Australian Government

Department of Health

HealthConsult

Better thinking. Better advice.

Department of Health

Development of a Framework for Secondary Use of My Health Record Data

Public Consultation Paper

HealthConsult Pty Ltd
ACN 1 18 337 821

Sydney Office: 3/86 Liverpool Street, Sydney, New South Wales, 2000 Phone (02) 9261 3707

Melbourne Office: 429/838 Collins Street, Docklands, Victoria, 3008 Phone (03) 9081 1640

1 September 2017

Table of Contents

Section Page

List of Abbreviations

ABC	Australian Broadcasting Corporation
ABS	Australian Bureau of Statistics
ACD	Australian Cancer Database
ACIR	Australian Childhood Immunisation Register
ACT	Australian Capital Territory
AIFS	Australian Institute of Family Studies
AIHW	Australian Institute of Health and Welfare
AODR	Australian Organ Donor Register
APPS	Australian Privacy Principles
CPSIC	Cross Portfolio Statistical Integration Committee
DHS	Department of Human Services (Commonwealth)
DQF	Data Quality Framework
EAP	Expert Advisory Panels (part of HealthConsult Consortium)
EHR	Electronic Health Record
EMR	Electronic Medical Record
EREC	External Request Evaluation Committee
GPRN	General Practice Research Network
HIN	Health Insurance Number
HMO	Health Management Organizations
HPF	High Power Field
HREC	Human Research Ethics Committee
ICES	Institute for Clinical and Evaluative Sciences
KP	Kaiser Permanente
NCRIS	National Collaborative Research Infrastructure Strategy
NEAF	National Ethics Application Form
NHHRC	National Health and Hospital Reform Commission
NHMRC	National Health and Medical Research Council
NHS	National Health Service
NPS	National Prescribing Service
NSW	New South Wales
OAIC	Office of the Australian Information Commissioner
OECD	Organisation for Economic Co-operation and Development
PAH	Potentially Avoidable Hospitalisations
PBS	Pharmaceutical Benefits Scheme
PCEHR	Personally Controlled Electronic Health Record
PCIS	Primary care information system
PHIPA	Personal Health Information Protection Act
PHRN	Population Health Research Network
PIN	Personal Identity Number
SA	South Australia
SLK	Statistical Linkage Keys
SURE	Secure Unified Research Environment
SUS	Secondary User Service
UK	United Kingdom
US	United States
YLL	Years of life lost

Glossary of Terms

- **Anonymisation of data:** Anonymisation refers to the conversion of identifiable data into data that cannot be used to identify an individual whether from that data itself, or from that data and other information to which the organisation has or is likely to have access to. This process may lead to the creation of 'de-identified' data.
- **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 optimisation'.[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.
- **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 producing source data, in addition to their role as a holder of datasets.'[2]
- **Data linking:** means 'the bringing together of two or more data sets to create a new, richer data set.'[3] 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 healthcare system. This has the potential to inform government policy making and decisions about improving service delivery.[3]
- **De-identified data:** involves a process by which information such as identifiers, names, addresses, gender, date of birth or other identifying information are removed from datasets entirely, or coded or encrypted. The aim of de-identification is to obscure the identifiable data items within a persons' records sufficiently that the risk of potential identification of the subject or a persons' record is minimised to acceptable levels.
- **Data re-identification:** refers to any process by which anonymised data is matched to its true owner after it has been released in de-identified form.
- **Identifiable data:** is data that can be used on its own or with other information to identify, contact, or locate a single person, or to identify an individual in context.
- **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. The AIHW, ABS and AIFS are the only three currently accredited Integrating Authorities in Australia.
- **Non-identified data:** It is worth noting that in Australia, the National Statement on Ethical Conduct in Human Research suggests the term: 'non-identified' data (due to inconsistent definitions of de-identification)[4]. Identifying information including unique numbers, name, address gender, date of birth or other identifying information can be removed from datasets and then coded, encrypted, or masked (e.g. changing data values or aggregation).[5]
- **Pseudo-anonymised data**: is a procedure by which the most identifying fields within a data record are replaced by one or more artificial identifiers, or pseudonyms. There can be a single pseudonym for a collection of replaced fields or a pseudonym per replaced field. When pseudo-anonymisation techniques are consistently applied, the same pseudonym is provided for individual patients across different data sets and over time. This allows for the linking of data sets and other information.
- **Statistical Linkage Keys:** A key that enables two or more records belonging to the same individual to be brought together. It is represented by a code consisting of the 2nd, 3rd and 5th characters of a

person's family name, the 2nd and 3rd letters of the persons' given name, the day, month and year when the person was born and the sex of the person, concatenated in that order.[6]

1 Introduction

HealthConsult, as leader of a consortium consisting of two other firms and eight subject matter experts, was engaged by the Department of Health (the 'Department') to:

"develop a Framework for the secondary use of data held in the My Health Record system for research, policy, system use, quality improvement and evaluation activities"

For the purposes of this project 'secondary use' is defined as using the information in the My Health Record system for purposes other than the provision of direct healthcare to the individual person, which is considered to be the primary use. The use of data solely for commercial and non-health related purposes is considered out of scope.

This Chapter briefly describes the My Health Record system, sets the context and identifies the scope of the project, the project objectives and the purpose of this Public Consultation Paper.

1.1 THE MY HEALTH RECORD SYSTEM

The My Health Record system (previously known as the Personally Controlled Electronic Health Record (PCEHR) system) was implemented in 2012, in response to recommendations made by the National Health and Hospital Reform Commission (NHHRC)[7]. The NHHRC considered that the introduction of a PCEHR for each Australian was one of the most important systemic opportunities to improve the quality and safety of healthcare, reduce waste and inefficiency, and improve health outcomes for individuals. The PCEHR was aimed at giving consumers better access to their own health information, promoting consumer participation, and supporting self-management and informed decision-making.

The system, as initially implemented, supports consumer control of record content and access through opting-in and nominating health practitioners and information[8]. Optional privacy features are available to consumers, giving them choice about what information is in their record and how that information is shared. These features include the capability to remove documents they do not want to share with healthcare provider organisations, set a record access code (a code they can give to healthcare provider organisations to allow them to access their record), restrict particular documents and control who can view those documents, and receive an email or SMS when a new healthcare provider organisation or one of their nominated representatives accesses their record.

Subject to these consumer controls, the My Health Record system allows healthcare provider organisations to access from anywhere important summary health information about individuals like allergies, medical conditions and treatments, medicine details, test or scan reports in digital form in one place. Aside from information entered directly by a consumer, the system contains either a copy or a summary of health information held by healthcare provider organisations. The original information is retained within the system of the healthcare provider that delivered the service.

In response to the 2013 PCEHR Review[9], the Government committed to further developing the system to improve its usability and clinical utility, strengthen eHealth governance and operations, and trial new participation arrangements. As a result, in January 2016, the PCEHR was renamed My Health Record, two trials of 'opt out' participation arrangements were initiated in the Nepean Blue Mountains and

Northern Queensland geographic areas, and the Australian Digital Health Agency (the 'Agency') was established to manage governance, operation and ongoing delivery for digital health from 1 July 2016'.[10]

During 2018, My Health Record participation arrangements will switch to opt out meaning a My Health Record will be created for every Australian, unless they choose to opt out. This will result in a significant increase in My Health Record participation and a rich dataset to support research and public policy making.

1.2 PURPOSE OF THE PROJECT

As indicated above, currently data collected through the My Health Record system is not permitted for any secondary use. At this time, the system and the data stored within it are only used for the purposes of providing healthcare. Individuals who have a My Health Record can set their personal access controls, and healthcare provider organisations may update and/or access each record accordingly.

Under the *My Health Records Act 2012* (the Act), health information in the My Health Record system may be collected, used and disclosed "for any purpose" with the consent of the healthcare recipient. In addition, one of the functions of the System Operator (the Agency) is "to prepare and provide de-identified data for research and public health purposes." Before the provisions of the Act will be implemented, a Framework for the secondary use of My Health Record system data (hereinafter referred to as the 'Framework') must be established. The role of HealthConsult is to develop a draft Framework and an associated draft Implementation Plan for consideration by the Department.

A key task in developing the Framework is to design and conduct a consultation process to facilitate a public conversation about the possible secondary uses of My Health Record data in the future. Submissions in response to this consultation paper, as well as the program of consultation activities which will occur concurrent to its release, will inform the development of processes and controls for the secondary use of data held in the My Health Record system. The Department seeks to understand the views of consumers and healthcare provider organisations about the grounds upon which it would be acceptable to allow the secondary use of My Health Record data. These views will be faithfully reflected in the Framework.

1.3 PROJECT CONTEXT

The questions that arise when considering the secondary use of health data are the same questions that arise in any attempt to achieve the public good. Secondary use of health data has the potential to enhance future healthcare experiences for patients by enabling the expansion of knowledge about disease and appropriate treatments, strengthening the understanding about effectiveness and efficiency of service delivery, supporting public health and security goals, and assisting providers in meeting consumer needs.[11] It allows a range of organisations to conduct research and innovate to improve health and healthcare outcomes, which can in turn improve well-being, productivity and the efficient use of resources.

The development of a Framework will assist government and individuals to protect privacy and data security, and to build a culture of trust and integrity in the secondary use of the data for community benefit. Accordingly, in developing the Framework, it is important that the relevant ethical, political, privacy, technical, and social issues are understood and appropriately addressed.[12] While not new[11], these issues will be given careful consideration in the context of the ever expanding volume of health data; the need to improve data access; and to provide coherent policies and standards of evidence-based best practice to support implementation of secondary use arrangements. This project will build on national and international learnings to ensure that the Framework reflects the views of stakeholders, addresses the issues and is evidence-based.

The Framework is being developed in an environment of growing demand from government, research and non-government organisations to use existing health data resources. Although, it is at a relatively early stage in its evolution, it is recognised that the My Health Record system could become one of Australia's most comprehensive health data resources. In this context, the Framework will establish the guidance for the future potential use of data held in the My Health Record system for purposes that may include policy analysis, health services program development, research, quality and safety measurement, public health, performance management and to develop and improve healthcare services and treatments. It is not the intention to use My Health Record system data to determine remuneration or appropriate rebate claiming patterns for healthcare providers.

The secondary use framework will be developed in the context of the broader Australian Government policy direction on big data. In a December 2015 policy statement the Prime Minister, the Honourable Malcom Turnbull MP, indicated that Commonwealth Government entities are committed to:

- specific actions designed to optimise the use and reuse of public data
- release non-sensitive data as open by default
- collaborate with the private and research sectors to extend the value of public data for the benefit of the Australian public.

To that end the Framework will be consistent with the work of the Data Integration Partnership of Australia, which has been tasked to transform the analysis of public data to improve policy and program implementation and expenditure.

While it is noted that the Framework will be developed ahead of any Australian Government action in relation to the Productivity Commissions' Data Availability and Use Inquiry, the public consultation process is in place to ensure transparency consistent with following key elements of reform it recommends:

- a legislative framework designed to provide a clear and modernised approach to data access and use
- new rights for consumers to enable them to share in the benefits of data — these demand-focused reforms would drive better choices and more competitive markets – establishing a scalable, risk-based institutional framework to allow integration and sharing or release of Australia's data
- recognising that some datasets are of such significance they should be treated as national assets.

