervention and also alert regulators to fraud in the healthcare system.³

2.3 These are just some of the potential benefits Australia may obtain if the Australian Government and the States and Territories combined and fully utilised their administrative datasets.

2.4 Over the last three years, Australian Public Service agencies have been working together to promote a new approach to using and releasing datasets held by the Australian Government.⁴

2.5 On 7 December 2015 the Prime Minister, the Hon Mr Malcolm Turnbull MP and the Minister for Industry, Innovation and Science, the Hon Mr Christopher Pyne MP, launched the National Innovation and Science Agenda.⁵ One of the agenda's key planks was for government to 'lead by example in the way Government invests in and

1 Lateral Economics, *Open for Business: How Open Data Can Help Achieve the G20 Growth Target*, June 2014, p. 23.

2 Lateral Economics, *Open for Business: How Open Data Can Help Achieve the G20 Growth Target*, June 2014, p. x.

3 Prof Fiona Stanley, Patron and former director, Telethon Kids Institute, *Committee Hansard*, 2 February 2016, p. 23.

4 Ms Gemma Van Halderen, General Manager, Australian Bureau of Statistics, *Committee Hansard*, 3 February 2016, p. 16.

5 The Hon Mr Christopher Pyne MP, Minister for Industry, Innovation and Science and the Hon Malcolm Turnbull MP, Prime Minister, joint media release, 'National Innovation and Science Agenda', 7 December 2015, p. 1, www.minister.industry.gov.au/ministers/pyne/media-releases/national-innovation-and-science-agenda (accessed 25 January 2016).

uses technology and data to deliver better quality services'.⁶ This announcement coincided with the release of the *Public Sector Data Management report* and the *Public Data Policy Statement*.⁷ The report and the statement are considered at paragraphs 2.50–2.56 below.

2.6 The committee has previously heard from the Population Health Research Network (PHRN) in October 2014 about some of the challenges faced in maintaining health data linkages and in encouraging custodians of health data to be more open in releasing their data sets.⁸ These and similar concerns from other witnesses prompted the committee to initiate this current examination of issues relating to big data and data linkage.⁹

2.7 This chapter will consider the meaning of data linking and the new opportunities for Australia to harness the full benefits of big data and data linkage. This will be considered having regard to the existing framework and the government's recently announced data policies.

2.8 There are some key concepts that are important for this report. These include: big data, data linkage, data custodianship, unit record level data and data linkage keys.

Big data

2.9 The phrase 'big data' has been defined to mean 'high-volume, high-velocity and/or high-variety information assets that demand cost-effective, innovative forms of information processing for enhanced insight, decision making, and process optimization'.¹⁰

6 The Hon Mr Christopher Pyne MP, Minister for Industry, Innovation and Science and the Hon Malcolm Turnbull MP, Prime Minister, joint media release, 'National Innovation and Science Agenda', 7 December 2015, p. 1.

7 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, 3 December 2015, www.dpmc.gov.au/pmc/publication/public-sector-data-management-report (accessed 5 January 2016); Department of the Prime Minister and Cabinet, *Australian Government Public Data Policy Statement*, 7 December 2015.

8 Professor Brendon Kearney, Chair and Dr Merran Smith, Executive Officer, PHRN, *Committee Hansard*, 9 October 2014, p. 23; PHRN, *Answers to Questions on Notice 13*, 9 October 2014, pp 1–4.

9 See Ms Alison Verhoeven, Chief Executive, Australian Healthcare and Hospitals Association, *Committee Hansard*, 4 November 2015, p. 55; Professor Julie Byles, Director, Research Centre for Generational Health and Ageing, University of Newcastle, *Committee Hansard*, 27 November 2015, p. 34.

10 Department of Finance and Deregulation, Australian Government Information Management Office, *The Australia Public Service Big Data Strategy*, August 2013, p. 8, www.finance.gov.au/sites/default/files/Big-Data-Strategy_0.pdf (accessed 8 January 2016).

2.10 Examples of big health data include:

- analysing the Australian Childhood Immunisation Registry and all childhood immunisation records in Western Australia and New South Wales, involving the analysis of 1.8 million records;¹¹ and
- an analysis of unplanned hospital stays for Western Australian seniors, requiring the linkage of 153 million digital records from six data collections.¹²

Data linking

2.11 Data linking is the bringing together of two or more data sets to create a new, richer data set.¹³ By bringing together sets of data that were previously isolated, researchers, clinicians and governments can deepen their understandings of the ways people actually use the health care system. This has the potential to inform government policy making and decisions about improving service delivery.¹⁴

Data custodianship

2.12 According to the National Statistics Service, data custodians are:

...agencies responsible for managing the use, disclosure and protection of source data used in a statistical data integration project. Data custodians collect and hold information on behalf of a data provider (defined as an individual, household, business or other organisation which supplies data either for statistical or administrative purposes). The role of data custodians may also extend to producing source data, in addition to their role as a holder of datasets.¹⁵

2.13 For example the Department of Health is the custodian of the Medicare Benefits Schedule data.¹⁶

11 Dr Heather Gidding, Senior Lecturer, *Committee Hansard*, 11 December 2015, p. 36.

12 E/Prof D'Arcy Holman, Additional Information 13, p. 20; Dr Sylvie Price et al, 'Use of case-time-control design in pharmacovigilance applications: exploration with high-risk medications and unplanned hospital admissions in the Western Australian elderly', *Pharmacoepidemiology and Drug Safety*, vol. 22, no. 11, pp 1159–1170.

13 The National Statistics Service is a network of Australian Government and State and Territory entities led by the Australian Bureau of Statistics that works together to improve Australia's statistics system. National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1, [www.nss.gov.au/nss/home.nsf/533222ebfd5ac03aca25711000044c9e/91242a5a14b12e26ca257ba8007b0819/\\$FILE/data%20linking%20w.pdf](http://www.nss.gov.au/nss/home.nsf/533222ebfd5ac03aca25711000044c9e/91242a5a14b12e26ca257ba8007b0819/$FILE/data%20linking%20w.pdf) (accessed 9 December 2015).

14 National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1.

15 National Statistics Service, 'Rights and responsibilities of data custodians', [http://www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/59fd060543b4e9e0ca257a4e001eacfe/\\$FILE/Rights,%20responsibilities%20and%20roles%20of%20data%20custodians_Dec2013.pdf](http://www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/59fd060543b4e9e0ca257a4e001eacfe/$FILE/Rights,%20responsibilities%20and%20roles%20of%20data%20custodians_Dec2013.pdf) (accessed 22 January 2016).

16 Department of Health, *Answers to Questions on Notice 59*, p. 2.

Unit record level data

2.14 A distinction needs to be made between individual unit records and aggregated data. Aggregated data provides information about a population as a whole and no individual can be identified from that data.¹⁷ An example of aggregated data is the Census.

2.15 This can be contrast with unit record level data which, according to the Australian Bureau of Statistics is:

...a file of responses to ABS surveys or censuses that have had specific identifying information about persons and organisations confidentialised. [The unit record level data files] contain very detailed information for each individual record - a record can be a person, a business, a family, household or a job for example.¹⁸

2.16 For researchers that wish to understand the health system or are interested in a particular pharmaceutical product, it is preferable to have de-identified unit level records as Dr Merran Smith, Chief Executive of the PHRN explains:

Aggregated data is valuable and even linked aggregated data is valuable. But it probably cannot do the sorts of things we are talking about for the health/medical research that really needs the detail.¹⁹

2.17 For that reason, researchers need access to de-identified unit record level data to achieve the best result.

Data linkage key

2.18 A data linkage key is a code that is constructed to replace identifying information, such as name, date of birth and address on a linked record in order to protect the privacy of the subjects of the study. By using a linkage key, researchers can link records that belong to the same person from multiple datasets without needing to know who the person is.²⁰

2.19 Additional terms used in this report may be found in the Glossary.²¹

17 Australian Bureau of Statistics, *2011 Census privacy & confidentiality*, 23 November 2015, abs.gov.au/websitedbs/censushome.nsf/home/privacy?opendocument&navpos=130, (accessed 15 February 2016).

18 Australian Bureau of Statistics, *About CURF microdata*, 10 February 2016, abs.gov.au/websitedbs/D3310114.nsf/home/About+CURF+Microdata, (accessed 15 February 2016).

19 Dr Merran Smith, Chief Executive, PHRN, *Committee Hansard*, 11 December 2015, p. 28.

20 National Statistics Service, *Data Linking: Deterministic linking and linkage keys*, Information Sheet 3, p. 2.

21 See p. xi.

Australia's potential

2.20 Data is an important and valuable government resource. Data linking has the capacity to maximise that resource and to create new opportunities for more complex and expanded evidence-based policy and research.²² Professor Stanley highlighted the benefits to government of using more linked data:

...[Australia] would be second to none in the world in enabling us to evaluate all the outcomes of all [government] services that are provided. [Australia] would be able to influence and evaluate evidence based practice; we would be able to look at the epidemiological trends and risk factors of major and costly problems.²³

2.21 In the medical sphere there are some shining examples of how data linking has improved health outcomes. For instance, data linking has helped to identify the role of folate in pregnancy in reducing neural tube defects, such as spina bifida.²⁴ The Northern Territory Government facilitated 'a study that reviewed the association between primary care utilisation and the number of hospital admissions for the NT remote Aboriginal population'.²⁵

2.22 Linked data sets have also been used to 'estimate the prevalence of dementia in the NT Aboriginal and non-Aboriginal populations' and analyse the 'cost effectiveness of primary care in the management of diabetes'.²⁶

2.23 The Commonwealth Scientific and Industrial Research Organisation (CSIRO) has used linked data to create a Patient Admission Prediction Tool (PAPT) that is helping to make hospitals more efficient.²⁷ The tool uses historical data from emergency departments and hospital data sets to model the number of patients that are likely to present at the emergency department and the numbers that are likely to require admission to wards. The CSIRO notes that improved access to hospital

22 National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1.

23 Prof Fiona Stanley, Patron and former director, Telethon Kids Institute, *Committee Hansard*, 2 February 2016, pp 18–19.

24 National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1.

25 Northern Territory Government, *Submission 148*, p. 3; Yuejen Zhao, et al, 'The relationship between number of primary health care visits and hospitalisations: evidence from linked clinic and hospital data for remote Indigenous Australians', *BMC Health Services Research*, vol. 13, p. 466.

26 Northern Territory Government, *Submission 148*, p. 4; Shu Qin Li, et al, 'Dementia prevalence and incidence among the Indigenous and non-Indigenous populations of the Northern Territory', *Medical Journal of Australia*, vol. 200, no. 8, pp 465–469; Susan L Thomas, 'The cost-effectiveness of primary care for Indigenous Australians with diabetes living in remote Northern Territory communities', *Medical Journal of Australia*, vol. 200, no. 11, pp 658–662.

27 CSIRO, *Submission 174*, p. 5.

datasets held by the Australian Government would ensure that decisions could be made on the most comprehensive data available.²⁸

2.24 Many witnesses argued that governments could facilitate a greater degree of health data linkage, thereby releasing significant untapped opportunities. For instance the Council of Academic Public Health Institutions Australia (CAPHIA) noted that linking State and Australian Government datasets has:

...the potential for national, state and local comparative effectiveness, clinical trials and registry research that has thus far been largely untapped, to drive health policy, redesign, quality improvement and evidence translation in health care. Additionally, it enables...the rigorous objective evaluation of health policy for government and key policy professionals; and the ability to compare trends nationally, to identify programs that deliver value for money and to avoid wasting resources on those that are not delivering. The result is better targeted, evidence-based and more cost-effective health policy, services and interventions for the Australian community.²⁹

2.25 In addition to the excellent research outlined in paragraphs 2.21–2.22, the Northern Territory submitted that the following opportunities may be possible if more Australian Government datasets were accessible:

Geographic distribution of Medicare and PBS [Pharmaceutical Benefits Scheme] funded service access mapped against state based services or health need,

Socioeconomic distribution of Medicare and PBS funded service access

Associations between utilisation of Medicare funded services and hospital and/or [Emergency Department] services...

The distribution of PBS funded items and measures of health need.

Quality and safety measures of primary care, by linking Medicare or PBS items and outcomes such as diabetic control, hospitalisation and mortality.³⁰

2.26 The Australian Government also acknowledged the latent potential of data linkage. For example Department of Health representative Ms Alanna Foster, First Assistant Secretary told the committee:

Linked data would also enable understanding of the full extent of patients' health-service usage—that is, it would be possible to follow patients' pathways through the system and answer questions about patient populations, such as: are the high users of primary care also high users of the hospital system? If we provide better access to chronic disease management in primary care are patients less likely to present to hospital?

28 CSIRO, *Submission 174*, p. 5.

29 CAPHIA, *Submission 173*, p. 1.

30 Northern Territory Government, *Submission 148*, p. 5.

What interactions do patients have, with their General Practitioners (GPs), when they leave hospital?

With big-data technologies linking and advanced analytic capabilities, we could, for example, use pattern mining to quickly identify adverse events that may arise from medical devices or health services, use cluster analysis to assign patients to like groups—for example, identifying groups with diabetes or cardiovascular conditions that may be amenable to policy intervention and then model the impacts of those imperfections, in terms of costs and patient outcomes. We could use pathways analysis to investigate how patients—for example, cancer patients—are moving through the health system and model the impact of policy interventions targeted at improving these pathways. These are just some of the tools that could be used when forming government decision making and the work of researchers.³¹

International standing

2.27 The Australian experience stands in stark contrast to those of other developed economies that have already liberalised their use of administrative data. In 2013 the Productivity Commission reported that:

In Denmark, Sweden, Finland and the Netherlands, linked administrative data are accessible for research purposes. Statistics Finland considers that statistics should be compiled from administrative records whenever possible — around 96 per cent of its data come from these sources. This openness promotes research — ‘microsimulation specialists pour into Nordic countries because of their liberal approach towards sharing statistics’...³²

2.28 Meanwhile, Australian researchers, frustrated at the relative inaccessibility of Australian datasets are choosing to use datasets from other countries. For instance Professor Philip Clarke, Professor of Health Economics at the University of Melbourne informed the committee:

Other countries have very good datasets. I have done work with Scandinavian registries in diabetes. They make those available... I am currently building a cardiovascular health policy model with funding from the NHMRC [National Health and Medical Research Council], but explicitly in my application I said I would be using New Zealand data, because there was no appropriate Australian data. I am able to work with researchers at the University of Auckland. There are half a million clinical records with cardiovascular patients that have had their cardiovascular risk assessed. Those have been linked to hospital records and medical records, and I am able to work with researchers almost immediately to start

31 Ms Alanna Foster, First Assistant Secretary, Department of Health, *Committee Hansard*, 11 December 2015, p. 64.

32 Productivity Commission, *2012-13 Annual Report*, p. 12 (references omitted).

analysing that. I would be dreaming if I thought that could happen in Australia within the next few years.³³

2.29 Australia is missing out on important opportunities to identify health risks for our own population because Australian Government datasets are inaccessible. This is particularly the case with pharmaceutical safety. Professor Sallie-Anne Pearson, Head of the Medicines Policy Research Unit at the Centre for Big Data Research in Health noted that data inaccessibility has meant that medicine safety research is not commonly undertaken in Australia:

...fewer than 30 studies have examined drug safety in the last 25 years. This needs to change. Australia is actually well-placed to deeply understand our return on PBS investment, and also other health programs. The data already exists. We have information that covers our entire population.³⁴

2.30 The lack of research is surprising when there are 190 000 hospitalisations caused by medications in Australia every year at a cost of \$660 million to the health care system.³⁵

2.31 Witnesses told the committee that Australia could safely exploit the existing PBS data for the benefit of Australians. Dr Barbara Mintzes, Senior Lecturer in Pharmacy at the University of Sydney informed the committee of the approach of several other developed countries:

The experience to date in Canada, the US, the UK and Scandinavia makes it clear that these databases are important tools for medication safety and protection of public health.³⁶

2.32 In some cases Australia has been collecting data for years but without fully utilising the data, its collection is rendered fruitless. As Professor Fiona Stanley identified:

My biggest anguish has been that over 30 years of setting up a birth defects registry to find the next thalidomide, another one could be happening all the time and we are unable to detect it.³⁷

2.33 In 2015 the Productivity Commission attempted to articulate why Australia was falling behind other developed countries in releasing administrative data. In its

33 Prof Philip Clarke, Professor of Health Economics, University of Melbourne, *Committee Hansard*, 11 December 2015, pp 43–44.

34 Professor Sallie-Anne Pearson, Head, Medicines Policy Research Unit, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 7.