1.4 PROJECT SCOPE

While the conceptual parameters of developing a Framework for secondary use of health data are broad, this public conversation must be limited by the explicit requirements of the existing legislation. Given the *My Health Record Act 2012* (the Act) stipulates that de-identified data from the My Health Record system may only be used "for research and public health purposes", the use of data solely for commercial and non-health related purposes is considered out of scope.

In keeping with the overarching focus of enabling secondary use of health data for the public good, the Framework will build upon existing privacy legislation and procedures. For example, the Framework will prevent the provision of data for secondary use by private health insurers without the explicit consent of individual/s concerned. It is envisaged that the Framework will address overlap between commercial and health related uses. For example, use of data for development of pharmaceuticals could be considered both a health related and commercial purpose. Similarly, the data may also be relevant for decision

support tools for healthcare provider organisations which could be developed by private industry as a commercial enterprise – but also fulfil an important health purpose which is in the public interest.

The Framework is also expected to enable the Agency, as the System Operator, to adhere to the proven security protocols and procedures that have been established in a sensible manner. For example, while this paper includes examples of data governance practices from a variety of institutions, the Framework is being developed with a view to defining a role for a single accountable authority for the management of My Health Record data for secondary use to minimise the risks associated with security breaches such as re-identification of data.

PURPOSE OF THIS DOCUMENT

Release of this Public Consultation Paper marks the start of a public conversation about the possible secondary use of My Health Record data. The paper outlines the context for the project, provides information on practices elsewhere in respect of the key areas to be covered in the Framework, and seeks to canvass the issues in the My Health Record context, so that they can be addressed in the drafting of the Framework. It also presents a number of case studies describing potential beneficial secondary uses of health data developed by members of the Consortium's Expert Advisory Panels (EAPs), which draw on relevant national and international experiences.

Stakeholders are invited to make a written submission against the questions in this Public Consultation paper and/or to participate in any other aspect of the consultation process. Full details of how to participate are at www.myhealthrecorddata.healthconsult.com.au.

2
About My Health Record data

This Chapter provides a summary of the My Health Record data. The reason My Health Record was developed was so that an individual's health information can be accessed through one system by authorised healthcare provider organisations and the individual anywhere at any time.

Table 2.1 describes the data that can currently be held in the My Health Record system.

Table 2.1: Summary of information that can currently be held in the My Health Record system, August 2016

Information	Authored by	Viewable by*	Type of information
Discharge Summary	Hospital Clinician	Authorised employees of a registered healthcare provider	A record of an individual's hospital stay and any follow up treatment required. The data may include discharge medications, relevant pathology results, presenting problem and diagnoses, treatments provided and recommendations on future care.
Event Summary	Any clinician	Authorised employees of a registered healthcare provider	Key information relating to one or more episodes of care such as a visit to an allied health professional, general practitioner, etc. that is relevant to future care.
Shared Health Summary	Medical practitioners, registered nurse or certain Indigenous health workers	Authorised employees of a registered healthcare provider	A clinical document summarising an individual's health status that includes important information such as allergies/adverse reactions, medicines, medical history and immunisations.
Specialist Letter	Medical specialist	Authorised employees of a registered healthcare provider	Report on a course of care provided by a medical specialist that is sent to the referring clinician. It may include diagnoses, treatments, relevant test results, recommendations for future care, etc.
Referral	Any clinician	Authorised employees of a registered healthcare provider	A request for advice or care from another clinician. It may include a summary of previous health and health care issues such as history, presenting problem, diagnoses, relevant test results, etc.
Prescription Record	Prescribing Clinician	Authorised employees of a registered healthcare provider	Information about medicines prescribed.
Dispense Record	Pharmacist - Community or hospital based	Authorised employees of a registered healthcare provider	Information on medicines dispensed.
Pathology Report	Pathologist	Authorised employees of a registered healthcare provider	Pathology test results
Diagnostic Imaging Report	Radiologist	Authorised employees of a registered healthcare provider	Report on the outcome of a diagnostic imaging examination

Information	Authored by	Viewable by*	Type of information
Advance Care Document	Consumer	Authorised employees of a registered healthcare provider	A written statement regarding a person's wishes for their future medical or health care treatment, which may formally appoint a substitute decision-maker.
Advance Care Document Custodian	Consumer	Authorised employees of a registered healthcare provider	Identity and contact details of the custodian of a person's advance care plan.
Personal Health Note	Consumer	Consumer only	Anything the consumer wishes to record.
Personal Health Summary	Consumer	Authorised employees of a registered healthcare provider	Consumer entered medicines, allergies and adverse reactions.
Personal Health Achievement	Consumer	Consumer only	Child development milestones.
Child Development Information**	Consumer	Authorised employees of a registered healthcare provider	Child development information.
Personal Health Observations**	Consumer	Authorised employees of a registered healthcare provider	Records of childhood development measurements such as height and weight.
Pharmaceutical Benefits Report (PBS)	DHS	Authorised employees of a registered healthcare provider	PBS subsidised dispense information.
Medicare/DVA Benefits Report (MBS)	DHS	Authorised employees of a registered healthcare provider	Medicare subsidised event information.
Australian Immunisation Register (ACIR)	Clinicians vis DHS	Authorised employees of a registered healthcare provider	Immunisation information.
Australian Organ Donor Register (AODR)	Consumer via DHS	Authorised employees of a registered healthcare provider	Australian Organ Donor status.

* Viewing is subject to consumer-applied access controls however consumers can see everything in their own record.
** currently only available in some States/Territories.

Stakeholders should consider that the My Health Record system will evolve with consumer and healthcare provider feedback. With this evolution will likely come more and different clinical content.

As at 27 August, 2017, 5 million Australians' have a My Health Record. It is anticipated that the number of Australians with a My Health Record will continue to rise as will the registration and use by healthcare provider organisations. In 2018, the My Health Record system will switch to an 'opt-out' participation arrangement.

The My Health Record system holds the first data set that has the potential to allow analysis around the full set of health services received by a person (as opposed to the services provided to a person by, usually, a single organisation or program). It can enable time series analysis and multi-service provider analysis. Improving the availability and use of health data to inform clinical, research and resource allocation decisions offers the opportunity to reduce costs and improve health outcomes. As the My Health Record system develops further, the expected improvements in completeness and coverage will yield a rich dataset that could not only advance quality, safety, continuity of care, health outcomes and efficiency, and reduce waste and duplication in the Australian healthcare system[8], but also has the potential to be an important resource for health, clinical and medical research. As the number of individuals with a My Health Record and the number of providers accessing and updating those records continues to increase overtime it will result in the My Health Record data becoming an even more valuable and sought after resource.

3
Secondary Uses and Users of the My Health Record data

This Chapter explores potential secondary uses and users of the My Health Record data based on national and international experiences of secondary uses of similar data sets.

3.1 POSSIBLE SECONDARY USES OF MY HEALTH RECORD DATA

There are many existing public and government health datasets being used for secondary purposes. Although there is no current secondary use of My Health Record data, existing health (e.g. Hospitalisation data; Cancer Registry data; Childhood Immunisation Records) and non-health (e.g. Air Travel data; Births, Deaths and Marriages Registrations) data sets are collected in Australia and around the world and have uses for secondary purposes. Hermon and Williams (2014) found that big healthcare data are being used predominantly for[13]:

- administration and delivery – managing healthcare delivery and cost
- clinical decision support; and
- understanding behaviour and lifestyle factors of both individuals and the public.

Four case studies highlighting examples of the beneficial secondary use of health data are provided in Appendix A (including references). In brief, the case studies cover:

- The pioneering work in Western Australia that used data linkage to establish that birth defects arising from folate deficiency, and issues associated with low birth weight babies and pre-term deliveries, cerebral palsy, and spina bifida research. As a result of this discovery, Australian federal and state governments agreed to introduce the compulsory enrichment of bread-making flour with folate.
- An analysis using data from the Pharmaceutical Benefits Scheme that measured the impact of the Australian Broadcasting Corporation's Catalyst program titled "Heart of the Matter", which was critical of the use of statins in patients with increased risk of cardiovascular disease. This research demonstrated the power of the media in influencing public opinion and behaviour, as well as illustrating the consequences when the media get it wrong.
- An analysis that linked primary care and hospital data for Indigenous people in the Northern Territory to measure the rate of Potentially Avoidable Hospitalisations (PAH). This research demonstrated that improved access to primary care can reduce the rate of PAH's in people with diabetes. It resulted in improved access to primary care in remote communities for the management of diabetes, which in turn resulted in net health benefits to patients and cost savings to government.
- The work done by the Institute for Clinical and Evaluative Sciences (ICES) in Canada that depends on the linkage of individual-level health data from a variety of sources. One ICES study showed that there were important variations across care settings in the occurrence of inappropriate shocks and deaths for patients who had received an implantable cardio defibrillator. These findings were different to results from clinical trials, and contributed to policy formulation in the Ministry of Health.

These case studies demonstrate national and international use of health data sets for a range of purposes, which importantly include improving the quality of care and making the health system work better to deliver good treatment and economic (in terms of value per healthcare dollar invested) outcomes. They cover the use of individual datasets such as the PBS example, as well as linked data sets such as the Western Australian, Northern Territory and ICES examples. They also demonstrate beneficial de-

identified data (e.g. PBS data to explore impact of catalyst program) and identified data uses (e.g. the linkage of primary care and hospital datasets for Indigenous people in the Northern Territory).

Question 1: What secondary purposes, if any, should My Health Record data be used for?

Question 2: What secondary purposes should My Health Record data not be used for?

3.2 POSSIBLE SECONDARY USERS OF MY HEALTH RECORD DATA

The Australian Government Public Data Policy Statement (December 2015) states that the "Australian Government commits to optimise the use and reuse of public data; to release non-sensitive data as open by default; and to collaborate with the private and research sectors to extend the value of public data for the benefit of the Australian public"[14]. Health information is generally considered to be sensitive, so it is important that the Framework identifies criteria for appropriate secondary users of My Health Record system data, in a way that is aligned with government policy around release of data.

In Australia, data held by government agencies are made available to a requesting party/organisation subject to the request meeting certain criteria. For example, requests received by the Department of Human Services (DHS) are subject to a clarification, assessment and approval process through the External Request Evaluation Committee (EREC). This process establishes if the request can be met and if the intended use of the data is considered appropriate (benefits will accrue). The Australian Bureau of Statistics (ABS) and the Australian Institute of Health and Welfare (AIHW) have similar processes.

Internationally, many OECD countries allow national health data to be released to those eligible organisations (based on assessing if the purpose is ethical, lawful, appropriate, uses data transparently, and evaluates benefits and harms) with research and development objectives including:

- eligible researchers (e.g. non-profit/academic);
- eligible governmental uses (including custodians of the data);
- accredited commissioning organisations/data processing organisations (i.e. purchasers of healthcare services); and
- eligible external parties (these may include private, non-governmental applicants and foreign applicants).

Data governance should also include the secure destruction of data. It is likely that there will be interest in using data held in the My Health Record data from a similar set of organisations in Australia.

Question 3: What types of organisations/individuals should be able to access My Health Record data for secondary purposes?

Question 4: Should access to My Health Record data for secondary uses be restricted to Australian users only or could overseas users be allowed access?

3.3 PRINCIPLES TO GUIDE SECONDARY USE OF MY HEALTH RECORD DATA

The inclusion of a set of principles in the Framework to guide the secondary use of My Health Record data may provide a valuable reference point when assessing secondary use requests for data access. Application of the principles is likely to improve the consistency of decision making with respect to processing data requests, as well as guiding the process for the release and subsequent use of the data.

Appendix B provides more detail (including references) on four examples of principles that are used in Australia and the UK, which guide the use and release of data by some organisations. In summary:

- In 2009, Australian Government Portfolio Secretaries established a Cross Portfolio Statistical Integration Committee (CPSIC), jointly chaired by the ABS and the (then) Department of Health and Ageing, to create an Australian Government approach to facilitate linkage of social, economic and environmental data for statistical and research purposes. The CPSIC has developed seven high level principles for the integration of Commonwealth data for statistical and research purposes.
- Under the ABS transformation agenda, microdata access is a priority. The ABS' microdata access process embraces the Trusted Access Model[15], adapted from international best practice. It is built on the recognition that mutual benefits flow from researcher access to public data and the value of partnerships that reflect trust and shared accountability. The model is being implemented in the ABS using the Five Safes Principles (Safe people, Safe project, Safe setting, Safe data, and Safe output) for the assessment of disclosure risk.
- The Department of Health released its "data access and release policy" in August 2015 with objectives including improving public benefit from increased data, protecting individual privacy, and ensuring efficient approval, extraction and release processes.[16]. The policy includes principles and guidelines for access and release of both "Low Risk De-Identified, Confidentialised or Non Re-Identifiable Data" (refer to principles 1-3) and "High Risk - Identifiable Data" (refer to principles 4-5).
- In 2016 the department established the executive level Data Governance and Analytics Committee to foster increasing access to health data assets and analysis to inform improvements in health outcomes, while maintaining data security and individual privacy. Under the Data Access and Release Policy the Department agreed a Data Governance and Release framework that is being progressively implemented across the Department during 2017. The secondary use Framework that is being developed aims to provide a consistent, transparent, reportable and auditable approach to data release and established a new risk assessment model for data releases based on the Five Safes framework as described by Desai, T., Ritchie, F. and Welpton, R. in *Five Safes: Designing data access for research* (University of the West of England, 2016).
- The National Health Service (NHS) in the UK has subsumed the Caldicott principles into the NHS confidentiality code of practice[17]. They describe general principles that health and social care organisations throughout the NHS should use when reviewing their use of client information. There were initially six principles as published in the 1997 Caldicott Report[18] but a seventh was added after a follow-up report was published in 2012[19].

Question 5: What principles, if any, should be included in the Framework to guide the release of data for secondary purposes from the My Health Record system?

4
Process for requesting and accessing data

This Chapter explores processes for the release of data for secondary uses and possible formats in which the data may be released, as well as governance arrangements to oversee the release of My Health Record data for secondary purposes.

4.1 GOVERNANCE ARRANGEMENTS

Data governance refers to the overall management of the availability, access and release, usability, integrity, and security of the data held by an organisation (e.g. government, non-government or private organisations). Data governance arrangements often include a governing body or council which oversees and implements a defined set of procedures (i.e. as will be described in the developed Framework), and a plan to execute those procedures (i.e. as will be described in the developed Implementation Plan). The data governance arrangements should ensure that data can be trusted and that people can be made accountable for any adverse event that happens because of low data quality or inappropriate release of data (or decision for non-release). The Framework will need to articulate the process for, and what the arrangements will be to govern, the release of data from the My Health Record system.

In Australia data custodians, such as AIHW, ABS, Health and DHS, play a central role in privacy protection and use of data for monitoring and research. Examples include:

- If a researcher from a private sector organisation requests Medicare data about a number of services provided by orthopaedic surgeons in the previous five years from DHS, there is a governance body to determine if the request has merit and ensure the data are provided in a suitable format.
- In a statistical data integration project, the data custodians are agencies responsible for managing the use, disclosure and protection of the source data used.[20] Accordingly, for any given data integration project there may be more than one data custodian (e.g. Commonwealth, state/territory agencies and non-government organisations such as universities and private sector businesses).

Governance arrangements in Australian agencies that are data custodians generally include multiple bodies. For example, the AIHW data governing bodies include the Board, the Director and the Ethics Committee. All bodies have a role in privacy protection and use of data for monitoring and research.

The role of data custodians includes maximising the inherent value of data assets and mitigating privacy concerns associated with the use of the data. The role often includes:

- safe storage of unit record level information;
- assessing the level of risk for each data integration project;
- ensuring compliance with relevant legislation, including privacy, for data release;
- entering into agreements with integrating authorities;
- safe transmission of data; and
- maximising the value of data holdings.

The Office of the Australian Information Commissioner (OAIC) is the independent regulator of the My Health Record system, and plays a key role in protecting the privacy of individuals' information. This is discussed further in Chapter 6 of this paper. Accidental misuse would likely be dealt with under the Australian Privacy Principles and may lead to an investigation by the OAIC.

Overseas, organisations that have data custodian responsibility for one or more data sets have established governance arrangements, which include either a single governing body (that may or may not be independent of the data custodian organisation) or multiple bodies responsible for data governance. For example in England, the secondary use service (SUS) within the NHS England operates a strict information governance process to ensure data are protected from unauthorised access[21]. In Scotland, NHS Research Scotland oversees access to NHS Scotland data and is responsible for complying with legislation and NHS policies which govern the access to data. In addition, NHS Research Scotland has established a bio-repository network of a range of datasets, but the point of difference is single sign off across Scotland (under the universities umbrella agreement of single contracting) to reduce regulatory burdens in ethics approvals[22]. In Wales, Secure Anonymised Information Linkage (SAIL) governance arrangements include seven different bodies (see Appendix C).

4.2 REQUESTING AND GAINING APPROVAL FOR THE DATA TO BE RELEASED

The Framework will need to describe a process by which organisations that are not seeking to access data solely for commercial and non-health related purposes can make a data request to use My Health Record data for secondary purposes.

There are already many established processes and models for requesting data from health data collections that the Framework can leverage, for example:

- DHS, the data custodian of MBS and PBS data, requires all ad hoc data requests for statistical information to be submitted by email. Once approved, the provision and release of data must meet the requirements set out in the bilateral agreement between DHS and the Department of Health. If approval has been granted to receive de-identified data, there is a requirement to destroy the data 12 months after initial supply (or as otherwise negotiated). Approved parties are also required to sign a Statutory Declaration and return it to DHS upon destruction of the data.

  If researchers have ethics committee approval to obtain identifiable data, DHS will send letters to the target patient group about the research and provide contact details of the researcher. It is then the responsibility of the researcher to obtain the appropriate levels of informed consent so that DHS can provide the identifiable data to the researcher. DHS never provides the patient contact details or any identifiable data to the researcher without patient consent[23].

  Principles from this model which could be adopted for the Framework for the secondary use of My Health Record system data include the requirement to destroy the data after initial supply, and/or requiring research organisations obtain the appropriate consent from consumers.

- A request for data held by the AIHW is made through the AIHW custom data request service and involves the organisation/researchers completing a web-based form. Data requests are clarified and considered by AIHW internal data committee and they also seek approval for release from the data custodians. Data requests that require ethical approval are considered by the AIHW Ethics Committee and there is an administration fee for each ethics committee application. The time taken to provide data varies from weeks to several months depending on the complexity of the request.

  Principles from this model which could be adopted for the Framework for the secondary use of My Health Record system data include ensuring an efficient approach to managing each request through a single committee.

- The process used by Population Health Research Network is similar to that of other organisations. Researchers submit a request, ethics approvals are required for accessing the linked data, the linkage is undertaken by PHRN and then the researcher is provided with de-identified linked data for their research purposes.

As in Australia, internationally the roles of ethics committees are key to requesting and gaining approval for using health data. There are numerous approval bodies within each country and within regions within each country, each with their own applications for data access, and approval processes.[24]

Question 6: Which of the governance models described above should be adopted to oversee the secondary use of My Health Record data?

Question 7: What principles, if any, should be adopted to enable organisations/researchers to request and gain approval for de-identified data from the My Health Record system to be provided for secondary purposes?

Question 8: What principles, if any, should be adopted to enable organisations/researchers to request and gain approval for identified data from the My Health Record system to be provided for secondary purposes?

4.3 HOW SHOULD DATA LINKAGE REQUESTS BE HANDLED

The Framework is being developed with a view to defining a role for a single accountable authority for the management of My Health Record data for secondary uses to minimise the risks associated with privacy and security breaches such as re-identification of data.

As presented in Chapter 3, there are potentially many reasons why organisations may want to link My Health Record data with other data sources. The Framework should describe under what circumstances, and how, data from other data sources can be linked to My Health Record data. This section explores the various data linkage request possibilities.

4.3.1 Australian data linkage organisations

Under Australia's *Privacy Act 1988* (Privacy Act), national health and medical research guidelines enable the requirement for patient consent to be waived by an authorised ethics committee in cases where the collection of consent is impractical or impossible and where the outcome of the project in terms of the public good outweighs any infringement of patient privacy. In the past, once approval was sought from the relevant data custodian, it was possible for researchers to access identifiable data to undertake their own data linkages for their approved project. In recent years, this practice has become rare. Researchers with an approved research project instead benefit from national or state-level data linkage centres that conduct data linkages on their behalf, and provide them with access to de-identified data. There are several data linkage units in Australia, involved in the PHRN collaboration (see Table 4.1).

Table 4.1: Australian Data Linkage Units[25]

Data Linkage Unit	Location	Jurisdiction
AIHW	AIHW, ACT	National
Centre for Data Linkage	Curtin University, WA	National
Data Linkage Branch	Department of Health, WA	Western Australia
SA-NT DataLink	University of South Australia	South Australia and Northern Territory
CHeReL	Ministry of Health, NSW	NSW and ACT
Centre for Victorian Data Linkage	Department of Health and Human Services	Victoria
Tasmanian Data Linkage Unit	University of Tasmania	Tasmania
Queensland Research Linkage Group	Queensland Health	Queensland

4.3.2 High risk data linkage projects

Under the national principles for data integration, projects involving identifiable data that are considered to be "high risk" (as determined through application of the risk assessment guidelines)[26], and therefore requiring extra protection of data privacy, would be processed only by an organisation that has been accredited as an Integrating Authority. The AIHW, the ABS and Australian Institute of Family Studies (AIFS) are currently the only three organisations accredited by the Cross Portfolio Data Integration Oversight Board as Integrating Authorities in Australia. These organisations can undertake high risk data integration projects involving Commonwealth data for statistical and research purposes. Integrating Authorities are required to provide a summary of all approved projects on their website.