35 Dr Barbara Mintzes, Senior Lecturer, Faculty of Pharmacy, University of Sydney, *Committee Hansard*, 11 December 2015, p. 1.

36 Dr Barbara Mintzes, Senior Lecturer, Faculty of Pharmacy, University of Sydney, *Committee Hansard*, 11 December 2015, p. 2.

37 Professor Fiona Stanley, Patron and former director, Telethon Kids Institute, *Committee Hansard*, 2 February 2016, p. 24.

Efficiency in Health research paper the Productivity Commission suggested several reasons including:

- concerns about privacy;
- that processes for accessing administrative data were poorly structured and did not encourage researchers;
- a lack of transparency about what data government holds; and
- a tendency for data owners to develop costly ad hoc datasets rather than developing enduring continuous datasets for use by multiple researchers.³⁸

2.34 The Productivity Commission concluded that:

The potential of administrative data is not being realised in Australia, and the lost opportunities will only grow as technology continues to open up new ways to use and analyse data. Calls to release and better link administrative datasets have been made previously by the Commission and by others.³⁹

Committee view

2.35 The evidence heard by the committee and received in submissions suggests that Australia has significant health data assets and medical research capabilities. The evidence also clearly demonstrates that in comparison to other countries Australia is failing to capitalise on its data potential.

Recommendation 1

2.36 The committee recommends that Australia forms partnerships with other countries engaged in data linking to ensure that Australian data access and linkage policies and regulations are developed to world's best practice.

Australian framework

2.37 As the Productivity Commission and other experts have noted, the factors that are holding Australia back are largely barriers erected by the legislative framework or its application by the public service. The blockage is not in technical expertise or infrastructure. Australia has a world leading data linkage system and many talented researchers and academics in the field.

38 Productivity Commission, *Efficiency in Health*, Research Paper, April 2015, p. 83 (references omitted).

39 Productivity Commission, *Efficiency in Health*, Research Paper, April 2015, p. 83 (references omitted).

Experience and history

2.38 Australia's modern data linkage capacity dates back to 1995. Before this time, some statistics were collected but as Emeritus Professor D'Arcy Holman, formerly a Professor of Public Health at the University of Western Australia noted 'what we could do with health statistics...was severely constrained by the technical infrastructure available to us'.⁴⁰

2.39 That changed in 1995 when the Western Australian Data Linkage System (WADLS) was established.⁴¹ The formation of the WADLS allowed population health researchers to:

...map over 30 pre-existing health databases on the people of WA. The links mean that the journeys of individuals through the health system can be followed anonymously over many years and thus their risk factors for major diseases, and the use and outcomes of health services can be evaluated using anonymous information.⁴²

2.40 More information on the change in the use of technology and how improvements in technology are being used to protect privacy can be found in Chapter 3.

2.41 At the Australian Government level there is a restriction on who can perform the data linkage function. The Australian Government requires that only certain accredited 'integrating authorities' may link Australian Government data. More information on integrating authorities can be found in Chapter 3.

2.42 Each State and Territory either has its own data linkage unit or is associated with a data linkage unit.⁴³ In 2004 the Australian Government established the National Collaborative Research Infrastructure Strategy (NCRIS). Through NCRIS the government provided \$20 million to establish the PHRN.⁴⁴ The PHRN is a national network that works to support collaboration between data linkage units and further Australia's linkage potential.

State / Commonwealth divide

2.43 Witnesses told the committee that Australia's federal constitution contributes to its data challenges. As Emeritus Professor Holman noted:

Australia differs from other federations, Canada for example, in that our [Australian] Government has not directed its financial support for these

40 E/Prof Holman, *Additional Information 13*, p. 3.

41 Bureau of Health Information, *Tabled Document 90*, p. 6.

42 E/Prof Holman, *Additional Information 13*, p. 5.

43 PHRN, *Who is involved?*, www.phrn.org.au/about-us/who-is-involved/ (accessed 16 February 2016).

44 NCRIS, *Population Health Research Network*, 25 November 2013, p. 1, https://docs.education.gov.au/system/files/doc/other/population_health_research_network.pdf (accessed 8 December 2015).

integral components of health care through the states, but has established itself as a separate vertical player.⁴⁵

2.44 This State / Commonwealth divide means that the Australian Government collects primary health and aged care data whilst the States collect hospital, births, deaths and cancer information. A list of the Australian Government's major health related data holdings can be found in Appendix 4.

2.45 One of the challenges to sharing data between the Australian Government and the States and Territories has been a reticence by Australian Government departments to release data based on privacy concerns. Ms Alanna Foster, First Assistant Secretary of the Department of Health insisted that 'due to the separate legislative requirements, it can be challenging to link these datasets while also adhering to strict privacy guidelines'.⁴⁶

2.46 One of these privacy guidelines requires that MBS [Medicare Benefits Schedule] and PBS data cannot be linked and another requires that Australian Government data linkages must be destroyed at the conclusion of the project.⁴⁷ These two restrictions will be considered in greater detail in Chapters 3 and 4 respectively.

2.47 Despite these restrictions, Professor Clarke told the committee that 'there have been linkages but they tended to be sporadic'.⁴⁸

2.48 However, Emeritus Professor D'Arcy Holman described the period between 2007 and 2012 in Western Australia when 'things were different'. This was because, as Emeritus Professor Holman recalled:

The two separate information systems [the Australian and Western Australian] were permitted to talk one with the other.

A short reprieve of different senior administration in the [Australian Government] led to a collaboration with the State to include the Medicare, pharmaceutical and aged care data within the WADLS system. This was the first and only instance since federation that the [Australian Government] and an Australian State agreed to integrate their data in a functional way to create a total picture of health system performance.⁴⁹

45 E/Prof Holman, *Additional Information 13*, p. 11.

46 Ms Alanna Foster, First Assistant Secretary, Department of Health, *Committee Hansard*, 11 December 2015, p. 64.

47 *Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programs*, Guideline 6.2, www.comlaw.gov.au/Details/F2008L00706 (accessed 25 January 2016).

48 Prof Philip Clarke, Professor of Health Economics, University of Melbourne, *Committee Hansard*, 11 December 2015, p. 42.

49 E/Prof Holman, *Additional Information 13*, p. 13.

Recent developments

2.49 In late 2015, government attitudes toward sharing data started to change. On 3 December 2015, the Department of the Prime Minister and Cabinet released the *Public Sector Data Management Report*.⁵⁰

2.50 The report sets out a roadmap towards the regular and systematic release of public sector data and highlights the need to reform certain areas to enable the Australian Public Service to get the most out of Australia's data holdings.⁵¹

2.51 On 7 December 2015, the Department of the Prime Minister and Cabinet released the *Australian Government Public Data Policy Statement*.⁵² The statement declares that Australian Government entities will:

- make high-value data available for use by the public, industry and academia, in a manner that is enduring and frequently updated using high quality standards...
- securely share data between Australian Government entities to improve efficiencies, and inform policy development and decision-making;
- engage openly with the States and Territories to share and integrate data to inform matters of importance to each jurisdiction and at the national level;
- ensure all new systems support discoverability, interoperability, data and information accessibility and cost-effective access to facilitate access to data.⁵³

2.52 Whilst this was seen as a welcome development, it was a surprise to many non-government witnesses who told the committee that they had not been consulted

50 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, 3 December 2015.

51 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, 3 December 2015, pp 21 and 36.

52 Department of the Prime Minister and Cabinet, *Australian Government Public Data Policy Statement*, 7 December 2015.

53 Department of the Prime Minister and Cabinet, *Australian Government Public Data Policy Statement*, 7 December 2015, p. 2.

and were not aware that the government had been working on the policy statement or the data management report.⁵⁴

2.53 When Ms Helen Owens, Assistant Secretary of the Department of the Prime Minister and Cabinet was asked who the government consulted she listed:

...organisations like Telstra, Google, the World Bank, the [Australian Broadcasting Corporation], [software producer] IBM, [software company] SAP. We also spoke with some research institutions—the Grattan Institute and the Crawford school at ANU. [The Department of the Prime Minister and Cabinet] then did some individual consultations with business leaders in the data space and open data space.⁵⁵

2.54 The Office of the Australian Information Commissioner was nominally consulted in the development of both the *Public Sector Data Management Report* and the *Public Data Policy Statement*.⁵⁶ However, the government did not consult the National Health Performance Authority (NHPA), the National E-Health Transition Authority (NEHTA) or the Australian Commission on Safety and Quality in Health Care in the development of either document.⁵⁷

2.55 Turning the report and the statement into a reality will take commitment and perseverance, something previous governments have promised in this space but not delivered.⁵⁸ As the Productivity Commission stated in their *2012-13 Annual Report*:

Realising these goals [harnessing administrative data to support research and evidence-based policy evaluation] requires political will, articulated at the highest levels, to persevere with a concerted strategy with clear timeframes based on the principle that open access to de-identified

54 See for example Prof Louisa Jorm, Director, Centre for Big Data Research in Health, Prof Sallie-Anne Pearson, Head – Medicines Policy Research Unit, Centre for Big Data Research in Health, Dr Julian Elliott, Australasian Cochrane Centre, *Committee Hansard*, 11 December 2015, p. 20; Prof Brendon Kearney, Chair, PHRN, *Committee Hansard*, 11 December 2015, p. 28; Prof Philip Clarke, Professor of Health Economics, University of Melbourne, Dr Heather Gidding, Senior Lecturer, University of New South Wales, Dr David Hansen and Prof David Preen, *Committee Hansard*, 11 December 2015, p. 46; Ms Fiona Brooke, Policy Adviser, NRHA, *Committee Hansard*, 11 December 2015, p. 53; Prof Fiona Stanley AC, Patron and former director, Telethon Kids Institute, *Committee Hansard*, 2 February 2016, p. 22; Ms Anne McKenzie, Consumer Advocate and Program Manager, University of Western Australia School of Population Health and Telethon Kids Institute, *Committee Hansard*, 2 February 2016, p. 22; Prof John Glover, Director, Public Health Information Development Unit, *Committee Hansard*, 3 February 2016, p. 4.

55 Ms Helen Owens, Assistant Secretary, Department of the Prime Minister and Cabinet, *Committee Hansard*, 3 February 2016, p. 17.

56 Office of the Australian Information Commissioner, *Correspondence 5*, 8 February 2016, pp 1–2.

57 Ms Bettina McMahon, Head of Assurance and Risk, NEHTA, *Committee Hansard*, 3 February 2016, p. 4; Department of Health, *Answer to Question on Notice 58*, p. 2.

58 Productivity Commission, *Annual Report 2012-13*, October 2013, pp 13–14.

information should be a default position. Realistically, it could take 5-10 years to rollout and embed systems before the ‘holy grail’ of relatively unimpeded remote access to high quality, de-identified and linked administrative data is achievable.

While there have been announcements and initiatives in the past and more recently, the lack of sustained tangible progress means that it is important that the 5-10 year timeframe does not become a motivation for more ‘false starts’, deferrals or eventual reprioritisation and non-delivery. International practices and over thirty years of experience in Western Australia suggest that the capabilities necessary to achieve a more open data culture could be developed by all Australian governments.⁵⁹

Committee view

2.56 The evidence presented to this committee demonstrates that Australia has the potential to create a world leading data linkage system that can both maintain data security and produce ground-breaking public health research.

2.57 The committee recognises that linking administrative data, which is already routinely collected, has the potential to reveal new insights about the ways Australians use the healthcare system and potential ways to improve the health outcomes of all Australians.

2.58 The opportunities Australia is squandering are not just possibilities for health improvements for future generations; but the ability to detect causes of harm to Australians. The committee has received evidence that Australia could be using its data resources to detect harmful prescription medications both in children and in adults. Instead, Australian researchers are forced to rely on studies conducted in other countries where such drug safety studies are possible. For the benefit of the health of all Australians we can and must do better.

2.59 Improving our data linkage system involves breaking down some of the historical barriers that have resulted from our federated system of government. We have seen in sporadic intervals that such cooperation is possible and can lead to highly beneficial outcomes.

2.60 Australia has the infrastructure and the knowledge to make a national data linkage system work but it will require legislative changes and cultural changes in the Australian Public Service. The nature of these challenges will be examined in greater detail in Chapters 3 and 4. These changes could catapult Australia to become a world leader in data linkage.

2.61 The committee welcomes the renewed focus on Australia's data assets and is encouraged by the attempt to coordinate efforts across government to make more datasets available. But the committee notes that there is still a long way to go to overcome many of the barriers currently faced by researchers and the valid community concerns regarding privacy.

59 Productivity Commission, *Annual Report 2012-13*, October 2013, p. 15.

2.62 The committee further notes that this is not the first time an Australian Government has promoted a more open approach to sharing data. The committee is concerned at the very limited nature of the government's consultation in developing its recent *Australian Public Data Policy Statement* and its *Public Sector Data Management report*. In compiling its most recent policies, the government obtained very limited input from key stakeholders, including those funded by the Australian Government. By failing to consult any health professionals it became manifestly clear that the use of health data was not a priority for the government. The committee is concerned by the low regard in which the government seems to hold health data and the research groups that work with it.

2.63 To ensure that the government's newly articulated approach to releasing data maximises Australia's big health data potential, while attending to valid community expectations about security and privacy around personal health data, the government must broaden its data policy engagement to include health-related academics, researchers and practitioners.

Chapter 3

Improving health policy development

Introduction

3.1 Australian governments rightly place a high priority on the health of their citizens. As a result Australia delivers some of the highest quality and best value hospitals and primary care in the world. However, a world-class healthcare system is an expensive business. In 2013-14, combined government health-related expenditure was greater than \$100 billion per annum.¹ The Commonwealth alone expended more than \$63 billion in that year, the equivalent of 25 per cent of Australian Government tax revenue.² Over the past decade overall health expenditure has grown significantly above the inflation rate at 5 per cent in real terms.³

3.2 At a time when the government is struggling to effectively manage the growing health budget, it is clear that new opportunities to evaluate current practices and deliver more effective and cost-efficient policies and programs should be vigorously pursued.

3.3 This chapter explores the new opportunity that big data provides to ensure that our health expenditure is as efficient as possible, and more importantly to improve the standard of healthcare in Australia.

The traditional approach

3.4 Witnesses told the committee that the current approach to government health policy evaluation and development lacks a firm evidence-base. For instance the Centre for Big Data Research in Health cited evidence underpinning the Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS) to highlight the limitations of the current approach:

Australian governments invest more than \$100 billion annually on healthcare, yet we have a relatively limited understanding of Australia's return on this investment. For example, the Medicare Benefits [Schedule] (Commonwealth spend approximately \$21 billion annually) consists of some 6000 items, but fewer than 5% have been assessed for safety, effectiveness and cost-effectiveness against contemporary evidence. Even when medical treatments have undergone extensive pre-market evaluation in randomised controlled trials, like all of the items listed on the Pharmaceutical Benefits Scheme (Commonwealth spend approximately \$9 billion annually), they are most often tested over relatively limited time

1 AIHW, *Health expenditure Australia 2013-14*, September 2015, p. viii.

2 AIHW, *Health expenditure Australia 2013-14*, September 2015, pp 15–16.

3 AIHW, *Health expenditure Australia 2013-14*, September 2015, p. viii.

frames, even if medicines are used for a lifetime, and in populations that do not represent the people using them in routine clinical care.⁴

3.5 In a recent research paper, the Productivity Commission pointed to potential negative consequences of data holders not permitting the release of administrative health data:

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.⁵

3.6 Dr Barbara Mintzes, a Senior Lecturer at the University of Sydney's Faculty of Pharmacy, gave the committee a concrete example of risks associated with the traditional approach to PBS listings:

When a medicine is first approved for marketing, we know very little about its safety, especially in the longer term. On average, drugs are approved based on the experience of around 2,000 people who have used the medicine for short periods of time. Once on the market, millions of people may use the same drug. This is what happened with the arthritis drug Vioxx [a prescription anti-inflammatory which was recalled in 2004]... In its five years on the market, Vioxx caused up to 140,000 heart attacks in the US.⁶

3.7 At a subsequent hearing, Professor Fiona Stanley who pioneered ground data-linked population health research in Western Australia in the 1970s and '80s, explained how the problems with Vioxx could have been mitigated through the use of data linkage:

I have one example around a PBS linkage to all the health outcomes... Vioxx was not picked up for a long time—perhaps for four or five years—because it caused a common outcome of heart attacks and heart-related deaths. However, if we had linked our PBS into all our health outcomes, how many deaths and serious, morbid and costly heart attacks could we have prevented in that four or five years? Hundreds and hundreds. In my opinion, not doing this linkage of PBS to health outcomes alone is actually negligent.⁷

3.8 The Australian Health Economics Society (AHES) pointed out that under current arrangements, certain 'basic questions' cannot be answered:

4 Centre for Big Data Research in Health, UNSW, *Submission 172*, p. 2.

5 Productivity Commission, *Efficiency in Health*, April 2015, pp 85–86, as cited by Research Australia, *Submission 182*, p. 9.

6 Dr Barbara Mintzes, Senior Lecturer, Faculty of Pharmacy, University of Sydney, *Committee Hansard*, 11 December 2015, p. 1.

7 Professor Fiona Stanley, AC, Patron and former Director, Telethon Kids Institute, *Committee Hansard*, 2 February 2016, p. 23.

Australia has been – and still is – lagging behind [the US, UK, Canada and New Zealand in the access and use of health care administrative data]. As a consequence, Australia has a poorer health economics and health services research infrastructure and many basic questions cannot be addressed (e.g. changes in the out-pocket payments by consumers using Medicare)...⁸

3.9 As a result the AHES submitted that Australia is foregoing 'considerable benefits in terms of understanding health system which can lead to both greater efficiency and improved health outcomes.'⁹ The AHES concluded that:

...research within government focuses on short term issues within electoral cycles and is not driven by an overarching research strategy that focuses on the key long term questions. As a consequence, key research questions and policy issues remain unanswered for decades and governments continue to revisit these issues inefficiently leading to waste.¹⁰

3.10 SA-NT DataLink highlighted the difficulties faced by state and territory governments in formulating their health policies due to the inaccessibility of Commonwealth data:

Lack of timely and affordable access to critical Commonwealth data (such as MBS, PBS, Centrelink) is a serious obstacle to developing a more informed understanding of health outcomes and services at the State/Territory levels.¹¹

3.11 Finally, the Centre for Big Data Research in Health argued that, given a multitude of modern-day pressures, the traditional approach is 'no longer adequate':

The increasing complexity of healthcare in terms of rapidly evolving and fragmented service delivery models, the disruptive impacts of new therapies and technologies, and people living longer with multiple health conditions means that traditional methods guiding health policy and practice are no longer adequate.¹²

New opportunities for health policy development

3.12 By contrast, a variety of submitters explained the significant benefits that could flow to the development of health policy if decision-makers had a more robust evidence-base.

3.13 The Department of Health provided the committee with a long list of 'significant benefits' which big data can bring to the health system:

- Better information to inform the government's policy decisions

8 Australian Health Economics Society, *Submission 184*, pp 3–4.

9 Australian Health Economics Society, *Submission 184*, p. 3.

10 Australian Health Economics Society, *Submission 184*, p. 4.

11 SA-NT DataLink, *Submission 181*, p. 2.

12 Centre for Big Data Research in Health, UNSW, *Submission 172*, p. 2.

- A clearer picture of the real experiences of patients as they engage with the health system
- A better understanding of what works, how well, for what cost, and in what circumstances
- Earlier detection of trends – both positive and negative
- Earlier detection of anomalous behavior and deviations from expected results
- A more efficient health system, by supporting the most cost-effective treatments, strategies and interventions on broad-based independent evidence.¹³

3.14 SA-NT DataLink drew the committee's attention to analysis by the Productivity Commission that highlighted the critical need for evidence-based policy:

[The Productivity Commission] recognised that the ability to undertake population based research by linking administrative data held by government agencies and other bodies is essential to supporting evidenced based policy. [The Productivity Commission] strongly argued the need for systematic evidence-based policy to ensure the effectiveness of the massive expenditure in the areas of health, welfare, education and other areas of Government activity. [The Productivity Commission] regarded the demonstrable effectiveness of this expenditure in achieving the planned for outcomes as critical, particularly in periods where there are very strong budgetary pressures on Government to reduce expenditure.¹⁴

3.15 The Council of Academic Public Health Institutions Australia (CAPIA) also provided a compelling account of the benefits of linking health datasets to deliver improved health policies:

The availability and accessibility of linked data collections is vital in working towards improvements in the health of Australians and in healthcare delivery. At a population level, data linkage provides a more complete understanding of health than is otherwise possible utilising alternative research methods. Providing approved researchers with access to a range of linked State and Commonwealth health and social data has the potential for national, state and local comparative effectiveness, clinical trials and registry research that has thus far been largely untapped, to drive health policy, redesign, quality improvement and evidence translation in health care. Additionally, it enables, for example, the rigorous objective evaluation of health policy for government and key policy professionals; and the ability to compare trends nationally, to identify programs that deliver value for money and to avoid wasting resources on those that are not delivering. The result is better targeted, evidence-based and more cost-

13 Department of Health, *Supplementary Submission 155*, p. 1.

14 SA-NT DataLink, *Submission 181*, p. 5.

effective health policy, services and interventions for the Australian community.¹⁵

Linking Commonwealth datasets

3.16 Health policy development at the Commonwealth level was a key area where submitters argued that significant improvements and efficiencies could be made. For instance Dr Julian Elliott, a Senior Research Fellow at the Australasian Cochrane Centre explained that the benefits of transforming existing datasets into 'evidence-informed policymaking':

So we need to have a capacity in Australia, particularly public agencies, to use the datasets that are becoming available to really drive effective decision making, evidence-informed policymaking... It is really about how we take data and then transform that into evidence-informed, up-to-date recommendations across the whole of the health sector, whether it is policymaking at a Commonwealth level or right down to the decision of an individual clinician. We have a huge opportunity to improve that cycle...¹⁶

3.17 Dr Elliott elaborated on the benefits of using big data analytics to effectively evaluate the impact of policies on health outcomes:

...it is really about how we monitor the effect that these systems then have on the outcomes that we are interested in. Is it actually changing healthcare practice? Are we getting a return on investment for our healthcare interventions? Ultimately, what effect is it having on patient outcomes? Those elements can also be collected and understood within these data systems.¹⁷

3.18 Witnesses referred the committee to an array of important Commonwealth data sources that could be beneficially linked to support evidence-based policy development as well as providing medical researchers with valuable source data. Perhaps the most comprehensive list was provided by SA-NT DataLink which suggested the following:

PBS, MBS, Immigration, Justice, Childcare Benefits, Private and Public School Education, Higher Education enrolment and academic results, Aged care, Family Tax Benefits, Employment related data from ATO Personal Income Tax and Company Tax/ABN GST and ABS.¹⁸

3.19 Of this list, two datasets, the MBS and the PBS, were virtually universally recognised by submitters as key Commonwealth data source for linkage. CAPHIA for instance submitted that:

15 Council of Academic Public Health Institutions Australia, *Submission 173*, p. 1.

16 Dr Julian Elliott, Senior Research Fellow, Australasian Cochrane Centre, *Committee Hansard*, 11 December 2015, pp 18–19.

17 Dr Julian Elliott, Senior Research Fellow, Australasian Cochrane Centre, *Committee Hansard*, 11 December 2015, pp 18–19.

18 SA-NT DataLink, *Submission 181*, p. 10.

MBS and PBS data represent two of the most important datasets in the Commonwealth repository, as they provide information on the uptake of primary care and specialist medical services, as well as use of medicines in the community, which are not available through routinely-collected State-based data collections. When combined with other data, they can provide a rich source of information to allow analysis of clinical outcomes, effectiveness of health policy, cost-effectiveness analyses and access to services across a range of dimensions, including residential location, socioeconomic status and Aboriginality.¹⁹

3.20 The current restrictions on linking MBS and PBS data were highly commented on during the inquiry. This issue is explored in greater detail in the next chapter. A related discussion regarding several Commonwealth departments' reticence to release de-identified data to other agencies is covered in Chapter 5.

3.21 Finally, it is worth mentioning a recent Commonwealth Government initiative, the Multi-Agency Data Integration Project, which is linking a series of related Commonwealth datasets. The ABS which is leading the project submitted:

A key example of these [data custodian] partnerships is the Multi-Agency Data Integration Project which brings together, for the first time, Census data with administrative data on health, income, and social security payments, to establish a foundational, linked data resource. The project aims to create an enduring integrated data resource that is:

... A comprehensive data source for evidence-based policy development across areas of broad social and economic concern...²⁰

3.22 This initiative is further discussed in Chapter 5.

Linking Commonwealth and state datasets

3.23 Due to the shared responsibility for the development of health policy in Australia, significant quantities of health data is collected at both the Commonwealth and state levels. In this regard the PHRN has recognised that:

Australia has a federated health system. The country also has high quality health data collections which can be used for planning and research. However, because of the federated system, information about a person's lifetime health journey is collected and stored in many places. For example, the States and Territories collect the birth, hospital and death data and the Commonwealth collects the childhood immunisation, Medicare Benefits Scheme (MBS), Pharmaceutical Benefits Scheme (PBS) and aged care data. In order to compare national trends and to evaluate the effectiveness of health policy for government and key policy decision makers it is necessary to be able to link this information together and use it in a timely fashion.²¹

19 Council of Academic Public Health Institutions Australia, *Submission 173*, p. 2.

20 Australian Bureau of Statistics, *Submission 192*, p. 2.

21 Population Health Research Network, *Answer to Question on Notice*, 9 October 2014, p. 1.

3.24 Professor David Preen from CAPHIA also noted how linking of Commonwealth and state health datasets can provide a sound evidence-base to government policy development:

Critically, the Commonwealth and state linked data provides for really robust, evidence based decision making in health care that can benefit not only the health system but also, ultimately, health consumers across the country. Also, we know it can be done effectively because there have been a lot of precedents over the last 10 years where people have used Commonwealth and interstate data successfully for research to address a number of issues that have directly informed government policy, at both a state and federal level.²²

3.25 The Centre for Big Data Research in Health spelt out some of the beneficial health outcomes that could flow from an evaluation based on linking Commonwealth and state datasets:

Data linkage, across national and state collections provides a platform for answering questions about access to, and outcomes of, population and individual health interventions, surveillance of disease and mortality, health system performance, policy impact and economic analysis. Put simply, it allows us to identify high-risk and low-value health services and high-risk population subgroups, and transfer this knowledge into evidence-based policies.²³

Examples of linked datasets

3.26 The committee received many examples of past, current or potential data linkage projects which strongly point to the benefits of the technique. One outstanding example of linked Commonwealth-state datasets was provided by researchers from the School of Public Health and Community Medicine at UNSW. Dr Heather Gidding and her colleagues are linking two Commonwealth datasets, the Australian Childhood Immunisation Register (ACIR) and the National Death Index, to de-identified health data from 1.8 million children across New South Wales and Western Australia to identify specific populations at risk of preventable diseases:

The ACIR alone is a significant resource, being one of only a handful of national population based immunisation registers. However, there is insufficient information on ACIR about each child to identify specific sub-populations at greatest risk of preventable diseases. Our study brings together a wealth of routinely collected information about each child to produce the first population-based estimates of effectiveness for vaccines continuing to cause morbidity in Australian children, a strategic priority area in the recently released National Immunisation Strategy. It is also the first population-based record linkage study in the world to provide estimates of vaccine impact in an indigenous sub-population, and identify

22 Prof . David Preen, Council of Academic Public Health Institutions of Australia, *Committee Hansard*, 11 December 2015, p. 38.

23 Centre for Big Data Research in Health, UNSW, *Submission 172*, pp 2–3.

specific maternal, infant and demographic characteristics of infants with delayed vaccination. We anticipate our findings will provide assurance of the effectiveness of vaccination in all Australian children and strengthen the rationale for improving vaccination timeliness, by quantifying its impact on disease burden.²⁴

3.27 The Department of Health is also leading a data linkage trial labelled the “Better targeting of mental health services” project:

The project will explore apparent disparities in provision of mental health services and assist in better targeting these services. The project is being undertaken in collaboration with the Australian Bureau of Statistics using already linked MBS and PBS data with the 2011 Census of Population and Housing data. The proposed demonstration project aims to conduct further data linkage by expanding the dataset and using more sophisticated analytical techniques to help explain the drivers of these disparities and, if appropriate, potential targets for policy intervention. A report detailing findings of the project, including both implications for mental health policy and implications for public sector data management, will be completed in early 2016.²⁵

3.28 The committee understands from the Department of Health that as at 13 April the report is yet to be finalised.

3.29 The CSIRO's submission provided a highly-practical example of a data linkage project designed to improve the efficiency of our hospital system:

Our Patient Admission Prediction Tool (PAPT) shows how the use of routinely collected administrative data can be used to make hospitals more efficient. PAPT uses a model built on historical data to forecast the number of patients to present at an emergency department within a certain time and the number that will go on to need admission to the main hospital wards. Access to reliable public health care is a key foundation to Australia's social and economic well-being. PAPT is designed to make improvements in resource allocation efficiency, reduce waiting times, and increase timely access to care and is now used in a number of Queensland Hospitals and is undergoing its first trial in a Victorian Hospital...

This important study required linking data across emergency department and hospital data sets. Sets from member hospitals of The Health Roundtable were provided, in accordance with their academic policy for use of collected data for research purposes...²⁶

24 Dr Heather Gidding, Senior Lecturer and NHMRC Early Career Research Fellow, School of Public Health and Community Medicine, University of New South Wales, *Submission 185*, p. 1.

25 Department of Health, *Supplementary Submission 155*, p. 3.

26 CSIRO, *Submission 174*, p. 5.

3.30 The CSIRO's submission went on to explain how the PAPT project could be improved through access to Commonwealth hospital datasets:

Although based on 12.5 million ED records, 11.6 million inpatient episodes and 46000 hospital deaths, the [PAPT] coverage represents only 79% of Australian tertiary hospitals and 40% of all Australian Hospitals. Improving access to hospital datasets held by the Commonwealth for quantitative analysis can avoid such limitations and ensure important policy decisions are based on the most comprehensive data available.²⁷

3.31 Finally, the Australian Commission on Safety and Quality in Health Care provide a future linked health data scenario which would enable the efficient monitoring of actual care and inform safety and quality improvement:

National guidelines for the management of acute coronary syndrome (ACS) specify that patients admitted to hospital for management of ACS be discharged on five medications...

However, studies show compliance with this guideline has been shown to vary across hospitals and hospital types highlighting this as an area for potential improvement.

Linking admitted national patient datasets...to PBS datasets using anatomic therapeutical chemical codes would allow analysis of adherence to national guidelines and variation from best practice, and provide valuable information for improving care of patients with ACS. Similar analyses could be conducted to monitor guideline compliance by healthcare facilities for a range of other conditions including recommended stroke discharge medicines.²⁸

Committee view

3.32 The committee is greatly encouraged by the strong interest expressed by government agencies, data linkage organisations and medical researchers, in the potential for improving evidence-based health policy development through data linking. There is clearly a wealth of worthy health policy proposals and evaluations that will commence once access to de-identified administrative data is more readily available. The novel insights that will flow from these projects will not only enable the development of innovative, evidence-based and more cost-effective health policy, it will also lead to better patient outcomes and improve the standard of healthcare in Australia.

3.33 The committee is however concerned by aspects of the Health Department's publicly stated approach to big data. Although the department supports the government's more open data policy, and also the recognises the significant potential

27 CSIRO, *Submission 174*, p. 5.

28 Australian Commission on Safety and Quality in Health Care, *Submission 187*, p. 2. The five ACS medications are 1. Beta blocker; 2. Aspirin; 3. Anti-platelet agents; 4. Cholesterol lowering medicines; and 5. ACS inhibitors.

of big data, it appears to be taking an intentionally slow approach to utilising data linkage in developing new health policies:

The use of Big Data technologies and analytics will be one of the focus areas in a broader activity that the Commonwealth Department of Health has just commenced to develop more comprehensive health analytic capabilities.

Initially the use of Big Data technologies will supplement the existing technology in the Department such as the Enterprise Data Warehouse (which provides for secure storage of health data for use across a range of health portfolio agencies) and the business intelligence Health capabilities. *Later stages may look at how a broader range of health data can be consolidated to develop deeper insights into the impact of health policy initiatives.*²⁹

3.34 This approach appears at best ambivalent, and at worst contrary to the government's public data policy statement which declares that 'Australian Government entities will...securely share data between Australian Government entities to improve efficiencies, and inform policy development and decision-making...'³⁰

3.35 The committee also notes the Department of Health has delayed its report into the data linkage project to better target mental health services. This further demonstrates that the department is not adequately prioritising important data linkage projects.

3.36 With Commonwealth funding of \$63 billion per annum at stake and recognising the importance of improving the health outcomes of all Australians, the committee urges the Department of Health to significantly increase its focus on data linkage.