In Australia an accredited Integrating Authority will link data for an approved project that includes personal administrative data held by Australian Government agencies as well as state-level authorities. To be approved the researcher has to demonstrate to the accredited Integrating Authority that approval has been secured from all of the data custodians and relevant HRECs. An accredited Integrating Authority may return a de-identified linked data file to a researcher for use if the linkage does not involve a "high risk" database. Research with "high risk" databases may only occur using a secure on-site data laboratory or within the secure remote data access facility called Secure Unified Research Environment (SURE). To access SURE, researchers must sign an agreement of use and complete SURE user training.

4.3.3 Organisations requesting linkage of their own data cohort

External researchers may request to have a cohort of data they have collected linked to the My Health Record data. For example, researchers may want to link My Health Record data to a database of clinical trial participants to investigate subsequent hospitalisations, diseases and death. Such linkages will provide very important information about the effectiveness and safety of treatments and clinical care. At the same time, such linkages pose additional risk to data protection because the researches involved have the ability to re-identify data within a de-identified database.

In Australia, the AIHW may agree to conduct a data linkage project involving a researcher's own data if all data custodians involved have approved the linkage and if a waiver of the need for consent has been provided by all human research ethics committees involved. Further, in most Australian States and Territories, deceased persons are not within the scope of the Privacy Act and the AIHW ethics committee has approved the linkage of death data without consent under certain conditions.

Question 9: Should there be specific requirements if researchers/organisations make a request that needs the My Health Record data to be linked to another dataset? If so, what should these requirements be?

4.4 ANONYMISATION AND DE-IDENTIFICATION

Prior to the release of My Health Record data for secondary purposes, the governing body will want to ensure that privacy is protected, and de-identification methods render the information not individually identifiable. There is a need to balance maximising the utility of data with the risk of breaching an individual's privacy, or causing harm to individuals in some way. The Framework will need to provide guidance on how and when certain techniques should be applied to the data prior to its release. Potentially data could be released for secondary purposes as patient identified, de-identified, anonymised or pseudo-anonymised (see Glossary). This section presents examples of current processes and practices.

4.4.1 Processes used in Australia to anonymise and/or de-identify data

The following techniques are used in Australia to ensure that releases and/or published data protects the identity of individuals:

(1) **Separation principle.** The separation principle provides a mechanism to protect the identities of individuals and organisations in datasets, applied as part of the linking and merging process used to form the integrated dataset. It means that no one working with the data can view both the personal information (such as name, address, date of birth or ABN) together with the content data (such as health information) in an integrated dataset[27]. This protocol is regarded by many as the best practice approach in data linkage as its application addresses privacy issues in research design.

(2) The **National Statistical Service** provides a number of principles for ensuring the confidentiality of data prior to release[28]. The techniques involved can be summarised as per the following:

- combining (or collapsing) categories;
- suppression (not releasing information for unsafe cells); and/or
- perturbation (altering the identifiable data in a small way without affecting aggregate results).

(3) **Statistical Linkage Keys** (SLK) maintains confidentiality by ensuring that identifying information (such as name, date of birth and address) is not transferred from one data set to another. They are derived using parts of a patient's name, gender, and date of birth, but are not unique since it is possible for a derived SLK to exactly match the SLK of another patient in a small percentage of patients. SLKs are then converted into unintelligible strings using hash algorithms. The use of project specific hash keys ensures that patient data from different research projects cannot be combined, and that individual patients cannot be re-identified. Due to the dependence on name and date of birth, the linkage is of lower quality and is less accurate, particularly for Vietnamese, Refugee and Indigenous populations[29].

(4) **The Secure Unified Research Environment (SURE)**[30] is a remote-access computing environment that allows researchers to access and analyse the approved linked data extracts provided by data custodians for their research projects. Within SURE each researcher is allocated a virtual computer that runs entirely on hardware physically located at, and controlled by, SURE. It aims to:

- minimise the risk of privacy and/or confidentiality breaches when conducting research using linked, de-identified health and other data, by supplying a remote-access computing environment with comprehensive security features, which replaces a researcher's local computing environment;
- improve the accessibility of linked data to accredited researchers undertaking ethics committee-approved population health, health services and related research; and
- prevent researchers from downloading potentially identifiable data or combining data with other external data sets.

National Health and Medical Research Council (NHMRC) approved Human Research Ethics Committees and Data Custodians also provide specific direction regarding the level of detail that can be publically released for any given research project. For example, a research project analysing the incidence of new HIV cases might be prohibited from releasing information at the postcode level and be required to categorise age in ten-year age groupings. A recent longitudinal cross-jurisdictional linked data analysis for end-of-life patients highlighted major inconsistencies in such directions between jurisdictions[31].

4.4.2 Processes used overseas to anonymise and/or de-identify data

In the US the Privacy Rule provides two de-identification methods[5] the:

- **Expert Determination Method:** takes a risk-based approach to de-identification that applies current standards and best practices from the research to determine the likelihood that a person could be identified from their protected health information. This method requires that a person with appropriate knowledge of and experience with generally accepted statistical and scientific principles and methods render the information not individually identifiable.
- **Safe Harbor Method:** the removal of specified individual identifiers (e.g. all dates must be generalised to year and zip codes reduced to three digits) as well as absence of actual knowledge by the covered entity that the remaining information could be used alone or in combination with other information to identify the individual[32].

As de-identification can limit the utility for surveillance of routine health data, the system in the US allows re-identification (by providers or regional health information organisations) through a process of randomising patient source codes, for the purposes of public health alerts or case reports[33].

In England, NHS England provide data as patient identifiable, anonymised or pseudo-anonymised as required. In Sweden, there is a unit in the National Board of Health that is permitted to access identifiable data and that undertakes data linkage. Linked files then have identifying numbers removed before they are provided to individuals within the Board for analysis. Data analysts never see identifiable data.

4.4.3 Re-identification

Data re-identification is any process by which anonymised data is matched to its true owner after it has been released in de-identified form. The primary risk of re-identification is to individuals who may be embarrassed or harmed as a result of re-identification.

In addition to the potential effects upon individuals whose privacy is compromised, instances of data re-identification also threaten to erode the social licence of data custodians tasked with facilitating the legitimate secondary use of public data.

In an environment where government data applications are more common and are accordingly facing increased scrutiny, debates on issues like this are appropriate. However, there is also a risk that concerns about data re-identification may be confounded, and the social licence of data custodians to facilitate and undertake legitimate secondary use analytics may be lost.

Question 10: What processes should be used to ensure that the data released for secondary purposes protects the privacy of an individual?

Question 11: What precautions should be taken to reduce the risk of de-identified data from the My Health Record system being re-identified after release?

4.5 PROCESSES BY WHICH MY HEALTH RECORD DATA COULD BE RELEASED

There are various ways in which data from the My Health Record system could be made available to users. It will be important for the Framework to consider these options and include the processes by which data can/will be released. This section describes ways in which other data sources are made accessible to those seeking access to the data.

Data held in the My Health Record system could be provided to users in a number of ways, including:

- **Routine publication.** Like the AIHW and ABS, the Agency could produce routine publications (or be part of their Annual Reports) including an analysis of the data held within My Health Record. Not only would the analysis include some descriptive statistics but it could also include a quality statement about the data (similar to the approach adopted by the ABS).
- **Data cubes.** Like the AIHW, ABS and DHS, the Agency could provide some de-identified, aggregate level data accessible through data cubes from its website. Access to data in this format would enable more users to explore potential uses of the My Health Record data and limit the number of requests made for access to data.
- **Provision of data files.** Many data custodians, such as the DHS and AIHW, provide the approved data as an Excel, TXT or CSV file. These data may be transmitted via email or through the post on a storage device (either as an encrypted file or not depending on the level of sensitivity of the data). The data files could also be provided to the approved user using secure file transfer service.
- **Restricted data platforms.** In Australia, researchers with approval to undertake a project involving personal health data considered to be "high risk" are able to access the linked de-identified data through a secure data linkage environment called SURE that is offered through the PHRN[34]. Files cannot be transferred between studies or between the study and the local computing environment. Researchers can look at records on screen to resolve issues with their analysis, but they cannot print the screen or download any data. All outputs of their results are checked for confidentiality.
- **Safe Havens.** Used in several countries (e.g. Scotland[35], Denmark, UK), Safe Havens provide data linkage services and access to de-identified health microdata to approved researchers through a secure real-time remote data access system. Approved researchers access the data held in Safe Havens via a dumb terminal in a secure access facility. The dumb terminals are configured so that the researcher cannot download or remove any of the data or outputs held at the Safe Haven. Analytical software is available within the Safe Haven for use by researchers. Safe Havens carry out statistical disclosure control on outputs to prevent accidental disclosure of identifiable information. Only aggregate results are allowed out of the Safe Haven environment.

Question 12: What arrangements should be considered for the preparation and release of My Health Record data and who should be responsible for undertaking and overseeing these arrangements?

Question 13: Whose responsibility should it be to make a quality statement about the My Health Record data and to ensure the data are of high quality?

5
Monitoring and assurance process

This Chapter presents information about existing processes used by organisations within Australia and overseas to ensure that the secondary use of data complies with the approved purpose. It also presents some of the risk mitigation strategies used.

5.1 PROCESSES USED TO ENSURE COMPLIANCE WITH APPROVED PURPOSE

Once data are released to researchers from the My Health Record system it may be important to monitor and assure compliance with the conditions of release, to ensure that there is no unapproved use. The processes by which the released data will be monitored and compliance with the release conditions assured would need to be included in the Draft Framework. This section presents processes used by Australian organisations as a means to ensure data usage is compliant with approved purposes.

5.1.1 Processes used in Australia to ensure data compliance

Researchers requesting data are often required to advise the data custodian:

- in what formats the data/information will be stored during and after the research project;
- what measures will be taken to ensure the security of information from misuse, loss, or unauthorised access while stored during and after the research project;
- in what format the information will be stored at the completion of the project (e.g. individually identifiable, re-identifiable, non-identifiable);
- how long the information will be stored after the completion of the project and why has this period was chosen;
- what arrangements are in place with regard to the storage of the information collected for, used in, or generated by, the project in the event that the principal researcher/investigator ceases to be engaged at the current organisation; and
- whether the information will be disposed of at some stage and how.

When ethics approval is required to access the data, ethics applications must be reviewed by the data custodian in charge of the records and by a HREC to ensure that only the information absolutely necessary for the fulfilment of the research project is provided. The use, disclosure and retention of information is also limited. For example:

- researchers are only permitted to use the information for the particular project they have received approval for, and in the precise way that has been approved;
- each researcher working on a project must be identified and approved and the information may not be given to another person; and
- the information may only be kept for the period of time approved for the research project and it must then be returned to the data custodian or destroyed - this condition is managed by either the contract the researcher has with a data custodian, their HREC approval conditions or by arrangement with the data linkage unit.