Recommendation 2

3.37 The committee recommends that the Department of Health, as a high priority, actively explore and then implement measures to advance cost-effective, evidence-based policy development through the use of data linkage.

Recommendation 3

3.38 The committee recommends that relevant government departments should include information in their annual reports which describes the processes and projects being undertaken to establish evidence-based policy based on data linkage as well as strategies they have adopted to contribute to the government's public data policy.

29 Department of Health, *Supplementary Submission 155*, pp 1–2. Emphasis added.

30 Department of the Prime Minister and Cabinet, *Australian Government Public Data Policy Statement*, December 2015, p. 2.

Chapter 4

Restrictions on linking MBS and PBS data

4.1 The committee is ever mindful of privacy concerns with regard to data, its storage, management, use and security. However, as noted in the previous chapter, many submissions indicated that significant health policy development and medical research could be advanced if linked Medicare Benefits Schedule (MBS) and Pharmaceutical Benefits Scheme (PBS) data were more readily accessible.¹ For example Professor Sallie-Anne Pearson told the committee:

The linkage of PBS, MBS and other Commonwealth collections, such as those held by the Department of Social Services, can expand our opportunities to explore value, real-world use and pivotal issues such as equity of access. Despite medicines being tested extensively in clinical trials, when they are PBS subsidised there is significant uncertainty about how they would perform in routine clinical care. When I talk to consumers, they are surprised to learn that comprehensive postmarket surveillance research does not occur routinely in Australia. Why is this the case? Activity of this kind actually requires Commonwealth and state based data holdings to be linked... The currently fragmented data systems in Australia make it difficult, if not impossible, to systematically capture these impacts.²

4.2 Professor Pearson also noted the practical problems for researchers in not being able to link MBS and PBS data:

The agencies, I believe, are prevented from linking their MBS and PBS data, but if you actually think about it—if you want to understand something as basic as does a person go to a specialist to get a particular medicine and then how is the medicine continued, is it continued by a specialist or by a general practitioner—we cannot find that out because we do not have the visits linked to the prescription. Very basic things around navigating through that system are really actually important—are people monitored after they are prescribed a medicine? We do not know that because that information cannot be linked. There are some really practical impediments to doing some very basic work in this regard.³

1 For example NPS MedicineWise, *Submission 94*, p. 2; Northern Territory Government, *Submission 148*, p. 5; CAPHIA, *Submission 173*, p. 2; SA NT DataLink, *Submission 181*, p. 2; Public Health Information Development Unit, *Submission 183*, p. 5.

2 Prof Sallie-Anne Pearson, Head, Medicines Research Unit, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 7.

3 Prof Sallie-Anne Pearson, Head, Medicines Research Unit, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p.15.

4.3 The importance and potential uses of these datasets were also recognised in the *Public Sector Data Management Report*. According to the report, linked MBS and PBS data is the fourth most requested data from the Australian Government.⁴

Source and content of the prohibition

4.4 Presently there are restrictions in both legislation and subordinate legislation that strictly constrain the linkage of MBS and PBS data.

4.5 The *National Health Act 1953* requires the Information Commissioner to make privacy rules.⁵ The National Health Act requires that the rules must:

- (d) prohibit agencies from storing in the same database:
 - (i) information that was obtained under the Medicare Benefits Program; and
 - (ii) information that was obtained under the Pharmaceutical Benefits Program; and
- (e) prohibit linkage of:
 - (i) information that is held in a database maintained for the purposes of the Medicare Benefits Program; and
 - (ii) information that is held in a database maintained for the purposes of the Pharmaceutical Benefits Program;

unless the linkage is authorised in the way specified in the rules;⁶

4.6 The current privacy guidelines were made by the Privacy Commissioner in 2008.⁷ These legally binding guidelines provide that data from the PBS and MBS databases may only be linked:

- if it is necessary to comply with law;
- to determine eligibility for a benefit under one program, where eligibility depends upon services provided by the other program;
- where Medicare reasonably believes that doing so would prevent or lessen a serious and imminent threat to life or health; or
- for release where a person has provided their consent.⁸

4 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, December 2015, p. 16, www.dpmmc.gov.au/pmc/publication/public-sector-data-management-report (accessed 5 January 2016).

5 *National Health Act 1953*, s. 135AA(3).

6 *National Health Act 1953*, s. 135AA(5).

7 *Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programs*, Guideline 1, www.comlaw.gov.au/Details/F2008L00706 (accessed 25 January 2016).

8 *Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programs*, Guideline 3.1.

4.7 If linkage is undertaken for medical research purposes, the claims data can only be released where an individual has consented to having their data released and where the researcher undertakes to destroy the claims information provided to them at the conclusion of the research.⁹

4.8 These strict limitations came about following a plan by the then Health Insurance Commission (now Medicare) to implement an online system that would allow pharmacists to claim reimbursement and to check whether patients were eligible for concession prices on pharmaceuticals at the time they were being dispensed.¹⁰

4.9 The scheme ultimately did not go ahead, but an amendment to protect the privacy of individuals given the large amount of data that would be collected under the MBS and PBS schemes was implemented as sections 135AA and 135AB of the *National Health Act 1953*.¹¹

4.10 The provision was last examined by the Parliament in 1993. At that time the aims of the amendment were encapsulated by the then Member for Macarthur, Mr Christopher Haviland MP who argued:

There is a need to ensure that legitimate privacy principles are balanced against the public interest, particularly in relation to the possible misuse of public money. This is the essential aim of this amendment—to clarify privacy provisions to ensure that legitimate privacy concerns of individuals are protected while enabling government agencies, in this case the Health Insurance Commission, to adequately safeguard against fraud and misuse of taxpayers' money.¹²

Changes in technology

4.11 The committee remains committed to that fundamental need to "ensure that legitimate privacy concerns of individuals are protected." It also notes though that while the evolutionary march of technology has both increased the ability to collect sensitive data about individuals it has also produced technology and techniques that can protect that information.

9 *Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programs*, Guideline 6; Prof Louisa Jorm, Director, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 15.

10 Office of the Australian Information Commissioner, *Report of the Privacy Commissioner's Review of the Privacy Guidelines for the Handling of Medicare and PBS claims information*, August 2006, p. 20.

11 Office of the Australian Information Commissioner, *Report of the Privacy Commissioner's Review of the Privacy Guidelines for the Handling of Medicare and PBS claims information*, August 2006, p. 21.

12 Mr Christopher Haviland MP, Member for Macarthur, *House of Representatives Hansard*, 26 May 1993, p. 1014.

4.12 Data linkage technology has transformed significantly in the last forty years. Emeritus Professor D'Arcy Holman of the University of Western Australia told an audience in July 2014 that:

My first employment, as a Public Health Medical Student Resident during the Christmas of 1972, was to tabulate figures...of intestinal parasite infections found in the patients of Swanbourne Hospital.

The information technology available to me wasn't of the digital electronic form, but consisted of a mechanical dinosaur known as the Hollerith card sorter. The point is that what we could do with health statistics, even as recent as the 1970s, was severely constrained by the technical infrastructure available to us.¹³

4.13 Emeritus Professor Holman later returned to the subject of data linkage and privacy in his lecture with the following exposition:

...one might query if this [data linkage] represents a significant invasion of privacy. To the contrary, the effects of data linkage on privacy have been exactly the opposite, with a profound privacy benefit compared with the way we did research before.

Here's what real medical records look like, courtesy of a patient who's given permission for them to be displayed. Lots of documents, and now computer screens, liberally plastered with the patient's name and address. During the first 20 years of my career, I waded through countless thousands of records like this. It was tedious and inefficient work, especially because often one had to pour through reams of paper to find just the one or two important facts to answer the research question. Data linkage has turned this approach on its head...so that during the last 20 years, what I've worked with has looked like this: No names and addresses, age rather than date of birth, contains only the information needed to answer the research question, and just a number is used to represent each person, although for any two research projects that system is different, so the patients don't even have a unique number. Nevertheless, use of the same number for the same anonymous person in each project, illuminates the crucial connections within and between different data collections, so that the outcomes can be measured.¹⁴

4.14 In his evidence, Mr Timothy Pilgrim, who is the Acting Australian Information Commissioner and currently performs the functions of the Privacy Commissioner, reminded the committee that sometimes legislation needs to be revisited in light of technological changes:

Something that we find with a number of the laws that I deal with is that there is a need to review some of those because the situations change quite

13 E/Prof D'Arcy Holman, *Additional Information 13*, pp 2–3.

14 E/Prof D'Arcy Holman, *Additional Information 13*, pp 8–9.

dramatically in terms of technologies you can use to bring together information sets and how they can be dispersed.¹⁵

4.15 The protection of sensitive personal information remains a key focus of both researchers and governments.

Calls for review

4.16 In the 23 years since the provision was last debated, the technologies available to protect privacy have increased dramatically. As technology has increased and researchers have become able to conduct more complex analysis of combined datasets, the demand for linked data has also grown.

4.17 This in turn has led to a number of prominent reports that have recommended that the National Health Act and the Privacy Guidelines be reviewed.

4.18 For example in 2009, the National Health and Hospitals Reform Commission wrote:

To better understand people's use of health services and health outcomes across different care settings, we recommend that public and private hospital episode data should be collected nationally and linked to MBS and PBS data using a patient's Medicare card number.¹⁶

4.19 In 2013, the Productivity Commission similarly suggested that the Privacy Guidelines be amended noting that in the present environment:

Protecting confidentiality is warranted but the current approach is too cautious and complex with the restrictions creating unnecessary downsides and delays for evidence-based policy formulation.¹⁷

4.20 In December 2015, the *Public Sector Data Management Report* called the current privacy arrangements 'over-cautious and cumbersome'.¹⁸

4.21 Days after this report was released, departmental witnesses who appeared before the committee were given the opportunity to explain how the current restrictions came to be in place and why they continue to be necessary. The answer provided by representatives of the Department of Health acknowledged the 'very strong concerns about privacy' which historically dominated departmental assessments of data requests from researchers.¹⁹ However officials noted the paradigm shift that has occurred:

15 Mr Timothy Pilgrim, Acting Australian Information Commissioner, *Committee Hansard*, 2 February 2016, p. 11.

16 Quoted in Productivity Commission, *Annual Report 2012-13*, p. 7.

17 Productivity Commission, *Annual Report 2012-13*, p. 7.

18 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, December 2015, p. 18.

19 Ms Alanna Foster, First Assistant Secretary, Department of Health, *Committee Hansard*, 11 December 2015, p. 67.

...what has happened fairly recently is that there has been a significant cultural shift in the way data is regarded. It is regarded as an asset; it is regarded as a key tool in informing policy development and research. I think we are shifting from a culture of protecting data at all costs to one of protecting data but also identifying ways we can use it.²⁰

4.22 Even the Acting Australian Information Commissioner supported the need to review the current legislative restrictions on linking MBS-PBS data:

...section 135 of the Health Act...came into force, I think, over 20 years ago. I would be the first to say that legislation should be reviewed regularly and, in fact, some years ago I actually proposed that when there were some challenges identified with the guidelines when the office [Office of the Australian Information Commissioner] was developing it at the time. We acknowledged that there seemed to be some challenges about data retention; being able to bring the two sets of data together into one database was another issue. At that stage—I think that that was in 2011—I said that we were certainly open to having that particular piece of legislation looked at because it was, for want of a better description, an old piece of legislation that was developed at a different time when there were different community expectations and different mechanisms to simply store the information.

So what I am saying is yes—I think it is entirely appropriate to have that piece of legislation reviewed, to look for other mechanisms which may be able to make more efficient use of that information in terms of...freeing up data for good social policy purposes. But at the same time I would then say that if we are going to do that, what can we build in to ensure there is the right level of protection about that information in a newer environment of how it is going to be used? That could be through mechanisms such as building up protections around security, giving it stronger protections where it is going to be held. Those sorts of issues are things we would want to look at.²¹

4.23 The *Public Sector Data Management Report* released by the Department of the Prime Minister and Cabinet in December 2015 recommended that:

Legislation should be reviewed to identify whether privacy and secrecy laws can be streamlined and modernised to enable data to be better used for policy and research...²²

Australian Information Commissioner's view

4.24 The Office of the Privacy Commissioner sits within the Office of the Australian Information Commissioner. The post of Privacy Commissioner is currently

20 Ms Alanna Foster, First Assistant Secretary, Department of Health, *Committee Hansard*, 11 December 2015, p. 67.

21 Mr Timothy Pilgrim, Acting Australian Information Commissioner, Office of the Australian Information Commissioner, *Committee Hansard*, 2 February 2016, p. 12.

22 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, December 2015, p. 37.

vacant and instead those functions are fulfilled by the Acting Australian Information Commissioner.²³

4.25 The Office of the Information Commissioner recognised that there are significant social benefits that can be obtained by using de-identified health data for policy development and research purposes. The Acting Australian Information Commissioner, Mr Timothy Pilgrim, wrote in his submission:

Taking into consideration the Committee's focus on improving access to and linkage between health data sets for policy development, I appreciate that personal information held by government can be, when it is handled appropriately, a valuable resource for policy, planning, research, innovation and providing better services.

If legislative and policy changes are made to facilitate or extend access to, and the use of, personal information in research and policy planning, it is important that an integrated approach to privacy management is taken from the beginning. This includes, for example:

- implementing legislative safeguards to limit the possibility of function creep
- considering whether any restriction on an individual's right to privacy that arises from changes to how health data sets are used is reasonable, necessary and proportionate to the expected benefits
- considering whether personal information is in fact required, or whether de-identified or anonymised information will suffice
- undertaking a Privacy Impact Assessment (PIA) for each project that uses personal or de-identified information.²⁴

4.26 In his testimony, Mr Pilgrim described a 'recurring theme' amongst certain government organisations that 'the Privacy Act was blocking the use of data' instead of them actively looking for ways to comply with the legislation and achieve the goals that researchers or others might be looking for.²⁵

4.27 Mr Pilgrim also pointed out that there were options that were available under the existing privacy arrangements:

One of the security provisions in Australian privacy principle 11, which deals with securing that information and keeping it safe, says that personal information that is no longer required should either be destroyed or de-identified. It does not set a time frame around that in particular to general personal information, so one of the mechanisms there which could allow that information to continue on is if it can be de-identified. Once information is de-identified it falls out of the definition of personal

23 Office of the Australian Information Commissioner, *Who we are*, www.oaic.gov.au/about-us/who-we-are/ (accessed 5 February 2016).

24 Office of the Australian Information Commissioner, *Submission 190*, p. 2.

25 Mr Timothy Pilgrim, Acting Australian Information Commissioner, *Committee Hansard*, 2 February 2016, p. 16.

information—as you would understand, if you cannot tell who the individual is, then it is not personal information—so there are mechanisms by which quite a bit of data, I would suggest, could be kept.²⁶

Proposed privatisation of Medicare payment systems

4.28 One week after the committee's final data linkage public hearing, the *West Australian* published an article about the government's 'secret' proposal to privatise Medicare's payments system:

The *West Australian* has learnt that planning for the ambitious but politically risky outsourcing of government payments is well-advanced, with a view to making it a key feature of Treasurer Scott Morrison's first Budget in May...

[The successful private sector provider] would administer claims and payments while overseeing eligibility criteria, meaning they would require access to people's sensitive private information.

Doctors would also have to open their books to the provider, which would be subject to regulatory oversight.²⁷

4.29 Although details of the government's privatisation process and timing are unclear, senior Health Department officials have stated that a new Request for Quote was issued in January 2016 'to start to look at how we might scope this type of work, stressing that obviously we are in an exploratory stage and no decisions have been made.'²⁸

4.30 Nevertheless, the Australian Medical Association has raised concerns that any move to privatise Medicare payments could 'compromise patient privacy and further fragment their care.' AMA Vice President Dr Stephen Parnis told ABC Radio that such a move would raise serious privacy issues:

There are concerns raised about the way that the administrators of these programs would handle confidential medical data; how their input may influence or undermine the doctor-patient relationship in terms of its funding.²⁹

26 Mr Timothy Pilgrim, Acting Australian Information Commissioner, *Committee Hansard*, 2 February 2016, p. 12.

27 Mr Andrew Probyn, Federal Politics Editor, *The West Australian*, *Govt eyes massive Medicare, health privatisation*, <https://au.news.yahoo.com/thewest/wa/a/30767296/government-eyes-massive-medicare-health-privatisation/> (accessed 29 April 2016).

28 Mr Martin Bowles, Secretary, Department of Health, Community Affairs Legislation Committee Additional Estimates, *Committee Hansard*, 10 February 2016, p. 8.