Both the HREC and the data custodians have the right to audit/monitor/check that the researchers are adhering to the agreed security and data disposal plans. If the researchers do not follow the agreed plans

they will be in breach of both their contracts with the data custodians and their HREC approval. Potential consequences of non-compliance include:

- HREC approval suspended or withdrawn (this would stop the researchers completing the project or publishing the results);
- data custodians refusing to provide data to the researcher in the future;
- data custodians refusing to provide data to the researcher's institution in the future; and/or
- legal action over the breach of contract.

It is also common for researchers to sign data agreements with each data custodian, which include agreed terms about disclosure, restrictions of use, indemnity and survival. Researchers may need to provide any output to the data custodians for review and seek data custodian approval prior to publishing. Breaches while expected to be rare must be reported to the nominated governance body.

5.1.2 Processes used overseas to ensure data compliance

An Organisation for Economic Co-operation and Development (OECD) study reported that national health research initiatives in most countries would be placed at risk by any serious breach in data security[36]. Data security measures adopted internationally include:

- requirement for employees to sign confidentiality protection agreements;
- training of staff and external researchers; differential levels of data access for staff;
- monitoring staff access to data;
- secure buildings; secure data storage and transfer; whole of government regulations and reporting requirements;
- data security audits; and
- legal penalties for breaches of data security.

In Finland, when a researcher applies to access data, their application must demonstrate how their institution or university respects data protection requirements. Data are provided to the researcher on a compact disc that has been encrypted and the encryption key is provided to the researcher in a separate communication. Only identified and approved individuals who have been named may access the data.

In Denmark, the project approval will describe to the researchers the retention period of the file and will bind the researcher to not linking the data to any other databases and to not disclosing the data to a third party. The data protection authority in Denmark is then responsible for follow-up with the researchers to ensure compliance and data security audits take place. Non-compliance is a legal violation and subject to penalties. At the data destruction date, the researcher is given the option to de-identify the data, if they would like to retain the data for a longer period.

In Scotland, access to data is via a secure real-time remote data access system. There are penalties for anyone who abuses personal data. Researchers are bound by a strict code, which prohibits disclosure of any personal identifying information. Each Safe Haven adheres to the Data Protection Act, Caldicott Principles (where required), Data Sharing Agreements, Governance Agreements, Ethics Approvals and other relevant agreements. The fact that researchers can only analyse the data within a Safe Haven environment and only aggregate results are allowed out of the Safe Haven environment enables data custodians to be confident that no one organisation will see linked identifiable data; that the data will never leave the Safe Haven environment; and that data will still be under their control. Thus, this approach has persuaded data custodians to contribute data to the Safe Havens.

In Switzerland, when data files are provided to an external researcher, their contract with the Federal Statistical Office binds them to protect the data and to follow the guidelines they are given. They are

warned that they will be required to destroy the data if there is any infringement of these requirements. In practice, researchers want to be able to continue to collaborate with the Statistical Office and will follow the requirements. There is no audit of external researchers but there is tracking of their external publications to ensure that their use of the data is consistent with the agreed-upon purpose.

Question 14: What monitoring and assurance processes, if any, should be considered to ensure My Health Record data *secondary* users comply with the Framework?

5.2 RISK MITIGATION STRATEGIES THAT HAVE BEEN IMPLEMENTED

There is variation in the risk management approaches and experiences across countries driven by the differences in attitudes towards allowing data access and ensuring privacy[37]. Sources of variation include:

- exemption to patient consent requirements may be granted;
- processes for release of data from data custodians with other government authorities;
- approval pathways and governance with defined criteria;
- mechanisms for respecting privacy and access to data.

The Commonwealth High Level Principles and governance arrangements introduce three elements unique to the Commonwealth arrangements that are considered fundamental to achieving an effective and safe environment for data integration projects, and hence are risk management strategies, involving Commonwealth data:

- projects are only undertaken where they are expected to deliver overall benefit to the public which outweighs the privacy imposition and risks to confidentiality;
- use of an accredited Integrating Authority (responsible for managing integration of datasets) supports projects where there is a high systemic risk determined; and
- registration of each project on the Public Register of Data Integration Projects[38] to ensure transparency.[39]

Recommendations from the Deeble Institute[8], relevant to mitigating risk in secondary use of health data include:

- comprehensive system security and privacy safeguards;
- tailored threat, risk and privacy impact assessments;
- flexible and clear policy and technical frameworks that are adaptable to clinical need; and
- action plan to implement a mix of technology, policy and process mechanisms aimed at strengthening security and privacy controls.

It is important to note that the misuse of information in either the My Health Record system or Healthcare Identifiers Service, and other activities that relate to the security and integrity of the My Health Record system, are subject to penalties under the *My Health Records Act 2012* and *Healthcare Identifiers Act 2010*. The serious penalties (see Appendix E) relating to the misuse of information do not apply to accidental misuse. The unauthorised collection, use or disclosure of information will only incur a penalty if the person knows or is reckless as to whether that action is unauthorised. This means that if a person accidentally collects, uses or discloses this information, for example, a healthcare provider organisation inadvertently or accidentally accesses an individual's My Health Record, they are not liable for a civil or criminal penalty (although there may still be an interference with privacy and the Australian Information Commissioner may still be able to investigate).

It is also important to note that all businesses and not-for-profit organisations with an annual turnover more than $3 million have responsibilities under the Privacy Act, subject to some exceptions. Also, some small business operators (turnover of $3 million or less) are covered by the Privacy Act including:

- private sector health service providers including:
  - traditional health service providers, such as private hospitals, day surgeries, medical practitioners, pharmacists and allied health professional;
  - complementary therapists, such as naturopaths and chiropractor;
  - gyms and weight loss clinics; and
  - child care centres, private schools and private tertiary educational institutions;
- businesses that sell or purchase personal information;
- credit reporting bodies;
- contracted service providers for a Commonwealth contract;
- employee associations registered or recognised under the *Fair Work (Registered Organisations) Act 2009*;
- businesses that have opted-in to the Privacy Act;
- businesses that are related to a business that is covered by the Privacy Act; and
- businesses prescribed by the *Privacy Regulation 2013*.

The Privacy Act does not cover:

- State or territory government agencies, including state and territory public hospitals and healthcare facilities (which are covered under state and territory legislation) except:
  - certain acts and practices related to My Health Records and Healthcare Identifiers;
  - entities prescribed by the Privacy Regulation.

Internationally, risk mitigation strategies include[40]:

- managing data security;
- minimising who has the data and what data is shared;
- using best practice for sharing, linking and analysing information;
- frequent monitoring and evaluation processes (including legislation compliance, data quality, safety, technology, privacy and confidentiality safeguards); and
- managing vendor's compliance with privacy laws[41].

Most countries note that risk needs to be assessed and mitigated from a number of perspectives and for all parties involved, including health, safety, quality, privacy, technology, information security, data users, and data collectors[42].

Question 15: What risk mitigation strategies should be included in the Framework?

Question 16: Should there be a public register which shows which organisations/researchers have requested data, the purpose, the status of their data request, what they have found by using the data; and any publications that have resulted from using the data?

Question 17: Are the existing penalties under the My Health Record Act sufficient?

6
Supporting legislation and policies

This Chapter describes the legislation and polices which support the release of data for secondary purposes from the My Health Record system and ensures the data remains secure and confidential.

6.1 THE AUSTRALIAN LEGISLATIVE FRAMEWORK

The following Australian legislation is relevant to the secondary use of health data:

- *My Health Records Act 2012;*
- *National Health Reform Act 2011;*
- *Private Health Insurance Act 2007;*
- *National Health Security Act 2007;*
- *Privacy Amendment (Private Sector) Act 2000;*
- *Australian Information Commissioner Act 2010;*
- *National Health Act 1953;*
- *Australian Bureau of Statistics Act 1975;*
- *Privacy Act 1988;*
- *Freedom of Information Act 1982;*
- *Healthcare Identifiers Act 2010;*
- *Human Services Legislation Amendment Act 2011;*
- *Australian Institute of Health and Welfare Act 1987;*
- *Census and Statistics Act 1905;*
- *Health Insurance Act 1973.*

In Victoria, New South Wales and the Australian Capital Territory state health records Acts may also apply[43]. In Victoria and NSW, responsibilities are further clarified under the Health Services Commissioner's (Vic) and NSW Privacy Commissioner's guidelines on the collection and use of health information for medical research[44].

6.2 MY HEALTH RECORDS ACT AND PRIVACY ACT

The *My Health Records Act 2012*, My Health Records Rule 2016 and My Health Records Regulation 2012 create the legislative framework for the Australian Government's My Health Record system. The Privacy Act will generally apply to the My Health Record system in respect of consumers' health information. Among other things, this triggers the OAIC ability to investigate certain matters. The My Health Records Act limits when and how health information included in a My Health Record can be collected, used and disclosed[45]. Unauthorised collection, use or disclosure of My Health Record information is both a breach of the My Health Records Act and an interference with privacy.

The OAIC regulates the handling of personal information under the My Health Record system by individuals, Australian Government agencies, private sector organisations and some state and territory agencies (in particular circumstances). The OAIC's role includes investigating complaints about the mishandling of health information in an individual's My Health Record. The OAIC can also conduct 'Commissioner initiated investigations'.

The functions and enforcement powers available to the OAIC under the *My Health Records Act 2012* and *Privacy Act 1988* include[45]:

- investigating and conciliating complaints;
- accepting enforceable undertakings;
- making determinations;
- seeking an injunction to prohibit or require particular conduct;
- seeking a civil penalty from the Courts; and

- accepting mandatory data breach notifications from the System Operator, healthcare provider organisations, contracted service providers, repository operators and portal operators.

6.3 RELEVANT AUSTRALIAN POLICY AND GUIDANCE INITIATIVES

The Australian Government has several current policy initiatives that support the development, implementation and compliance of national standards and legislation relating to secondary use of health data, including:

(1) **The OECD draft Council recommendation on health data governance**. These recommendations describe the key elements for development of national health data governance[36].

(2) **Australian Government Public Data Policy Statement**[46]. This Statement outlines risk management, intellectual property and data ownership, agency collaboration, jurisdictions standards and interoperability.

(3) **Australian Public Service Better Practice Guide for Big Data**[47]. This Guide includes guidelines for consent for future analysis and use.

(4) **Protective Security Policy Framework and Information Security Manual**.[48] This Framework is directed at ensuring public trust in the storage, access and approved use of information.

(5) **Australian Public Service Big Data Strategy**[49]. This Strategy provides six principles to guide and assist Agencies including: thinking of data as a national asset, privacy by design, data quality and transparency of processes, sharing of skills and capabilities, collaboration with industry and academia, and enhancing the use of open data.

(6) **National Statement on Ethical Conduct in Human Research**[50]. The Statement outlines expectations of organisations and individuals conducting research, including the responsibilities of data custodians providing data, and those receiving, storing and using data.

(7) **Principles on open public sector information**[51]. The Principles were developed by the OAIC through a process of public consultation. They draw on work in Australia and overseas that defines standards and principles to shape government information management practices.

(8) **Productivity Commission Inquiry Report into Data Availability and Use[52].** This report was released on 8 May 2017 and represents one of the most recent iterations in the public conversation about data use in Australia.