29 Australian Medical Association, *Privacy Risk on Medicare Outsourcing*, <https://ama.com.au/ausmed/privacy-risk-medicare-outsourcing> (accessed 29 April 2016).

Committee view

4.31 The committee notes the real risk to privacy, improved public policy planning and to the delivery of universal healthcare if the ideological attack on Medicare expands into the actual privatisation of the Medicare payments system and associated data.

4.32 The committee also notes that the restriction on linking MBS and PBS data that is embodied in section 135AA of the National Health Act is over 20 years old and is prescriptive given technological progression in protecting data and other restrictions on accessing data.

4.33 The committee agrees with the Acting Australian Information Commissioner and other witnesses that privacy is always an important consideration in the policy making process and that it ought to be afforded serious consideration in the making and altering of access arrangements in this space. The evidence of witnesses, such as Professor Stanley and Professor Pearson, clearly indicate that there are significant harms in failing to do so.

4.34 The evidence received however, indicates that current legislative restrictions on linking MBS and PBS data are unnecessarily placing Australian lives at risk. As Professor Stanley noted (see chapters 2 and 3), there could be another thalidomide crisis or hundreds of people needlessly dying of heart attacks and we would be unable to detect it because we currently do not have the evidence-based data available.

4.35 These significant health care imperatives must be weighed against competing public policy priorities. Privacy is and must continue to be a key consideration in the formation of public policy. However, the evidence presented to this committee, drawing on the long history of data linkage both domestically and internationally, demonstrates that data linkage is undertaken securely with successful containment of risk to the privacy of individuals while leading to significant improvements in health outcomes.

4.36 While the committee is confident about the thoroughly tested processes underpinning the use of de-identified health datasets in data linkage projects, the government's proposed privatisation of the Medicare payment system raises real privacy concerns. The committee is concerned that the government's privatisation plans risk an unintended disclosure of highly sensitive MBS and PBS data. In the committee's view it is important to maintain a clear distinction between the linkage of de-identified health datasets and the wholesale privatisation the Medicare payment systems.

4.37 The committee is heartened that there seems to be such strong public support to utilise de-identified data that is already routinely collected to improve the health of the populace.

4.38 Time and again, as demonstrated above, the committee heard that consumers were surprised that government did not already use administrative data for these purposes.

4.39 There is now a renewed focus on data in the public service. Given the significant opportunities to improve Australia's healthcare outcomes, the committee urges the government to adopt the following recommendations:

Recommendation 4

4.40 The committee recommends that given the changes in technology, and mindful of the capacity and moral obligation for governments to hold and strongly secure personal data and privacy, the government review the operation of section 135AA of the *National Health Act 1953*, with the aim of improving access to de-identified MBS and PBS data for the purpose of health policy evaluation and development as well as research undertaken in the public interest.

Recommendation 5

4.41 The committee recommends that the Australian Information Commissioner, in consultation with privacy advocates, data custodians, academics and healthcare consumers, review the *Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programs* in order to ensure that the government:

- retains ownership and management of Australian MBS and PBS data and improves technological capacity to ensure the privacy of all Australians health data; and
- develops a strategy to improve access to de-identified MBS and PBS data for the purpose of health policy evaluation and development as well as research undertaken in the public interest, in ways that don't decrease privacy.

Chapter 5

Issues and barriers for researchers

Introduction

5.1 During the course of this inquiry, the committee was alerted to the many imperfections in Australia's data linkage system. This chapter examines a number of the most commonly raised issues and makes a number of recommendations aimed at progressing public health research in Australia.

5.2 This chapter examines issues that researchers and others may encounter in the process of attempting to obtain data held by government agencies. In particular this chapter will consider custodianship issues, the approvals process, linkage and the desultory departmental approach.

5.3 Australia's health data holdings are currently dispersed. The dispersed nature of the data means that the issues that follow are not a problem in all parts of Australia or are necessarily problems for all agencies. They are however, the issues that have been continuously raised with the committee.

5.4 Dr Merran Smith, Chief Executive of the PHRN conveyed the sentiment of many of the researchers, musing:

...some of the problems with the Commonwealth that we are facing at this point in time is that, in my experience over many years, the Commonwealth government agencies often do not look outside. They are busy doing what they need to do in Canberra. Perhaps they do not appreciate, or are not in a position to leverage, the value and benefit that can be obtained from working with other groups within the nation.¹

5.5 Certainly there is a view in the research community that 'departments tend to want to own that information and control it and use it for their purposes as opposed to a public purpose'.²

5.6 Ms Helen Owens, an Assistant Secretary at the Department of the Prime Minister and Cabinet, told the committee that the Prime Minister moved data policy into his department to try and improve the coordination of government data policy.³

5.7 As part of that effort, the Department of the Prime Minister and Cabinet have produced a *Public Sector Data Management report* and a *Public Data Policy*

1 Dr Merran Smith, Chief Executive, PHRN, *Committee Hansard*, 11 December 2015, p. 26.

2 Prof Brendon Kearney, Chair, PHRN, *Committee Hansard*, 11 December 2015, p. 28.

3 Ms Helen Owens, Assistant Secretary, Department of the Prime Minister and Cabinet, *Committee Hansard*, 3 February 2016, p. 13.

Statement. However, as Professor Kearney, Chair of the PHRN identified 'it is a long and slow process and a lot more needs to be done'.⁴

Custodianship issues

Accessing data

5.8 The first step in any data linkage project is to determine whether the required data is available and what permissions are required to access it. That process can be more difficult than one might anticipate.

5.9 Dr Heather Gidding, a Senior Lecturer in the School of Public Health and Community Medicine at the University of New South Wales observed that in undertaking her data linkage project:

It took us a long time to find the data custodian for that ACIR [Australian Childhood Immunisation Register] data. When we did find someone who did understand the dataset, they are very hard to access. They are overworked and hard to get a hold of. It certainly does make the research quite difficult.⁵

5.10 The frustration and uncertainty about which department holds particular data was explained by Professor John Glover, Director of the Public Health Information Development Unit (PHIDU):

One of the difficulties with the datasets we cannot get is...who owns which data, who the data custodians are. We go to the Department of Human Services and put a case to them for health data. They have a committee. Health [the Department of Health] sits on that committee and then basically decides whether or not you can have the data. The DHS people do not really know anything about the data and the policy issues; that comes from Health, and that decision is then made and passed back by email. You then complain and you make some comment, 'What about so and so, can we do this or that?', and then it goes back again. The same with PBS.

It is very hard to engage totally with whoever the custodian really is—in this case it is Health, even though you are dealing with the Department of Human Services, who would run the tables.⁶

5.11 If researchers can work out which department holds the relevant data, Dr Smith explained that there is no clear process to access data:

Part of the challenge with accessing the Commonwealth data is that there is not an explicit process. When you are accessing state data, most states have online an application form and an explanation for how you need to go about it. But even now when you go online and try to find how to access data,

4 Prof Brendon Kearney, Chair, PHRN, *Committee Hansard*, 11 December 2015, p. 28.

5 Dr Heather Gidding, Senior Lecturer and NHMRC Early Career Researcher, School of Public Health and Community Medicine, University of New South Wales, *Committee Hansard*, 11 December 2015, p. 36.

6 Professor John Glover, Director, PHIDU, *Committee Hansard*, 3 February 2016, p. 11.

there is a little bit on the AIHW site, hardly anything useful on the Department of Health's site or on the Department of Human Services' site. To acknowledge the work of my colleagues in the Commonwealth, they certainly did as a result of this process go through some conversations and streamline processes, but it is still not really explicit in the way that a researcher who says, 'Look, I think I want to use this Commonwealth data set' needs. It is not really clear how they go about the process to get the approval.⁷

5.12 Part of the problem is that Australian Government departments are often unaware of what data they hold and no list is publicly available.

5.13 This problem was highlighted when Ms Michelle Wilson, a General Manager at the Department of Human Services was unable to tell the committee what data the Department of Human Services held.⁸ However, the Department of Human Services were not alone. The committee placed questions on notice to obtain a list of the health related data sets held by several health-focussed Australian Government departments and agencies.⁹ Some departments required more than two weeks to compile their lists. The AIHW provides a list in its annual report.¹⁰ The lists provided by the departments appear in Appendix 4.

5.14 Adding to the complication facing researchers of discovering what data sets are held by each department, some departments hold data that are owned by other departments. Ms Wilson explained:

The Department of Social Services are the data custodians of [Centrelink] data, so while the Department of Human Services holds [the Centrelink] data, we are not the custodians of it. It is the same as with the Medicare MBS and PBS data that we collect in the course of our business: we are not the custodians of that data; we are the holders of it.¹¹

5.15 Insofar as the Department of Human Services holds data on behalf of other departments, it considers itself to be a 'steward' of those data. Ms Wilson explained that a data 'steward', is 'somebody who sees the data through and ensures that it is well looked after and that the guidelines are adhered to'.¹²

7 Dr Merran Smith, Chief Executive, PHRN, *Committee Hansard*, 11 December 2015, p. 30.

8 Ms Michelle Wilson, General Manager, Department of Human Services, *Committee Hansard*, 11 December 2015, p. 69.

9 Department of Health, *Answer to Question on Notice 59*, pp 2–3; Department of Social Services, *Answer to Question on Notice 55*, p. 17; Australian Bureau of Statistics, *Answer to Question on Notice 54*, p. 1.

10 AIHW, *Annual Report 2014-15*, pp 201–204.

11 Ms Michelle Wilson, General Manager, Department of Human Services, *Committee Hansard*, 11 December 2015, p. 69.

12 Ms Michelle Wilson, General Manager, Department of Human Services, *Committee Hansard*, 11 December 2015, p. 69.

5.16 If researchers manage to navigate the challenges in finding the dataset and the custodian and obtain the relevant permissions, it is possible that the department may not provide the data in the level of detail required for the project. In this regard, Professor Glover explained to the committee:

...a major gap in the data we have relates to the Medicare Benefits Schedule data and the Pharmaceutical Benefits Scheme data. We cannot get it at what I believe is an appropriate geographic level—a level that will show the major variations that occur within our cities and within our regional and rural areas, particularly in variations in health outcomes, health service youth and social determinants of health.¹³

Under-resourced

5.17 The committee heard from multiple witnesses that the data custodians they dealt with were not sufficiently resourced to ensure that the data asset could be maximised. For example Professor Sallie-Anne Pearson, Head of the Medicines Policy Research Unit, Centre for Big Data Research in Health acknowledged that:

One of the greatest challenges [custodians] face is that this is incredibly under-resourced, and designated roles and having people responsible to do this work would be a very pragmatic and significant advance in getting some of those current logjams freed.¹⁴

5.18 Professor Louisa Jorm, Director of the Centre for Big Data Research in Health echoed those comments:

I would second that. In fact, several data custodians have said to me over the years: 'Providing data for research is not core business for our agency.' It is viewed as an add-on... If the leaders of their agencies are saying, 'This is your core business and we're going to resource it,' then things will start to move.¹⁵

5.19 The committee attempted to gauge whether recent shifts in government priorities may mean that appropriate resources will now be allocated to data custodians.

5.20 Ms Foster, First Assistant Secretary with the Department of Health explained to the committee some of the evolving processes within the Department of Health:

In terms of access to data and better use of data, the Department of Health has been through a number of different reviews that talked about better use of data, better valuation, better research to inform policy development towards better health outcomes. In fact, that was the reason the division I head up was established. It was to try and make better use of data and

13 Prof John Glover, Director, PHIDU, *Committee Hansard*, 3 February 2016, p. 2.

14 Professor Sallie-Anne Pearson, Head, Medicines Policy Research Unit, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 12.

15 Professor Louisa Jorm, Director, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 12.

ensure better use of evaluation. Certainly, in my KPIs it would be a fairly key feature that there be better use of data. In the department's vision...there is that reference to better health outcomes and one of the key underpinnings of that is that better use of research and data. So it really is quite widespread throughout the department. There is a major cultural change that is occurring.¹⁶

5.21 However that statement was not reflected in the Department of Health's Answer to a Question on Notice. When asked about the KPIs the Department of Health uses to evaluate data usage the department provided the following cryptic response:

The Department facilitates a Data Governance Council that includes representation from the Department, the AIHW, the Australian Bureau of Statistics, Department of Human Services and other Health portfolio agencies. The Council is responsible for ensuring effective policies and governance for the Department's approaches to data collection, management, interrogation, sharing, access and release.¹⁷

5.22 The Australian Bureau of Statistics identified that it has a KPI entitled 'Use of Microdata Increases'.¹⁸ The AIHW has measures relating to data tied into its Portfolio Budget Statements.¹⁹ However, it appears that neither the Department of Health, the Department of Human Services or the Department of Social Services have any KPIs relating to data.²⁰

The changing data culture

5.23 Professor Pearson pointed out that data custodians play an important role in releasing health-related data:

The custodianship of the data is a very important position to be placed in. I think traditionally custodians have been very risk averse. There are different interpretations of legislation guidelines, et cetera, about the nature and the type of data that can be released. I have actually seen a massive change over the last few years where custodians are now thinking about arguments as to why data should not be released...rather than thinking, 'We have to stop this.'²¹

16 Ms Alanna Foster, First Assistant Secretary, Department of Health, *Committee Hansard*, 11 December 2015, p. 73.

17 Department of Health, *Answer to Question on Notice 61*, p. 7.

18 Australian Bureau of Statistics, *Answer to Question on Notice 53*, p. 6.

19 AIHW, *Answer to Question on Notice 56*, p. 9; AIHW, *Annual Report 2014-15*, www.aihw.gov.au/aihw-annual-report-2014-15/ch1/ (accessed 9 February 2016).

20 Department of Social Services, *Answer to Question on Notice 55*, p. 15; Department of Human Services, *Answer to Question on Notice 63*, p. 14.

21 Professor Sallie-Anne Pearson, Head, Medicines Policy Research Unit, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, pp 9–10.

5.24 Professor Pearson also noted:

There clearly is an enormous responsibility on the part of custodians to protect the identity of individuals, but the challenges that they face relate to the interpretation of legislation, which can be different with different individuals. The other great challenge that they face is that oftentimes this role is done on top of a highly powered, high-level and highly committed role. It is really about goodwill and trying to serve the research community.²²

5.25 Apparently a new culture, with regard to the use and release of administrative data held by Australian Government departments, is in its infancy. While change is occurring, the speed at which parts of the public service are adapting varies. An example of the more cautious approach was exemplified by Ms Foster of the Department of Health:

...the whole senior executive [are] helping to drive that cultural change to make data available.

But having said that, while we are moving from that attempt to make data as available as possible to researchers and for policy development purposes there is that very key issue of privacy that we are concerned about. I think that for many years there was that culture, 'We must absolutely protect this data at all costs.' But, of course, as techniques—computing and statistical techniques and all sorts of things—get more sophisticated there are more ways to 'perturb' the data, as I think you referred to it, or to confidentialise the data so we can actually protect people's privacy and still be able to make information available for use by researchers.²³

5.26 Ms Foster continued:

There are very strict guidelines under the National Health Act, the Health Insurance Act, the privacy guidelines and the Privacy Act. We also observe those provisions very strictly. Indeed, sometimes those rules can limit our own potential to use data internally. But we do have to be very conscious of those privacy rules to ensure that people's information cannot be identified.²⁴

5.27 Ms Foster's answer can be starkly contrasted with the dynamic response of Mr David Dennis, a Branch Manager with the Department of Social Services who submitted:

We have endeavoured, especially in recent times, to be at the vanguard of testing technologies and new techniques and developing new techniques to

22 Professor Sallie-Anne Pearson, Head, Medicines Policy Research Unit, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 12.

23 Ms Alanna Foster, First Assistant Secretary, Department of Health, *Committee Hansard*, 3 February 2016, p. 23.

24 Ms Alanna Foster, First Assistant Secretary, Department of Health, *Committee Hansard*, 3 February 2016, p. 23.

allow more data to be released to intelligent sources... Internally, we have worked with DHS to streamline our approval processes and, in so doing, have devolved a lot of routine approvals to DHS directly so that they no longer need to seek recourse to DSS where routine approvals are involved, so it moves a lot more quickly.

Where we have been successful in this regard is by not treating every data application equally. In so doing, we have published to data.gov.au large amounts of information that is now open, albeit in a perturbed or confidentialised manner. That has allowed us to turn our minds to what we see as the current challenge...of arranging for researchers to have access to individual unit record data in a safe and appropriate environment, because this is what they need to perform the statistical analyses that they want to perform. We have trialled a number of models to facilitate that...using innovative products...such as the Secure Unified Research Environment... so that finally we can put an enduring dataset of DSS data in a safe place that credentialled individuals can access directly at the unit record level and where they can withdraw the products of those analyses, and we are assured that the data is safe and the researchers get the access that they need...²⁵

5.28 The Australian Bureau of Statistics also appears to be embracing the new data culture. The Australian Bureau of Statistics now operates a number of linked data projects including the Mental Health Services-Census Integrated Dataset.

5.29 The Mental Health Services-Census Integrated Dataset is a good example of linking diverse data sets to provide a solid evidence base. It combines 2011 Census data with mental health related items from the MBS and PBS along with demographic data. The Australian Bureau of Statistics is currently investigating adding state and territory hospital data to this dataset to enable 'a more complete picture of the patterns of service usage'.²⁶

Approvals processes

5.30 Various approvals to access data and conduct research are necessary to ensure that the data being released are used for the public benefit and that appropriate measures are being taken to ensure the privacy and security of personal information.

Ethics approvals

5.31 To ensure that health information is used appropriately, delegated legislation provides that researchers must undertake ethics approval if they wish to use health data.²⁷

25 Mr David Dennis, Branch Manager, Department of Social Services, *Committee Hansard*, 3 February 2016, p. 25.

26 Australian Bureau of Statistics, *Submission 192*, p. 4.

27 See *Privacy Act 1988*, ss 16B, 95, 95A and *Guidelines under Section 95 of the Privacy Act 1988*.

5.32 Ethics approvals play an important part in upholding high ethical standards in research. However, as Dr Gidding told the committee, having to satisfy multiple approvals can add substantial time to a research project:

I am leading a team of researchers in New South Wales and Western Australia that predominantly have linked up the Australian Childhood Immunisation Register, the vaccination records of Australian children to records for children in both WA and New South Wales—1.8 million children... We had to get nine approvals for our project to go ahead, which also included six ethics applications that took two years.²⁸

5.33 The Australian Commission on Safety and Quality in Health Care (ACSQHC) endorsed the idea that multiple ethics approvals hindered data linkage:

One of the main issues is...[the] complicated ethics approvals systems in the various jurisdictions. In Australia, projects using linked data must be approved by the data linkage unit; the data custodian responsible for each data set; and one or more Human Research Ethics Committee(s)... In order for data on safety and quality to be most effective in driving improvement, regular and timely review is required.²⁹

5.34 Dr Nicky Antonius, Acting Assistant Secretary of the Department of Health and Mr Warren Richter, Head, Chief Information Officer Group, AIHW both told the committee they would not request an additional ethics approval where the researcher had already obtained an approval that met the research ethics standards of the National Health and Medical Research Council (NHMRC). However, ethics approvals from other jurisdictions were more problematic. This continues to be a problem even within Australia where Commonwealth departments will not accept the ethics approval of a State, such as Victoria.³⁰

5.35 In his evidence to the committee, the Acting Australian Information Commissioner Mr Timothy Pilgrim questioned whether ethics approvals were necessarily the best mode of protecting privacy:

I have heard researchers remark regularly about some of the challenges around ethics committees, and there may be a good argument for looking at other mechanisms.³¹

28 Dr Heather Gidding, Senior Lecturer and Early Career Researcher, University of New South Wales, *Committee Hansard*, 11 December 2015, p. 36.

29 ACSQHC, *Submission 187*, pp 2–3.

30 Dr Nicky Antonius, Acting Assistant Secretary, Department of Health and Mr Warren Richter, Head, Chief Information Officer Group, AIHW, *Committee Hansard*, 11 December 2015, p. 72.

31 Mr Timothy Pilgrim, A/g Australian Information Commissioner, *Committee Hansard*, 2 February 2016, p. 11.

Delays

5.36 Once a researcher has obtained ethics approval, the committee received evidence that there have often been significant delays in obtaining data from departments.

5.37 In 2009 Professor Fiona Stanley and Emeritus Professor D'Arcy Holman wrote a letter to the then Secretary of the Department of Health and Ageing, Ms Jane Halton. The letter notified the Secretary:

Of the 23 research projects [that had applied for Commonwealth held health data], only nine have received useable Commonwealth data. All nine have experienced lengthy delays in obtaining data, some as much as two years from the time of application. Those still to receive data have been waiting for as long as three years. It would be fair to say that the affected researchers have found it extremely difficult to engage with the Australian Department of Health and Ageing in progressing a solution to these delays and solving attendant issues such as data queries and explanations.³²

5.38 Similarly when Professor David Preen, Deputy Head of the School of Population Health at the University of Western Australia, was asked about the source of delays in cross-jurisdictional linkage, he responded:

The approval processes through the Commonwealth were slow... [In] a number of cases we had queries or had made submissions but did not hear back for a period of time, and we would follow up. But it was not clear to whom we should be following up with... But, there being delays, what we could do to facilitate that and what else may needed to provide in terms of information, et cetera, was very unclear. That, in itself, caused further delays. Whether that was due to inaction or a lack of transparency with the existing processes, I guess, is debatable. But the end result is a huge delay in provision of these data.³³

5.39 It is important to note that there is a large financial cost to researchers from these delays. In particular, the lengthy delays have led to a significant wastage of public research funds.³⁴

5.40 The degree of wastage was articulated by Emeritus Professor D'Arcy Holman who estimated that in 2008:

...over 100 medical researchers waiting for [Commonwealth] data, had \$11 million in precious public research funds, mostly from NHMRC, simply going to waste.³⁵

32 Prof Fiona Stanley and E/Prof D'Arcy Holman, *Tabled Document 98*, p. 2.

33 Prof David Preen, Deputy Head, School of Population Health, University of Western Australia, *Committee Hansard*, 11 December 2015, p. 41.

34 Telethon Kids Institute, *Data Linkage Roundtable: Summary and actions*, p. 1, <http://telethonkids.org.au/media/1033902/Data-Linkage-Roundtable-Report.pdf> (accessed 12 February 2016).

5.41 The problem of significant delays in obtaining data is not limited to researchers. It also affects government agencies. In its most recent annual report the AIHW explained the impact of the delays in producing its reports:

This was because we did not gain access to records from the Medical Benefit Schedule and the Pharmaceutical Benefits Schedule as expected, which meant that the committee [the AIHW Ethics Committee] did not receive applications from external researchers for ethical clearance of projects involving use of these particular data sets.³⁶

Linkage

Linkage logjam

5.42 The committee heard that one of the greatest impediments to the widespread use of linked data is the lack of facilities accredited to link Commonwealth data.

5.43 Professor Louisa Jorm described the current situation:

...the bottleneck is our limited capacity for data linkage. Currently only the Australian Bureau of Statistics and the Australian Institute of Health and Welfare are accredited as integrating authorities for Commonwealth health data, and yet we have state-of-the-art health data linkage facilities operating in states and territories, including New South Wales, South Australia and the Northern Territory, and Western Australia. National capacity for data linkage would be dramatically boosted if these state and territory based units were accredited to link Commonwealth as well as jurisdictional data. In particular, it makes a lot of sense for these units to service projects that involve linkage of data from a single jurisdiction with Commonwealth data, rather than these projects adding to the linkage logjam and linkage queue that currently exists at the Australian Institute of Health and Welfare.³⁷

5.44 Similarly, the Centre for Big Data Research in Health strongly advocated for State and Territory data linkage units to be accredited by the Commonwealth:

Australia has world-renowned health data linkage facilities yet relatively few of these are accredited to link and integrate Commonwealth data collections. Health data linkage activities have increased substantially over the last decade and this activity will continue to grow in the coming years. The Commonwealth data linkage facilities based at the ABS and AIHW are not necessarily equipped to manage this growth and are not always an efficient option for projects involving linkage of Commonwealth data with

35 Emeritus Professor D'Arcy Holman, *Additional Information 13*, p. 14.

36 AIHW, *Annual Report 2014-15*, p. 6.

37 Professor Louisa Jorm, Director, Centre for Big Data Research in Health, *Committee Hansard*, 11 December 2015, p. 7. The Australian Institute of Family Studies is also an integrated authority but is yet to complete a data linkage project.

collections from one or more States. One solution is to let others with proven track records do this too.³⁸

5.45 According to the National Statistics Service, the only Integrating Authorities accredited to link Commonwealth health data are the ABS, the AIHW and the Australian Institute of Family Studies.³⁹

5.46 Dr Smith advised the committee that attempts to have State data linkage units accredited have been unsuccessful:

One of the state linkagers has sought accreditation to be able to receive Commonwealth data and the view from the Commonwealth was that they could not accredit it because it was a state agency, so they did not have jurisdiction.⁴⁰

Cost of linkage for researchers

5.47 The committee heard evidence that data linkage can be a costly barrier to the development of new knowledge and innovation.

5.48 Some Australian Government departments charge researchers on a cost recovery basis to access the data for linkage projects.⁴¹ In some cases, the amount charged is cost-prohibitive. SA-NT Datalink provided two examples:

For one project with a cohort of about 10,000 individuals and linking 4 datasets, the SA-NT DataLink cost was estimated at \$10,000. The researchers also wished to link to a Commonwealth dataset for which they quoted approximately \$160,000.

Because of the high Commonwealth costs, the researchers could not include this data.

Another project with a cohort of about 240 individuals and linking 4 datasets, the SA-NT DataLink cost was estimated at \$8,500. The researchers also wished to link to another Commonwealth dataset for which they quoted approximately \$40,000.

Again, because of the high Commonwealth data costs, the researchers could not include this data.⁴²

38 Centre for Big Data Research in Health, *Submission 172*, p. 3.

39 National Statistics Service, *Accredited Integrating Authorities*, www.nss.gov.au/nss/home.nsf/pages/Data%20Integration:%20Accredited%20Integrating%20Authorities (accessed 14 January 2016). Integrating Authorities are currently accredited by the Cross Portfolio Data Integration Oversight Board.

40 Dr Merran Smith, Chief Executive, PHRN, *Committee Hansard*, 11 December 2015, p. 32.

41 See for example AIHW, *Answer to Question on Notice 58*, p. 7; Department of Human Services, *Answer to Question on Notice 63*, pp 11–12.

42 SA-NT DataLink, *Answer to Question on Notice 65*, 2 February 2016, p. 1.

5.49 Then linkage units can also charge researchers depending on the work performed. Officials at the Australian Institute of Health and Welfare told the committee that they operate on a cost-recovery basis:

As you may be aware, the Australian Institute of Health and Welfare receives about 30 per cent of its funding from appropriations, so 70 per cent of our revenue comes from the provision of goods and services to others. We run our data-integrating authority and data-linkage services on a cost-recovery basis. That is, essentially, the cost of a salary plus the overheads associated with running buildings... There are other organisations involved in the chain. If you have a state-linkage organisation, a Centre for Health Record Linkage in New South Wales, they are also running on a cost-recovery basis, so the researchers have to pay the costs that are incurred to deliver the service they are asking for.⁴³

5.50 When the committee asked NACCHO [the National Aboriginal Community Controlled Health Organisation] what it cost to link data through the AIHW Dr Robert Starling replied 'it is in the tens of thousands of dollars—it is not 50 bucks here or there.'⁴⁴

5.51 In response to a question on notice, the AIHW advised that the average cost of a linkage project in 2015 was \$4414 inclusive of GST.⁴⁵

5.52 The ABS notified the committee that the average fee of obtaining Custom Data Information was \$1970 but that average cost of using Information Consultancies carried out by the Health subject matter area was \$4645 inclusive of GST.⁴⁶

5.53 Researchers may also be charged to use laboratories such as the SURE research facility to analyse data. Mr Wells, the Deputy Chief Executive Officer of The Sax Institute informed the committee:

On the researcher side, we do charge researchers. But that charge does not reflect the full cost, because we have a subsidy through NCRIS funding. So we charge them a charge which is not the full cost of providing it. But, certainly, it is a very costly facility to run. I am not saying that is a bad thing, but it is costly. We are constantly looking at enhancing its efficiency et cetera, but there is very much a cost to operating and providing this service.⁴⁷

43 Mr Warren Richter, Chief Information Officer, AIHW, *Committee Hansard*, 11 December 2015, p. 70.

44 Dr Robert Starling, Chief Information Officer, NACCHO, *Committee Hansard*, 11 December 2015, p. 56.

45 AIHW, *Answer to Question on Notice 51*, p. 2.

46 ABS, *Answer to Question on Notice 53*, p. 5. ABS have an arrangement with Australian universities that all but the most complex work is provided to researchers free of charge. Complex work is still charged on a cost recovery basis.

47 Mr Robert Wells, Deputy Chief Executive, The Sax Institute, *Committee Hansard*, 11 December 2015, pp 26–27.

5.54 One strategy that has the potential to bring down the cost of linkage is the creation of enduring linked datasets.

Enduring data linkage

5.55 At the end of a linkage project there is a legal requirement that researchers must destroy the links between the datasets as Mr Crettenden, Assistant Secretary, Department of Health explained:

The Privacy Commissioner's guidelines...say that Medicare...and the [Department of Health] are only able to link the datasets under very specific circumstances and that the dataset that is used to create the linkage must be destroyed within one month of it being created.⁴⁸

5.56 However that creates problems for ongoing research according to Dr Smith:

If you go to a lot of effort to create a high-quality linkage and then, after the project is finished, you destroy the links and start doing the same thing again for the next project, first of all, it is highly inefficient; it is both expensive and time consuming. Second, it does not lead to improvements in linkage quality. When we talk about the importance of enduring linkage, that is why it is important.⁴⁹

5.57 The benefits of being able to maintain a linked dataset is that the data can be accessed very quickly and it can become a permanent resource that researchers can use.⁵⁰

5.58 Australia has created enduring datasets before as Professor Preen recalled:

It was a formalised and accepted process for linkages de-identified unconsented or non-consented Medicare, PBS and state data. In WA a memorandum of understanding was developed and signed by the Commonwealth and by the state government to set up what was referred to as the cross-jurisdictional linkage facility to specifically link the projects after approval. These data would be linked in unconsented form but they would be provided in de-identified form for researchers. My understanding is that 25 projects went through, were approved and received data through that process. That system was functional, and the fact that the MOU was agreed to and signed means that it should be a possibility, obviously, within the existing legal framework. That has now been replaced because of some of the issues with the layers, with integrating authorities.⁵¹

48 Mr Ian Crettenden, Assistant Secretary, Department of Health, *Committee Hansard*, 11 December 2015, p. 67.

49 Dr Merran Smith, Chief Executive, PHRN, *Committee Hansard*, 11 December 2015, p. 23.

50 Mr Warren Richter, Group Head, Chief Information Officer Group, AIHW, *Committee Hansard*, 11 December 2015, p. 66.

51 Prof David Preen, Deputy Head of the School of Population Health, University of Western Australia, *Committee Hansard*, 11 December 2015, pp 42–43.

5.59 Departmental officials indicated that they were currently undertaking an ambitious Multi-agency Data Integration Project (MADIP). The MADIP aims to combine 'information from existing surveys, administrative collections and censuses' to develop 'a more complete picture of the circumstances of individuals, households and businesses'.⁵² The MADIP 'brings together, for the first time, Census data with administrative data on health, income, and social security payments'.⁵³

5.60 Unlike the 'link and destroy' model described above, MADIP is going to be an enduring linkage.⁵⁴

5.61 Promisingly, there is a commitment amongst several key government agencies to move towards greater use of enduring data sets. Ms Gemma Van Halderen, the General Manager of the Strategic Partnerships and Projects Division at the Australian Bureau of Statistics told the committee:

...the Commonwealth secretaries and some dep secretaries across the Australian Bureau of Statistics, Department of Human Services, Department of Social Services and the Department of Health have been working...since about 2009 on some Commonwealth arrangements for data sharing and data use. The secretary commissioned a review of those arrangements in 2014. That review was undertaken by a consultant, Ms Susan Linacre, who used to be a deputy at the Australian Bureau of Statistics. She recommended to the deputy secretaries and the secretaries that the Commonwealth should move away from a link-and-destroy model to a create-and-reuse-and-keep model, so we are currently moving towards that... There is a very strong appetite within the Commonwealth agencies and the members of that data integration initiative to put in place a research create-and-reuse-and-keep model, not a create-and-destroy model for data linking.⁵⁵

The desultory departmental approach

5.62 As the above evidence reveals, there is no consistent policy between Australian Government departments in their approach to releasing data to researchers.