There are also several Australian government policy initiatives underway which may influence the development of the Framework. These are described in Appendix D.

Question 18: What policy changes, if any, need to be considered to support the release of de-identified data for secondary uses from the My Health Record system?

APPENDIX A: EXAMPLES OF SECONDARY USES OF HEALTH DATA

AUSTRALIA: The Folate Story	
Context	Neural tube defects are birth defects of the brain or spinal cord. They happen usually in the first month of pregnancy even before many women know they are expecting. Anencephaly is one form of neural tube defect where the brain and skull do not properly develop and babies born with this disorder do not survive. Spina Bifida is the other well-known neural tube defect where the foetal spinal column does not close properly and children are often left paralysed. There is no cure for spina bifida but the research undertaken in WA provided evidence on how to prevent this crippling condition.
Data base/sets used	The Western Australian Birth Defects Registry was established in 1980 and records birth defects diagnosed prenatally and in children up to the age of six years throughout the state of Western Australia. The sources of notification used by the Registry are both statutory (midwives' notifications, death registrations, hospital morbidity data) and voluntary. The latter include clinicians (obstetricians, paediatricians, paediatric surgeons, orthopaedic surgeons), outpatient clinics, child health nurses and specialist services (genetics, pathology, cytogenetics, antenatal ultrasound, newborn screening, rural paediatrics, disability services). Following the suggestions from early studies overseas that folic acid may prevent neural tube defects, a case control study was undertaken on dietary folate and neural tube defects in the early 1980s. Cases of neural tube defects and a comparison group of non-neural birth defects were ascertained from the Birth Defects Registry. Cases were compared with the non-neural controls as well as a group of matched control infants without birth defects.
Findings/Benefits	The initial study found that mothers with high folate levels had low rates of neural tube defects, and this finding contributed to worldwide research showing that neural tube defects could be reduced by up to 70% with enough folate in very early pregnancy. Following this discovery, there were education campaigns to encourage women to improve their folate intake either through supplements or by choosing folate-rich foods, and whilst some progress was made the real breakthrough came in 2007 when Australian federal and state governments agreed to introduce the compulsory enrichment of bread-making flour with folate. All bread (wheat flour) products made in Australia are now folate-fortified and this has two key health benefits: • Having folate in everyday food means a baby is better protected from the very beginning, although additional folate supplements are also recommended if trying to fall pregnant or once a women is aware she is pregnant. • As well as reducing the risk of neural tube defects, folate has also been linked to general health benefits in all ages including a possible reduction in the risk of some cancers and improved cardiovascular health.
Enablers	The Registry brings together multiple sources of notification which have high ascertainment. The WA Government have provided stable ongoing funding since the early 1990s'.
Relevance	Through combining health data sets, researchers have discovered a finding which has reduced the number of births with neural tube defect which results in cost savings to both the health and disability sectors and to families due to a reduction in the lifetime costs associated with caring for a child with a neural tube defect.
Reference	Bower C, Stanley FJ (1989). Dietary folate as a risk factor for neural-tube defects: evidence from a case-control study in Western Australia. *Med J Australia*; 150: 613–19. Bower C, Stanley FJ (1992) Periconceptional vitamin supplementation and neural tube defects; evidence from a case-control study in Western Australia and a review of recent publications. *J Epidemiol Community Health,* 46: 157–161. Bower C (2006). Primary prevention of neural tube defects with folate in Western Australia: the value of the Western Australian Birth Defects Registry. *Congenital Anomalies*; 46, 118–121

AUSTRALIA: Did the ABC's Catalyst program change statin use in Australia	
Context	Statins are recommended nationally and internationally both for primary prevention of cardiovascular events in people at increased risk of cardiovascular disease, and for secondary prevention in those with established cardiovascular disease[53,54]. They are the most commonly prescribed medicines in Australia[55], used by over 30% of the population aged 50 years and older[56]. On 24th and 31st October 2013, the Australian Broadcasting Corporation (ABC) aired a two-part special edition of the science journalism series, Catalyst, titled "Heart of the matter", that was critical of HMG-CoA reductase inhibitors ("statins"). The program questioned the link between high cholesterol levels and cardiovascular disease, and suggested that the benefits of statins had been overstated and the harms downplayed[57]. Nearly 1.5 million Australians are estimated to have viewed each part of the program[58].
Data base/sets used	Used a 10% random sample, from 1st July 2009 to 30th June 2014, of dispensing records from the Pharmaceutical Benefits Scheme (PBS) data set. Using only long-term concessional beneficiaries' data (i.e. individuals' dispensed only medicines attracting a concessional co-payment) the dataset contained more than 190,000 people who were dispensed statins. The analysis included an interrupted time-series analysis to assess the impact of the Catalyst program on dispensing and discontinuation of statins.
Findings/Benefits	Following the Catalyst program, there was a 2.60% reduction in statin dispensing, equivalent to 14,005 fewer prescriptions dispensed Australia-wide every week. In the week the Catalyst program aired, there was a 28.8% increase in discontinuation of statin use, which decayed by 9% per week. An estimated 28,784 additional Australians ceased statin treatment. Up until 30th June 2014, there were 504,180 fewer dispensing of statins, and the authors estimated that this affected 60,897 people. The changes in statin use occurred despite warnings in the Catalyst program that its content should not be taken as medical advice, and public criticism of the program. This research demonstrated the power of the media in influencing public opinion and behaviour, as well as illustrating the consequences when the media get it wrong.
Enablers	Availability of data for a representative sample of all Australians ever dispensed a PBS-subsidised medicine, and of data for a long-term concessional beneficiary population, ensuring complete capture of statin dispensing in these concessional beneficiaries.
Relevance	Secondary use of My Health Record data will enable monitoring of the impact of policy changes, media reports and other events on the use of prescription medications in Australia, while also allowing analysis according to patient risk, and assessment of patient outcomes, using detailed information that is not currently captured in PBS data.
Reference	Schaffer AL, Buckley NA, Dobbins TA, Banks E, Pearson SA. The crux of the matter: Did the ABC's Catalyst program change statin use in Australia? Med J Aust. 2015 Jun 15; 202(11):591-5.

AUSTRALIA: The cost-effectiveness of primary care for Indigenous Australians with diabetes living in remote Northern Territory communities	
Context	Rates of potentially avoidable hospitalisations (PAHs) are indicators of access to primary care and include hospitalisations that may have been avoided by preventing illness or managing chronic disease[59]. In the Northern Territory (NT) Indigenous population, undiagnosed or poorly controlled diabetes often results in serious complications leading to PAH, disability and premature death. Between 1998–99 and 2005–06, Indigenous people in the NT were hospitalised for potentially avoidable causes at four times the rate of non-Indigenous people. This high rate was largely attributable to diabetes complications, and highlights barriers to accessing effective primary care[60].
Data base/sets used	Use of two administrative databases: the primary care information system (PCIS) and the Caresys hospital admission data system which were linked using patients' unique registration numbers. Data from 54 remote clinics and all five public hospitals in the NT for 10 years were extracted for statistical analysis. The government accounting system was used to extract financial data for primary care costs. Operational and personnel expenditures were allocated based on activity. Expenditures covered patient travel, property maintenance and cleaning, and salaries for doctors, nurses and Aboriginal health workers.
Findings/Benefits	Individuals were categorised to one of three groups (low, medium or high) based on their level of use of primary care services. The study found that overall, compared with the low use group, the medium and high use groups (patients who used primary care two or more times annually) experienced lower rates of annual hospitalisation, PAH and death and fewer years of life lost (YLL). The net health benefits in saved hospitalisations provide a summary measure for the value-for-money of primary care. The net health benefit, as measured by hospitalisations saved per person per year, is achieved at a lower cost when primary care is used between two and 11 times per year. The cost of preventing one hospitalisation for diabetes was $248 for those in the medium-use group and $739 for those in the high-use group. In both cases the cost was much less than the mean cost of one hospitalisation, $2,915.
Enablers	The study used reliable, linked data to provide new evidence that there are significant cost savings and better health outcomes for patients with diabetes when access to primary care is improved.
Relevance	Combining health databases even where those databases are limited in scope was used to improve access to primary care in remote communities for the management of diabetes, which results in net health benefits to patients and cost savings to government. The case study also lends weight to the importance of collecting data that provides comprehensive population representation over a significant time period enabling longitudinal analysis.
Reference	Thomas SL, Zhao Y, Guthridge SL, Wakerman J (2014) MJA 2014; 200: 658-662

CANADA: "Why Are Implantable Cardioverter Defibrillator Outcomes in Practice Different from Clinical Trials?"	
Context	The Institute for Clinical and Evaluative Sciences (ICES) is a university-based research centre providing population-based health services research for Canada's largest province, Ontario. The ICES research programme depends on the linkage of individual-level data from a variety of sources. In March 2014, ICES launched a new Data and Analytics Service where academic and non-profit researchers can apply for access to ICES data and linked data and the process for applying for data access has been publicised on the ICES web site. ICES is considering the development of a public benefits test for external applicants.
Data base/sets used	The Ontario Health Insurance Number (HINs) is used for all patient encounters for public healthcare services, including physician claims, drug benefits, and hospital encounters. The Ontario Registered Persons Database includes all HINs associated with individuals by name, address and birth date. ICES receives fully identifiable personal health data from various data custodians, including administrative health data from the Ontario Ministry of Health and Long-term Care.
Findings/Benefits	ICES research has informed the effects of treatments in real-world settings which can differ from results of clinical trials. For example, a recent study of patients who had received an implantable cardio defibrillator determined that after six months there were important variations across care settings in the occurrence of inappropriate shocks and deaths. This evidence contributed to policy planning and evaluation within the Ontario Ministry of Health and Long-term Care.
Enablers	The ICES initiative operates in a favourable political environment. The Ontario provincial government welcomes scientific evidence and research results influence policy. The legislative environment permits data use with privacy protections. The data privacy and security framework is strong. Data are increasingly accessible and secondary data use is regarded seen as positive by the Canadian public. The Ontario HIN is used for all patient encounters for public healthcare services, including physician claims, drug benefits, and hospital encounters. In addition, the Personal Health Information Protection Act (PHIPA) is an Ontario statute that identifies a small number of organisations as prescribed entities that may collect and process personal health information.
Risks	Canada and Australia are both federal systems. The ICES case illustrates ongoing difficulties with achieving integration of data at sub-national level within national federations. In the ICES case there is little progress of any national studies involving the sharing of person-level de-identified data. Such studies are not impossible but they remain difficult and expensive.
Relevance	The Canadian case demonstrates that the collection and use of the information is in the public good by furthering medical research. Note that ICES is also beginning to collect data from electronic medical records. In Ontario there is not yet a central archive for EMR data. There are also efforts at the provincial level to develop a legislation that would authorise an entity to collect Ontario EMR data.
Reference	Krishnakumar N., J.V. Tu, and D.S. Lee (2011), "Why Are Implantable Cardioverter Defibrillator Outcomes in Practice Different from Clinical Trials?" Cardiac Electrophysiology Clinics, Vol. 3, No. 4, pp. 511-520.

APPENDIX B: PRINCIPLES TO GUIDE SECONDARY USE OF MY HEALTH RECORD DATA

A number of organisations or collaborations of organisations have principles which guide the use of data. This section presents high level principles which guide the use and release of data by some organisations.

Cross Portfolio Statistical Integration Committee

In 2009, Australian Government Portfolio Secretaries established a Cross Portfolio Statistical Integration Committee (CPSIC), jointly chaired by the ABS and the (then) Department of Health and Ageing, to create an Australian Government approach to facilitate linkage of social, economic and environmental data for statistical and research purposes.

On 3 February 2010, the Portfolio Secretaries Meeting (now Secretaries Board) endorsed a set of high level principles for the integration of Commonwealth data for statistical and research purposes[61]. These principles (see Table B.1) were developed in recognition of two key issues:

- the potential benefits for research and evaluation that can be achieved through the integration of existing Commonwealth data for statistical and research purposes; and
- the need to protect the personal information of individuals, as set out in the *Privacy Act 1988*, and the confidentiality of data provided by individual businesses.

Table B.1: High level principles for the integration of Commonwealth data for statistical and research purposes

Strategic resource	**Principle 1** Responsible agencies should treat data as a strategic resource and design and manage administrative data to support their wider statistical and research use.
Custodian's accountability	**Principle 2** Agencies responsible for source data used in statistical data integration remain individually accountable for their security and confidentiality.
Integrator's accountability	**Principle 3** A responsible 'integrating authority' will be nominated for each statistical data integration proposal.
Public benefit	**Principle 4** Statistical integration should only occur where it provides significant overall benefit to the public.
Statistical and research purposes	**Principle 5** Statistical data integration must be used for statistical and research purposes only.
Preserving privacy and confidentiality	**Principle 6** Policies and procedures used in data integration must minimise any potential impact on privacy and confidentiality.
Transparency	**Principle 7** Statistical data integration will be conducted in an open and accountable way.

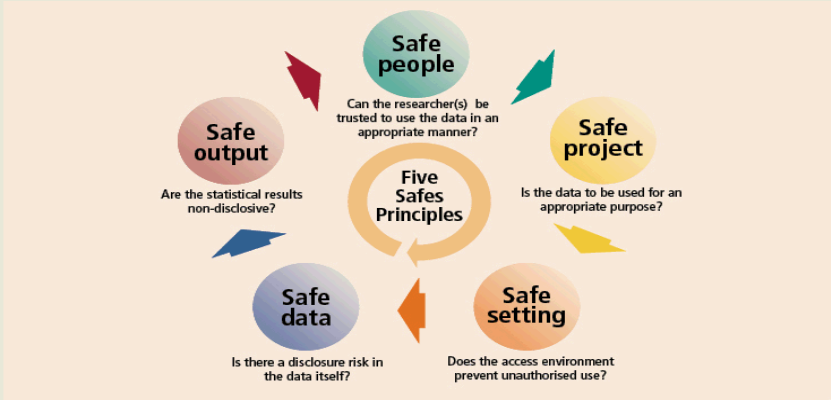
Source: Data Integration Involving Commonwealth Data for Statistical and Research Purposes: Governance and Institutional Arrangements. 6 Oct 2010

Australian Bureau of Statistics

Microdata[62] access is a priority under the ABS transformation agenda and is consistent with the objectives of the Department of the Prime Minister and Cabinet for public sector data management to: *"...optimise the use and reuse of public data; to release non-sensitive data as open by default; and to collaborate with the private and research sectors to extend the value of public data for the benefits of the Australian public."*

The microdata access process embraces the Trusted Access Model[63], adapted from international best practice. It is built on the recognition that mutual benefits flow from researcher access to public data and the value of partnerships that reflect trust and shared accountability. The model is being implemented in the ABS using the Five Safes Principles for the assessment of disclosure risk.

Figure B.1. Five Safes Principles



Department of Health

The Department of Health released their data access and release policy in August 2015[64]. The policy objectives are to:

- improve public benefit from increased data use;
- timely information release;
- relevant information release;
- protect individual privacy; and
- ensure efficient approval, extraction and release processes.

The Policy includes Principles and Guidelines for access and release of both "Low Risk De-Identified, Confidentialised or Non Re-Identifiable Data" (Principles 1-3) and "High Risk - Identifiable Data" (Principles 4-5) as shown in Table B.2.

Table B.2: Principles included in the Department of Health Data Access and Release Policy

Types of data	Principles
Low Risk De-Identified, Confidentialised or Non Re-Identifiable Data	**Principle 1:** Data that can be made public should be made public **Principle 2:** Health should grant structured access[65] to data as well as the delivery of data as a package **Principle 3:** Australian government data is a strategic national asset and agencies, such as Health, should permit researchers, other agencies and the public as much access as possible, while recognising and minimising any risks associated with data exposure.
High Risk - Identifiable Data	**Principle 4:** The Minister and relevant departmental delegates retain all relevant legal responsibility for their identified or identifiable unit record data at all times **Principle 5:** Where data by nature of its level of detail is considered to be a high risk to release publicly, only the elements of data relevant and essential to meet the purpose of a reasonable request shall be made accessible.

NHS Caldicott Principles

The Caldicott principles have been subsumed into the NHS confidentiality code of practice[17]. They describe general principles that health and social care organisations throughout the NHS should use when reviewing their use of client information. There were initially six principles as published in the 1997 Caldicott Report[66] but the seventh one was added after a follow-up report was published in 2012[19]. The seven principles are:

- **Principle 1: Justify the purpose(s).** Every proposed use or transfer of personally identifiable information within or from an organisation should be clearly defined and scrutinised, with continuing uses regularly reviewed by the appropriate guardian.
- **Principle 2: Do not use personally identifiable information unless it is absolutely necessary.** Personally identifiable information items should not be used unless there is no alternative.

- **Principle 3: Use the minimum personally identifiable information.** Where the use of personally identifiable information is considered to be essential, each individual item of information should be justified with the aim of reducing identifiably.
- **Principle 4: Access to personally identifiable information should be on a strict need to know basis.** Only those individuals who need access to personally identifiable information should have access to it.
- **Principle 5: Everyone should be aware of their responsibilities.** Action should be taken to ensure that those handling personally identifiable information are aware of their responsibilities and obligations to respect patient/client confidentiality.
- **Principle 6: Understand and comply with the law.** Every use of personally identifiable information must be lawful. Someone in each organisation should be responsible for ensuring that the organisation complies with legal requirements.
- **Principle 7: The duty to share information can be as important as the duty to protect patient confidentiality.** Professionals should in the patient's interest share information within this framework. Official policies should support them doing so.

In 2016 a further follow-up report was produced[67] following controversy over the care.data initiative.

APPENDIX C: GOVERNANCE ARRANGEMENTS

SAIL stands for Secure Anonymised Information Linkage. SAIL is a Wales-wide research resource focused on improving health, well-being and services. Its databank of anonymised data about the population of Wales is world recognised. SAIL receives core funding from the Welsh Government's Health and Care Research Wales. A range of anonymised, person-based datasets are held in SAIL, and, subject to safeguards and approvals, these can be anonymously linked together to address important research questions.

SAIL governance arrangements include seven different structures including:

SAIL Advisory Board

SAIL has an Advisory Board to provide strategic input from members of the public and leaders in the field of health information and research in Wales. Its aims are to:

- Ensure that the work of SAIL remains in harmony with developments in information strategy for health and social care in Wales and the UK
- Maximise the ability of SAIL to conduct and support research.
- Ensure that the work of SAIL meets the health and wellbeing needs of people in Wales.
- Support the further development and sustainability of SAIL

SAIL Consumer Panel

SAIL benefits from an active Consumer Panel to provide a public perspective on data linkage research. The role of the Panel is to:

- Act as advisors on issues in research
- Advise on how best to engage with the public
- Guide on how to recruit people to study steering groups
- Provide views on data protection issues
- Discuss proposals for research
- Review information designed for a lay audience
- Act as advocates for data linkage research

SAIL Management Board

SAIL has a Management Board, comprised of the SAIL Directors and Management Team, to oversee and direct operational arrangements. Its function is to:

- Oversee infrastructural and governance arrangements for the SAIL Databank
- Plan and oversee system developments, including resource usage, within the SAIL Databank
- Review progress reports, and discuss approaches to achieving deliverables connected with the SAIL databank
- Be advised by, and exchange information with, the SAIL Consumer Panel, Advisory Board and Principal Investigators Group
- Receive reports on data set priority lists and decisions taken by the SAIL Data Management Committee
- Consider issues, referred by the Data Management Committee, and decide on appropriate actions
- Receive notification of new projects and Information Governance Review Panel (IGRP) status reports

SAIL Principal Investigators Group

The SAIL Principal Investigators Group provides a forum to discuss issues connected with use of the SAIL Databank. Its terms of reference are to:

- Be advised of latest developments regarding SAIL
- Learn about updates to SAIL policy and procedures in connection with data use
- Receive reports on progress with new datasets and data refreshes
- Discuss any issues or difficulties encountered with using particular datasets and propose solutions
- Receive notification of Information Governance Review Panel (IGRP) status reports
- Discuss future collaborative opportunities with others that are using SAIL.

Information Governance Review Panel (IGRP)

SAIL established an IGRP that provides independent advice on Information Governance and reviews all proposals to use SAIL data to ensure that they are appropriate and in the public interest. The membership of the IGRP is comprised of representatives from:

- British Medical Association (BMA)
- National Research Ethics Service (NRES)
- Public Health Wales
- NHS Wales Informatics Service (NWIS)
- Consumer Panel

SAIL Data Management Committee

The SAIL Data Management Committee reports to the SAIL Management Board, and manages the operational issues connected with the acquisition, processing and preparation for use of SAIL datasets. Its terms of reference are to:

- Set priorities for dataset acquisition and loading
- Manage the administration of bringing datasets into SAIL
- Manage resource allocation for data acquisition
- Receive reports on progress with new datasets and data refreshes
- Discuss any issues or difficulties encountered with particular datasets and propose solutions
- Consider issues and permissions needed around data use cases for scoping and exploratory work
- Receive notification of new projects and Information Governance Review Panel (IGRP) status reports
- Review and prioritise project data requirements
- Provide visibility to management on dataset priority lists and decisions taken

SAIL User Forum

The SAIL User Forum provides an opportunity for all SAIL users to exchange information about the use of the SAIL system. Its aim is to:

- Meet others that are using SAIL
- Learn about the research projects that are using SAIL
- Demystify the policy and procedures about using SAIL
- Develop and share technical and methodological skills for using SAIL
- Identify professional development training needs for using SAIL

APPENDIX D: EMERGING POLICY FRAMEWORKS

This section presents several Australian government policy initiatives currently underway which may influence the development of the Framework.