5.63 Part of the reason for the haphazard approach may be attributed to the lack of coordination of government policy in this area. Ms Owens, Assistant Secretary, Department of the Prime Minister and Cabinet informed the committee that her department is responsible for the coordination of this area of policy:

Our role is really the coordination of the policy element. We would still rely upon our colleagues here at the table to do the actual coordination within

52 Australian Bureau of Statistics, *Submission 192*, p. 2.

53 Australian Bureau of Statistics, *Submission 192*, p. 2.

54 Dr Nicky Antonius, Acting Assistant Secretary, Department of Health, *Committee Hansard*, 3 February 2016, p. 17.

55 Ms Gemma Van Halderen, General Manager, Strategic Partnerships and Projects Division, Australian Bureau of Statistics, *Committee Hansard*, 3 February 2016, p. 22.

their own thematic areas. We are not actually going to do the doing in this space...⁵⁶

5.64 The problems identified in this report highlight the failures of some departments to adapt to a new, more flexible approach to releasing de-identified data. Those failures must be addressed and rectified before Australia can make the most of its big data potential.

Release of information between government departments

5.65 The complexity of the Australian Government's data arrangements were reinforced when Senator Moore tried to ascertain what data was held by the Department of Human Services:

Senator MOORE: I am interested to know whether the human services data includes only Medicare data or whether it has Centrelink and DVA [Department of Veterans' Affairs] data as well...

Ms Wilson: ...The Department of Social Services are the data custodians of that data, so while the Department of Human Services holds that data, we are not the custodians of it. It is the same as with the Medicare MBS and PBS data that we collect in the course of our business: we are not the custodians of that data; we are the holders of it.

Senator MOORE: Are you custodians of anything?

Ms Wilson: We are custodians of the data about how our customers behave in our service delivery systems.⁵⁷

5.66 Dr Nicky Antonius, Acting Assistant Secretary at the Department of Health also tried to assist the committee explaining:

If DHS [the Department of Human Services] was asked to divulge data to AIHW [Australian Institute of Health and Welfare] that would depend what dataset and who is the custodian of that dataset. For MBS, PBS and Medicare enrolments data it would be the Department of Health but we are the ones who will pursue the public interest certificates and then share that with the Department of Human Services to forward it to the requester. If the requester is asking for social services data then they would be directed to the Department of Social Services.⁵⁸

5.67 Mr Warren Richter of the AIHW and Ms Michelle Wilson from the Department of Human Services told the committee that the process of releasing data between Australian Government department could be very involved:

56 Ms Helen Owens, Assistant Secretary, Department of the Prime Minister and Cabinet, *Committee Hansard*, 3 February 2016, p. 14.

57 Ms Michelle Wilson, General Manager, Department of Human Services, *Committee Hansard*, 11 December 2015, pp 68–69.

58 Dr Nicky Antonius, A/g Assistant Secretary, Department of Health, *Committee Hansard*, 11 December 2015, p. 69.

Mr Richter: We have a schedule agreement that was signed last week with the department [Department of Human Services]. We have a public interest certificate which is in the process of being signed within the department and we have arrangements in place with the Department of Human Services to receive the Medicare enrolments data, which we need to do the linkage—to identify the links. As soon as that public interest certificate has been signed—I do not know if you know, Michelle, but we have your people all on tap to hit the button—

Ms Wilson: We are looking at the piece of equipment that is holding the data to make sure it is ready to go. We will wait until the public interest certificate has been tied up with a bow and signed by everybody. It is going through the final processes, as is the multi-agency data integration project. The public interest certificate is also at that final stage of processing. Each of our departments has to go through its own legal processes.⁵⁹

5.68 That somewhat confusing state of affairs was compounded by Ms Wilson who went on to argue:

We [the Department of Human Services] relied on a public interest certificate and a collaborative process to look at the privacy and secrecy issues with the Department of Health. Indeed, with the Department of Social Services data that relates to social security information we also have a collaborative process with the Department of Social Services. So it is generally not just up to the Department of Human Services about how the data is used. In fact we cannot really make decisions ourselves about how that data is used.⁶⁰

5.69 The *Public Sector Data Management Report* provides additional examples of current inefficiencies:

In one case where an important dataset was being linked with the Census, the negotiations on the MOU [Memorandum of Understanding] took up to 18 months while the linking only took two weeks.⁶¹

5.70 And:

An agency reported having up to 11 MOUs with the same department to access data.⁶²

5.71 Accordingly, the Public Sector Data Management Report concluded:

59 Mr Warren Richter, Chief Information Officer, AIHW and Ms Michelle Wilson, General Manager, Department of Human Services, *Committee Hansard*, 11 December 2015, p. 68.

60 Ms Michelle Wilson, General Manager, Department of Human Services, *Committee Hansard*, 3 February 2016, p. 24.

61 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, December 2015, p. 18, www.dpmc.gov.au/pmc/publication/public-sector-data-management-report (accessed 5 January 2016).

62 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, December 2015, p. 18.

Overall, the lack of data sharing prevents feedback on policy and hinders the potential of data to improve future service delivery.⁶³

Committee view

5.72 The committee is deeply concerned by the lack of transparency regarding the data held by government departments. If Australia is to maximise the value of its data resources, researchers need to know what data sets are available and how to access them.

5.73 During the course of this inquiry it became obvious that some departments were uncertain about what datasets they held. In some cases, when asked to provide a list to the committee on notice some departments required two weeks to list their data holdings.

5.74 It ought to be clear to both researchers and departments what data each department collects and what area within the department is responsible for data custodianship. The *Public Data Policy Statement*, the *Public Sector Data Management Report* and a chorus of witnesses all agree that non-sensitive data should be publicly available as a right and that sensitive data should be able to be accessed subject to appropriate privacy and security constraints. For this purpose, a list of data sets held by each department ought to be publicly available on their websites and on data.gov.au. Departments also ought to publish a clear statement that identifies exactly how requests for data are made and how those requests will be dealt with.

Recommendation 6

5.75 The committee recommends that each Australian Government agency develop and maintain on its website a list of datasets held by the agency along with the contact details of the data custodian. This list should be updated at least twice annually.

Recommendation 7

5.76 The committee recommends that all datasets held by the Commonwealth be listed on www.data.gov.au, identifying which agency is the data custodian.

Recommendation 8

5.77 The committee recommends that each Australian Government agency that is a data custodian develop and publish on its website guidance for researchers detailing its process for data requests and approvals.

63 Department of the Prime Minister and Cabinet, *Public Sector Data Management Report*, December 2015, p. 19.

5.78 In certain departments the current system of approving access to data is at times chaotic. The evidence demonstrates that seeking multiple ethics approvals was time consuming without necessarily being of significant additional benefit.

5.79 The committee is of the view that there ought to be a consistent method of applying for approval to access sensitive data. The committee notes that ethics approvals of other states are currently not recognised by Australian Government departments. Such approvals ought to be nationally recognised to ensure consistent and systematic application of process. Therefore, the relevant departments need to work at both the Commonwealth level and with the States and Territories.

Recommendation 9

5.80 The committee recommends that the government take a whole-of-government approach to streamlining the ethics approval process and the authorising environment in consultation with the Privacy Commissioner, privacy advocates, the NHMRC, data custodians, academics, consumers and the States and Territories. The government should also work with the States and Territories to establish a national accreditation system so that ethics approvals from accredited jurisdictions are recognised by the Commonwealth.

5.81 The evidence received by the committee demonstrates that data custodians are currently not being adequately resourced. Furthermore, it was clear from the lack of KPIs in most government departments that the responsibilities of data custodianship is not a priority for departments despite the considerable benefits that would flow from a sharper focus.

5.82 The burdens placed on custodians have undoubtedly made them more hesitant to release data. The relevant departments need to empower the officials in these positions and give them the resources and confidence to be able to release datasets where researchers have obtained the approval that demonstrates their ability to maintain the security and privacy of the data.

5.83 The committee was shocked to discover that precious Commonwealth funded grants for important medical research projects were going to waste because Australian Government departments would not provide the data required. The committee believes that the \$11 million figure cited by Emeritus Professor Holman is highly conservative.

5.84 The committee notes that the Productivity Commission and others have urged government to make de-identified datasets more readily available to researchers. Where data has been de-identified, open access ought to be the default position.

5.85 It is absolutely unacceptable that researchers should have to wait years to access data to facilitate important research projects in the public interest. This is a situation that requires immediate attention.

5.86 Therefore the committee makes the following recommendations aimed at prioritising data access and encouraging appropriate open access.

Recommendation 10

5.87 The committee recommends that relevant government agencies give greater priority to, and adequately resource, their data custodians.

Recommendation 11

5.88 The committee recommends that relevant government agencies provide guidance to data custodians to assist them in their decision-making, with a view to making more de-identified data available on an enduring basis.

Recommendation 12

5.89 The committee recommends that the government adopt the Productivity Commission's proposed principle that open access to de-identified datasets should be the default position.

Recommendation 13

5.90 The committee recommends that the government should direct relevant agencies to release de-identified datasets on an enduring basis as the default position.

Recommendation 14

5.91 The committee recommends that departments that have data custodianship responsibilities must establish and publish realistic Key Performance Indicators for the timely consideration and approval of datasets requests. These departments must publicly report on their KPIs in their annual reports.

If after 5 years departments continue to delay the release of datasets, then the committee recommends that the government establish binding timeframes for processing applications for data. Failure to comply with the timeframe should trigger appeal rights similar to those found in other information access regimes.

5.92 Currently, the release of various datasets between Australian Government departments and agencies at times appears shambolic. The evidence provided to the committee indicates that several departments, despite the imprimatur of the new data policy, were very protective of their data and demonstrated great reticence to release it to other government agencies. Whilst the committee understands and supports the need to protect privacy concerns and to act in accordance with legislation, the current practices are inefficient and stifle innovation.

Recommendation 15

5.93 The committee recommends that Government encourage collaboration on data linkage projects between government agencies, as well as academia and industry to provide for evidence-based policy development and facilitate research that is undertaken in the public interest.

5.94 The committee is confident that increasing the availability of linked data will make Australia a more attractive research destination that produces significant public health research and better evidence-based policy proposals.

5.95 To ensure that researchers can be supported into the future, consideration ought to be given to the linkage units. The committee heard that some units were almost already at capacity and that there was at the very least the potential for a 'linkage logjam'.

5.96 Australia has a number of excellent linkage units at the state level, some of whom would like to be accredited to link Australian Government data. The government should give the prospect serious consideration.

5.97 The committee also notes that depending on the nature of the project some researchers may be charged multiple times (potentially at the Commonwealth and state levels) for accessing various services in order to facilitate that research. To ensure that the medical research sector is as vibrant as it can be, the government ought to consider the cumulative costs incurred by researchers across the linkage project.

Recommendation 16

5.98 The committee recommends that government consider accrediting State data linkage units to link Commonwealth data with State data collections, subject to comprehensive privacy and security protocols.

Recommendation 17

5.99 The committee recommends that the Government review the cost of data access and linkage work undertaken by Commonwealth entities with a view to facilitating research and innovation in the national interest.

5.100 Australia has the capacity to be a world-leader in this field. For the benefit of the health of all Australians the committee strongly urges the government to implement all recommendations to ensure that Australia realises its big data potential.

Senator Deborah O'Neill

Chair

Appendix 1

Witnesses who appeared before the committee¹

The committee held three hearings focusing specifically on big data on 11 December 2015 and on 2 and 3 February 2016.

Friday, 11 December 2015 – Sydney

University of Sydney

Dr Barbara Mintzes, Senior Lecturer, Faculty of Pharmacy

Roundtable

Centre for Big Data Research in Health, University of New South Wales

Professor Louisa Jorm, Director

Professor Sallie-Anne Pearson, Head, Medicines Policy Research Unit

Dr Julian Elliott, Senior Research Fellow, Australasian Cochrane Centre

National Health Performance Authority (NHPA)

Dr Diane Watson, Chief Executive Officer

Population Health Research Network (PHRN)

Dr Merran Smith, Chief Executive

Professor Brendon Kearney, Chair

Centre for Data Linkage

Associate Professor James Boyd, Director

Associate Professor Anna Ferrante, Deputy Director

¹ The hearings and witnesses listed in this appendix relate to the committee's hearing on 11 December 2015 and 2–3 February 2016 which focussed on big data, data linkage and related issues. A full list of the committee's hearings and witnesses is at the committee's website:
www.apf.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Public_Hearings.

The Sax Institute

Professor Sally Redman AO, Chief Executive Officer

Mr Robert Wells, Deputy Chief Executive Officer

Roundtable**Council of Public Health Institutions Australia (CAPHIA)**

Professor David Preen, Director, Centre for Health Services Research, University of Western Australia

Australian e-Health Research Centre

Dr David Hansen, Chief Executive Officer

Australian Health Economics Society

Professor Philip Clarke, Professor of Health Economics, University of Melbourne

University of New South Wales

Dr Heather Gidding, Senior Lecturer and NHMRC Early Career Research Fellow, School of Public Health and Community Medicine

National Rural Health Alliance

Mr Gordon Gregory, Chief Executive Officer

Mr Andrew Phillips, Policy Adviser

Ms Fiona Brooke, Policy Adviser

National Aboriginal Community Controlled Health Organisation

Dr Robert Starling, Chief Information Officer

Royal Australian College of General Practitioners (RACGP)

Dr Nathan Pinski, Chair of the RACGP Expert Committee – eHealth and Practice Systems

Government panel**Department of Health**

Ms Alanna Foster, First Assistant Secretary, Strategic Policy and Innovation Division

Mr Ian Crettenden, Assistant Secretary, Strategic Policy and Innovation Division

Dr Nicky Antonius, Acting Assistant Secretary, Information Knowledge Management Branch

Australian Institute of Health and Welfare (AIHW)

Mr Warren Richter, Chief Information Officer

Dr Nick von Sanden, Unit Head, Statistical and Analytical Support Unit

Department of Human Services

Ms Michelle Wilson, General Manager, Strategic Information Division

Tuesday, 2 February 2016 – Canberra

Roundtable**Northern Territory Government**

Dr Steven Guthridge, Director, Health Gains Planning

SA-NT DataLink

Mr Andrew Stanley, Director

Mr Christopher Radbone, Associate Director

Office of the Australian Information Commissioner

Mr Timothy Pilgrim PSM, Acting Australian Information Commissioner

Ms Angelene Falk, Assistant Commissioner, Regulation and Strategy

University of Western Australia School of Population Health and Telethon Kids Institute

Ms Anne McKenzie AM, Consumer Advocate and Program Manager

Telethon Kids Institute

Professor Fiona Stanley AC, Patron and former director

Wednesday, 3 February 2016 – Canberra

Roundtable**National eHealth Transition Authority**

Mr Peter Fleming, Chief Executive Officer

Ms Bettina McMahon, Head of Risk and Assurance

Australian Commission on Safety and Quality in Health Care

Dr Robert Herkes, Clinical Director

Ms Catherine Katz, Director, Safety and Quality Improvement Systems and Inter-governmental Relations

Public Health Information Development Unit

Professor John Glover, Director

Roundtable**Department of the Prime Minister and Cabinet**

Ms Helen Owens, Assistant Secretary

Mr Tim Neal, Senior Advisor

Department of Health

Ms Alanna Foster, First Assistant Secretary

Mr Ian Crettenden, Assistant Secretary

Dr Nicky Antonius, Acting Assistant Secretary

Department of Social Services

Mr David Dennis, Branch Manager, Policy Evidence Branch

Department of Human Services

Ms Michelle Wilson, General Manager, Strategic Information Division

Australian Bureau of Statistics

Ms Gemma Van Halderen, General Manager, Strategic Partnerships and Projects Division

Australian Institute of Health and Welfare

Mr Geoff Neideck, Group Head, Chief Information Officer Group

Dr Nick von Sanden, Unit Head, Statistical and Analytical Support Unit

Appendix 2

Submissions received by the committee in relation to data linkage¹

- 1 National Health Performance Authority (supplementary submission)
- 43 Australian Health and Hospitals Association (supplementary submission)
- 69 Queensland Government (supplementary submission)
- 94 NPS MedicineWise (supplementary submission)
- 119 Aids Council of New South Wales (supplementary submission)
- 148 Northern Territory Government (supplementary submission)
- 155 Department of Health (supplementary submission)
- 167 Royal Australasian College of Surgeons (supplementary submission)
- 171 Centre for Data Linkage
- 172 Centre for Big Data Research in Health
- 173 Council of Academic Public Health Institutions Australia (CAPHIA)
- 174 Commonwealth Scientific and Industrial Research Organisation (CSIRO)
- 175 National E-Health Transition Authority (NEHTA)
- 176 The Sax Institute
- 177 Australian Institute of Health and Welfare (AIHW)
- 178 Family Medicine Research Centre, University of Sydney
- 179 Dr Julian Elliott
- 180 Australian Longitudinal Study on Women's Health
- 181 SA-NT DataLink
- 182 Research Australia
- 183 Public Health Information Development Unit (PHIDU)
- 184 Australian Health Economics Society

1 The submissions listed in this appendix relate to the committee's inquiry into big data, data linkage and related matters. A full list of submissions received by the committee is available on the committee's website:
www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Submissions.

- 185 Dr Heather Gidding
- 186 National Rural Health Alliance
- 187 Australian Commission for Safety and Quality in Healthcare (ACSQHC)
- 188 Population Health Research Network (PHRN)
- 189 University of Western Australia and Telethons Kids Institute
- 190 Office of the Australian Information Commissioner (OAIC)
- 192 Australian Bureau of Statistics (ABS)

Appendix 3

Additional information and answers to questions on notice¹

Tabled Documents

- No. 89 Tabled by Dr Barbara Mintzes, Senior Lecturer, Faculty of Pharmacy at a public hearing in Sydney on 11 December 2015
- No. 90 Tabled by Professor Louisa Jorm, Director, Centre for Big Data Research in Health, UNSW at a public hearing in Sydney on 11 December 2015
- No. 91 Tabled by Professor Sally Redman AO, Chief Executive Officer, The Sax Institute at a public hearing in Sydney on 11 December 2015 – SURE: safe sharing of sensitive data
- No. 92 Tabled by Professor Philip Clarke, Professor of Health Economics, University of Melbourne at a public hearing in Sydney on 11 December 2015 – Long term disability associated with war-related experience among Vietnam veterans
- No. 93 Tabled by Mr Gordon Gregory, Chief Executive Officer, National Rural Health Alliance at a public hearing in Sydney on 11 December 2015
- No. 94 Tabled by Professor John Glover, Director, Public Health Information Development Unit at a public hearing in Canberra on 3 February 2016 – SA3 Level data – Table
- No. 95 Tabled by SA-NT DataLink at a public hearing in Canberra on 2 February 2016
- No. 96 Provided by Ms Anne McKenzie AM, Consumer Advocate and Program Manager, University of Western Australia and Telethon Kids Institute and tabled by Chair on 3 February 2016
- No. 97 Provided by Ms Anne McKenzie AM, Consumer Advocate and Program Manager, University of Western Australia and Telethon Kids Institute and tabled by Chair on 3 February 2016

1 The documents listed in this appendix relate to the committee's inquiry into the proposed privatisation of Australian Hearing and related matters. A full list of documents is available at the committee's website:
www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Additional_Documents.

- No. 98 Provided by Ms Anne McKenzie AM, Consumer Advocate and Program Manager, University of Western Australia and Telethon Kids Institute and tabled by Chair on 3 February 2016
- No. 99 Provided by Ms Anne McKenzie AM, Consumer Advocate and Program Manager, University of Western Australia and Telethon Kids Institute and tabled by Chair on 3 February 2016
- No. 100 Tabled by Ms Bettina McMahon, Head of Risk and Assurance, National eHealth Transition Authority at a public hearing in Canberra on 3 February 2016 – Opening Statement

Answers to Questions on Notice

- No. 52 Answer to question on notice – public hearing 11 December 2015, Sydney – Australian Institute of Health and Welfare
- No. 53 Answer to question on notice – public hearing 11 December 2015, Sydney – Dr Julian Elliott
- No. 54 Answers to questions on notice – public hearing 11 December 2015, Sydney – Australian Bureau of Statistics
- No. 55 Answers to questions on notice – public hearing 11 December 2015, Sydney – Australian Bureau of Statistics – Summary of data collections
- No. 56 Answers to written questions on notice – public hearing 11 December 2015, Sydney – Department of Social Services
- No. 57 Answers to questions on notice – public hearing 11 December 2015, Sydney – Department of Health
- No. 58 Answers to written questions on notice – public hearing 11 December 2015, Sydney – Australian Institute of Health and Welfare
- No. 59 Answers to questions on notice – public hearing 11 December 2015, Sydney – Department of Health
- No. 60 Answers to questions on notice – public hearing 11 December 2015, Sydney – Department of Health
- No. 61 Answers to questions on notice – public hearing 11 December 2015, Sydney – Australian Institute of Family Studies
- No. 62 Answers to questions on notice – public hearing 11 December 2015, Sydney – Department of Health
- No. 63 Answers to questions on notice – public hearing 11 December 2015, Sydney – Department of Human Services
- No. 64 Answers to questions on notice – public hearing 2 February 2016, Canberra – Office of the Australian Information Commissioner

- No. 65 Answers to questions on notice – public hearing 2 February 2016, Canberra – Mr Andrew Stanley, Director, SA-NT DataLink
- No. 66 Answers to questions on notice – public hearing 3 February 2016, Canberra – Department of the Prime Minister and Cabinet
- No. 67 Answers to questions on notice – public hearing 3 February 2016, Canberra - Department of Health

Additional Information

- No. 13 Valedictory lecture by Emeritus Professor D'Arcy Holman, formerly of the University of Western Australia

Correspondence

- No. 5 Letter from Mr Timothy Pilgrim PSM, Acting Australian Information Commissioner, clarifying evidence provided at the public hearing on 2 February 2016 in Canberra

Appendix 4

Data collections held by departments

Answer:

Table 1 lists the major data collections held by the Department of Health. The term 'major data collections' has been defined as containing greater than one million lines of data and updated annually, or at a higher frequency, and potentially available for research purposes.

Table 1: Major Datasets held by Health

Ageing and Aged Care Division	
Branch Managing the collection	Data Collection
Aged Care Policy Branch	Residential Care data
	Home Care data
	Home and Community Care Minimum Data Set
Home Support Branch	Commonwealth Home Support Programme Data Set
My Aged Care Operations Branch	Ageing and Aged Care Data Warehouse
Health Workforce Division	
Branch Managing the collection	Data Collection
Rural Access Branch	Medicare Provider Number Approval System
Health Training Branch	Australian General Practice Training (AGPT) Programme data

Medical Benefits Division	
Branch Managing the collection	Data Collection
Medicare Financing and Listing Branch	Medical Benefits Schedule
Office of Hearing Services	Voucher Scheme clients and data
Pharmaceutical Benefits Division	
Branch Managing the collection	Data Collection
Pharmaceutical Policy Branch	Pharmaceutical Benefits Scheme
Research, Data and Evaluation Division	
Branch Managing the collection	Data Collection
Health Analytics Branch	Private Hospital Data Bureau (PHDB)
	Hospital Casemix Protocol 1 (HCP1)
	General Treatment Dental (GTD)
Office of Health Protection Division	
Branch Managing the collection	Data Collection
Health Protection Policy Branch	National Notifiable Diseases Surveillance System (NNDSS)
Immunisation Branch	Australian Childhood Immunisation Register
	National Human Papillomavirus Vaccination Register

**ANSWERS TO QUESTIONS ON NOTICE SOCIAL SERVICES PORTFOLIO
FOLLOWING PUBLIC HEARING - SYDNEY 11 DECEMBER 2015**

Question No: 15

Topic: Access to Australian Government Data Hansard page: Written

Senator Question Submitted by: Senator O'Neill, Chair, Senate Select Committee on Health

Please provide a list of the major data collections held by the department arranged by the area responsible for managing or updating the collection in a manner similar to Appendix 7 of AIHW Annual Report.

Answer:

No	Group responsible for handling the collection	Data collection / asset
1	Policy Office – Social Security Stream	Basic Data Set (BDS)
2	Social Security Performance and Analysis – Social Security Stream	Bluebook Dataset
3	Social Security Policy – Social Security Stream	Carers Data Set (CDS)
4	Social Security Policy – Social Security Stream	Cyril – Pensions Dataset
5	Social Security Policy – Social Security Stream	Household, Income and Labour Dynamics in Australia Survey (HILDA)
6	Social Security Policy – Social Security Stream	The Longitudinal Study of Australian Children (LSAC)
7	Social Security Policy – Social Security Stream	The Longitudinal Study of Indigenous Children (LSIC)
8	Social Security Policy – Social Security Stream	Journeys Home
9	Social Security Policy – Social Security Stream	Building a New Life in Australia (BNLA): The Longitudinal Study of Humanitarian Migrants
10	Multicultural Settlement Services and Communities – Families and Communities Stream	Humanitarian Entrants Management Systems (HEMS)
11	Multicultural Settlement Services and Communities – Families and Communities Stream	Settlement Database – Settlement Reporting Facility (SDB)
12	Housing Homelessness and Assurance– Disabilities and Housing Stream	Housing Data Set (HDS)
13	Housing Homelessness and Assurance– Disabilities and Housing Stream	Journeys Home: Longitudinal Study of Factors Affecting Housing Stability
14	Social Security Policy – Social Security Stream	JASON (A longitudinal episodic collection of datasets derived from data from the Department of Human Services)

Appendix 7

Data collections

This appendix details data collections managed by the AIHW at 30 June 2015.

Group and Unit managing the collection	Data collection
Chief Information Officer Group	
Data Linkage Unit	National Death Index
Community Services and Communication Group	
Child Welfare and Prisoner Health Unit	Adoptions Child Protection Juvenile Justice Prisoner Health Intensive Family Support Services (Child Protection) Survey on the views of children and young people in out-of-home care
Disability and Ageing Unit	Disability Services National Minimum Data Set Home And Community Care Minimum Data Set Hospital Dementia Services Survey National Aged Care Data Clearinghouse Younger people with disability in residential aged care
Health Group	
Cancer and Screening Unit	Australian Cancer Database BreastScreen Australia Database National Bowel Cancer Screening Dataset National Cervical Cancer Screening Database Cervical Screening (Safety Monitoring) Dataset
Cardiovascular, Diabetes and Kidney Unit	National (insulin-treated) Diabetes Register
Population Health and Primary Care Unit	Adult Vaccination Survey data collections Australian Infant Feeding Survey 2010 National Mortality Database Bettering the Evaluation and Care of Health (BEACH) Survey (before 1 July 2011) Risk Factor Prevalence Surveys

continued

Group and Unit managing the collection	Data collection
Health Group	
	Selected veterans and defence health databases and nominal rolls
	National Health Survey
	National Physical Activity Surveys
	National Survey of Lead in Children 1995
Hospitals, Resourcing and Classifications Group	
Group as a whole	Injury presentations to selected hospital emergency departments
	National Coronial Information System data
Health Performance Indicators Unit	Australian Spinal Cord Injury Register
	Radiotherapy waiting times
Hospitals Data Unit	National Hospital Morbidity Database
	National Public Hospitals Establishments Database
	National Elective Surgery Waiting Times data collections
	National Non-admitted Patient Emergency Department Care Database
	National Outpatient Care Database
	National Non-admitted Patient Care Database
	National Emergency Access Target Database
	National Elective Surgery Target Database
	Hand Hygiene Collection
	State and territory infection surveillance data collection
	Hospital Utilisation and Costs Study
	Medical Indemnity National Collection
Expenditure and Workforce Unit	Health and Welfare Expenditure Database
	Aboriginal and Torres Strait Islander health expenditure database
	Health Labour Force Collections
	Medical Schools Outcomes Database
	Dental: Health Surveys
	Expenditure Output data collection

continued

Group and Unit managing the collection	Data collection
Housing and Specialised Services Group	
Housing and Homelessness Collection Operations Unit	Specialist Homelessness Establishment Database
Housing and Homelessness Collection Processing Unit	Specialist Homelessness Services
	Public Rental Housing
	Mainstream Community Housing data collection
	Indigenous Community Housing data collection
	Australian Government Housing data
	Private Rental Assistance data collection
	Home Purchase Assistance data collection
	National Social Housing Survey
	Supported Accommodation Assistance Program
	Commonwealth Rent Assistance Survey 1998
	ABS Household Data Collection
	Commonwealth–State Housing Agreement Collection
	Community Housing Mapping Data Collection 1998
	Community Services Commission data collection
	High and Complex Needs Survey
	Youth Homelessness Pilot Program Collection
Mental Health and Palliative Care Unit	National Community Mental Health Care Database
	National Mental Health Establishments Database
	ABS Survey of Mental Health and Wellbeing 1997
	National Community Mental Health Establishments Database
	National Residential Mental Health Care Database
	National Survey of Mental Health Services
Tobacco, Alcohol and Other Drugs Unit	Alcohol and Other Drug Treatment Services
	National Opioid Pharmacotherapy Statistics Annual Data
	National Drug Strategy Household Surveys
Statistical Advisor	Census of Population and Housing Sample File 2011

continued

Group and Unit managing the collection	Data collection
Indigenous and Children's Group	
Group as a whole	National Perinatal Data Collection
Indigenous Analyses and Reporting Unit	Child Health Check (CHC) data collection National Aboriginal and Torres Strait Islander Survey 1994
Indigenous Community and Health Services Reporting Unit	Online Services Report data collection Indigenous Primary Healthcare National Key Performance Indicators Closing the Gap Clearinghouse
Maternal Health, Children, Youth and Families Unit	Footprints in Time—the Longitudinal Study of Indigenous Children Growing Up in Australia—the Longitudinal Study of Australian Children Household, Income and Labour Dynamics in Australia survey

Australian Bureau of Statistics

Area	Data collection/survey
Health and Disability Section	<ul style="list-style-type: none"> • National Health Survey (NHS) • Australian Health Survey (AHS) • National Health Measures Survey (NHMS) • National Nutrition and Physical Activity Survey (NNPAS) • National Nutrition Survey 1995 (NNS) • Patient Experience Survey (PEX) • Survey of Disability, Ageing and Carers (SDAC) • Apparent Consumption of Alcohol • Apparent Consumption of Foodstuffs • Coordination of Health Care Survey (CHC) [Data collection underway with results expected to be available from early 2017].
Mental Health Section	<ul style="list-style-type: none"> • National Survey of Mental Health and Wellbeing (SMHWB) • Mental Health Services – Census Integrated Dataset
National Centre for Aboriginal and Torres Strait Islander Statistics (NCATSIS)	<ul style="list-style-type: none"> • Australian Aboriginal and Torres Strait Islander Health Survey (AATSIS) • National Aboriginal and Torres Strait Islander Health Survey (NATSIS) • National Aboriginal and Torres Strait Islander Nutrition and Physical Activity Survey (NATSINPAS) • National Aboriginal and Torres Strait Islander Health Measures Survey (NATSIHMS) • National Aboriginal and Torres Strait Islander Social Survey (NATSISS)
Health and Vitals Statistics Unit	<ul style="list-style-type: none"> • Causes of Death • Private Hospital Establishments Collection (PHEC)
Demography Section	<ul style="list-style-type: none"> • Life tables, States Territories and Australia • Life Tables for Aboriginal and Torres Strait Islander Australians
Living Conditions Section	<ul style="list-style-type: none"> • General Social Survey (GSS)

Senate Select Committee on Health**PUBLIC HEARING – 11 DECEMBER 2015
ANSWER TO QUESTION ON NOTICE****Department of Human Services****Topic:** Big Data and Data Linkage**Question reference number:** DHS – Question 15**Senator:** O'Neill**Type of question:** Written**Date set by the committee for the return of answer:** 26 January 2016**Number of pages:** 2**Question:**

Can you provide a list of the major data collections held by the department arranged by the area responsible for managing or updating the collection?

Answer:

The Department of Human Services (the Department) delivers a range of government and other payments and services to almost every Australian.

As a consequence of the wide range of interactions the Department has with the Australian public, large amounts of information are collected and held in the Department's systems. The data is not held as a formal collection.

The Department's interactions include:

- Centrelink payments and services for retirees, the unemployed, families, carers, parents, students, people with disabilities, Indigenous Australians, people from culturally and linguistically diverse backgrounds, people living overseas and provision of services at times of major change, including disaster recovery payments.
- Aged care payments to services funded under the Aged Care Act 1997 including residential care, home care and flexible care services, and assessment of means and assets.

- Medicare services and payments that support the health of Australians such as Medicare, the Pharmaceutical Benefits Scheme, eHealth, Private Health Insurance Rebate, the Australian Childhood Immunisation Register, the National Bowel Cancer Screening Register and the Australian Organ Donor Register.
- Child Support services for separated parents to provide the financial and emotional support necessary for their children's wellbeing.

The Department also delivers other services including Tasmanian Freight Equalisation Scheme and Early Release of Superannuation, as well as whole of government services such as myGov.

The Department, in consultation with the policy departments, publishes data on public facing websites. Currently, the Department provides the following datasets on data.gov.au:

- Centrelink Programme Demographic data;
- Medicare Programme information including the Pharmaceutical Benefits Scheme and the Australian Organ Donor Register; and
- Annual Report.