Guide to big data and the Australian Privacy Principles

On the 5th August, 2016 the OAIC completed the public comment process on a draft guide to big data and the Australian Privacy Principles (APPs)[68] that it has developed to facilitate big data activities while protecting personal information. The draft Guide outlines key privacy requirements and encourages the implementation of the Privacy Management Framework. Taking this approach will embed 'privacy by design' in entities' culture, systems and initiatives from the design stage onwards.

Australian Population Health Research Network review

The PHRN recently conducted a strategic planning process to produce a five-year strategic plan. The strategic planning process follows an independent review of the PHRN in 2014, which highlighted the importance of a strategic vision for data linkage in Australia. In December 2015, the Australian Government announced that it will allocate $1.5 billion over 10 years from 2017-18 for the National Collaborative Research Infrastructure Strategy (NCRIS) program[69]. This funding is of major significance as it will provide opportunities for long-term investment in Australia's national research infrastructure.

NSW Health Analytics Framework

Many Australian jurisdictions are considering secondary use initiatives to support clinical pathways analysis and measurement of clinical outcomes. For example, in January 2016, NSW Health released the "Analytics Framework - transformed health through data and insights"[70]. That report suggested that analytics will enable NSW Health to provide world-class and truly integrated healthcare by delivering data and insights that support evidence-based decision making, planning and performance.

APPENDIX E: PENALTIES UNDER MY HEALTH RECORD ACT

Action	Penalty
Misuse of information	
Unauthorised collection, use or disclosure of health information in a My Health Record Sections 59 and 60 of the *My Health Records Act 2012*	Civil penalty of up to 600 penalty units ($25,200 for individuals and $126,000 for bodies corporate). Criminal penalty of up to two years imprisonment and/or 120 penalty units ($25,200 for individuals and $126,000 for bodies corporate)
Unauthorised use or disclosure of healthcare identifiers or other information obtained for the purposes of the Healthcare Identifiers Service Section 26 of the of the *Healthcare Identifiers Act 2010*	Civil penalty of up to 600 penalty units ($25,200 for individuals and $126,000 for bodies corporate). Criminal penalty of up to two years imprisonment and/or 120 penalty units ($25,200 for individuals and $126,000 for bodies corporate)
Security and integrity	
If a person accesses the My Health Record system on behalf of a registered healthcare provider organisation and fails to provide enough information to the System Operator to identify that person without seeking more information Section 74 of the *My Health Records Act 2012*	Civil penalty of up to 100 penalty units ($21,000 for individuals and $105,000 for bodies corporate)
Failing to notify an actual or potential data breach in which they were directly involved Section 75 of the *My Health Records Act 2012*	Civil penalty of up to 100 penalty units ($21,000 for individuals and $105,000 for bodies corporate)
Failing to give written notice within 14 days if the entity ceases to be eligible to be registered Section 76 of the *My Health Records Act 2012*	Civil penalty of up to 80 penalty units ($16,800 for individuals and $84,000 for bodies corporate)
Holding, taking, processing or handling, records held for the purposes of the My Health Record system outside Australia, or causing someone else to do so Section 77 of the *My Health Records Act 2012*	Civil penalty of up to 600 penalty units ($25,200 for individuals and $126,000 for bodies corporate). Criminal penalty of up to two years imprisonment and/or 120 penalty units ($25,200 for individuals and $126,000 for bodies corporate)
Failing to comply with the My Health Records Rules that apply to the entity Section 78 of the *My Health Records Act 2012*	Civil penalty of up to 100 penalty units ($21,000 for individuals and $105,000 for bodies corporate)
Failure to notify the Healthcare Identifiers Service Operator of changes to their organisation's information within 20 days Section 25E of the *Healthcare Identifiers Act 2010*	Civil penalty of up to 100 penalty units ($21,000 for individuals and $105,000 for bodies corporate)
Failure to retain identifying information about a person requesting disclosure of healthcare identifiers (if not provided at the time of disclosure) Regulation 7 of the of the *Healthcare Identifiers Regulations 2010*	Civil penalty of up to 50 penalty units ($10,500 for individuals and $52,500 for bodies corporate)

APPENDIX F: REFERENCES

[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] National Statistics Service, 'Rights and responsibilities of data custodians', www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/59fd060543b4e9e0ca257a4e001eacfe/$FILE/Rights,%20responsibilities%20and%20roles%20of%20data%20custodians_Dec2013.pdf (accessed 22 January 2015).
[3] The National Statistics Service is a network of Australian Government and State and Territory entities led by the Australian Bureau of Statistics that works together to improve Australia's statistics system. National Statistics Service, *Data Linking: What is data linking?*, Information Sheet 1, p. 1, www.nss.gov.au/nss/home.nsf/533222ebfd5ac03aca25711000044c9e/91242a5a14b12e26ca257ba8007b0819/$FILE/data%20linking%20w.pdf (accessed 9 December 2015).
[4] http://www.ands.org.au/working-with-data/enabling-data-reuse/de-identifying-data
[5] O'Keefe C, Connolly C (2010) Privacy and the use of health data for research. MJA 193 (9): 537-541 https://www.mja.com.au/journal/2010/193/9/privacy-and-use-health-data-research
[6] http://meteor.aihw.gov.au/content/index.phtml/itemId/349895
[7] Final Report of the National Health and Hospital Reform Commission (NHHRC). June, 2009, *'A Healthier Future for all Australians'*. http://www.health.gov.au/internet/nhhrc/publishing.nsf/Content/nhhrc-report.
[8] Partel, K (2015). Toward better implementation: Australia's My Health Record. Deeble Institute Issues Brief 13.
[9] Hobbs MST, McCall MG. Health statistics and record linkage in Australia. J Chron Dis 1970; 23: 375-81.
[10] Budget Papers 2015-16, Australian Government 2015, p. 113
[11] Safran, C., Bloomrosen, M., et al. (2007). Toward a national Framework for the secondary use of health data: an American Medical Informatics Association white paper. Journal of the American Medical Informatics Association, 14: 1-9
[12] Oderkirk J, et al. International comparisons of health system performance among OECD countries: Opportunities and data privacy protection challenges. Health Policy (2013). http://www.oecd.org/publications/strengthening-health-information-infrastructure-for-health-care-quality-governance-9789264193505-en.htm
[13] Hermon R and Williams PAH (2014). Big data in healthcare: What is it used for? Originally published in the Proceedings of the 3rd Australian eHealth Informatics and Security Conference. Held on the 1-3 December, 2014 at Edith Cowan University, Joondalup Campus, Perth, Western Australia.
[14] https://www.dpmc.gov.au/sites/default/files/publications/aust_govt_public_data_policy_statement_1.pdf
[15] ABS. (2016). 1015.0 - Information Paper: Transforming Statistics for the Future, Feb 2016. Retrieved 8 August 2016 from: http://www.abs.gov.au/AUSSTATS/abs@.nsf/be4aa82cd8cf7f07ca2570d60018da27/e4d483bab4e1ad93ca257f4c00170bb6!OpenDocument
[16] http://www.health.gov.au/internet/main/publishing.nsf/Content/Data-Access-Release-Policy August 2015
[17] NHS confidentiality code of practice. Department of Health. 7 November 2003.
[18] The Caldicott Committee (December 1997). The Caldicott Report. Department of Health. Retrieved 2011-11-21.
[19] The Information Governance Review: To Share or Not to Share (PDF). Department of Health. 2013-03-21. Retrieved 2015-05-14.
[20] NSA (2016). Rights, responsibilities and roles of data custodians. Retrieved 3 August 2016 from http://www.nss.gov.au/nss/home.nsf/pages/Data+Integration+-+Roles+and+responsibilities+of+data+custodians?opendocument
[21] http://www.hscic.gov.uk/sus
[22] SANT Datalink (2015) – Internal presentation
[23] DHS. (2016). Statistical information and data. Retrieved 8 August 2016 from: https://www.humanservices.gov.au/corporate/statistical-information-and-data
[24] Oderkirk J (2016) Case Study notes – Farr Institute UK – Supporting UK-wide data sharing, linkage and use.
[25] http://www.phrn.org.au/about-us/who-is-involved/australian-data-linkage-units/
[26] http://www.nss.gov.au/nss/home.NSF/pages/Data%20integration%20projects%20-%20how%20to%20determine%20the%20risk%20level%20-%20Risk%20assessment%20guidelines
[27] https://statistical-data-integration.govspace.gov.au/topics/applying-the-separation-principle/
[28] http://www.nss.gov.au/nss/home.nsf/pages/Confidentiality+-+How+to+confidentialise+data:+the+basic+principles
[29] Internal SANT DataLink report.
[30] https://www.saxinstitute.org.au/our-work/sure/
[31] McGowan C. Changing Patterns of End of Life Health Care Use. PhD Thesis 2015 (unpublished).
[32] http://www.hhs.gov/hipaa/for-professionals/privacy/special-topics/de-identification/index.html#rationale
[33] Centre for Disease Control and Prevention (2015) Federal Public Health Laws Supporting Data Use and Sharing
[34] Sax Institute. (2016). SURE. Retrieved 3 August 2016 from https://www.saxinstitute.org.au/our-work/sure/
[35] Safe Havens in Scotland were established as part of a national need for delivering research excellence and the need for rapid access to high quality health data for research purposes. They were developed in line with the SHIP blueprint which outlined a program for a Scotland-wide research platform for the collation, management, dissemination and analysis of anonymised Electronic Patient Records (EPRs). Safe Havens operate independently to provide advice, support and a secure environment for access to a wide range of datasets (including national datasets through to specialised local datasets) but also as a federated network across Scotland working to common principles, and standards and common processes optimising the safe and secure flow of data between the different Safe Havens
[36] OECD (October 2015) Health Data Governance: Privacy, Monitoring and Research - Policy Brief http://www.oecd.org/health/health-systems/Health-Data-Governance-Policy-Brief.pdf
[37] Oderkirk J, Klazinga N, Ronchi E (2012) (PPT) Follow up on strengthening health information infrastructure
[38] http://www.nss.gov.au/nss/home.NSF/pages/Data+Integration+Find+A+Project?OpenDocument
[39] https://statistical-data-integration.govspace.gov.au/topics/scope-of-the-commonwealth-arrangements/
[40] Oderkirk J, Klazinga N, Ronchi E (2013) (Conference PPT) Strengthening health information infrastructure for quality management
[41] http://managedhealthcareexecutive.modernmedicine.com/managed-healthcare-executive/content/secondary-use-health-data-increases-security-risks
[42] Herzig TW (2010) Information Security in Healthcare: Managing Risk [accessed at: https://books.google.com.au/books?id=ckswBQAAQBAJ&dq=managing+risk+in+secondary+use+of+health+data&source=gbs_navlinks_s]
[43] Health Records Act 2001 (Vic). Health Records (Privacy and Access) Act 1997 (ACT)
[44] OHSC. (2016).Use and Disclosure of Health Information. Retrieved 8 August from http://www.legislation.act.gov.au/a/1997-125/current/pdf/1997-125.pdf
[45] https://www.oaic.gov.au/privacy-law/other-legislation/my-health-records

[46] PMC (2016).Australian Government Public Data Policy Statement. Retrieved 3 August 2016 from pmc.gov.au/sites/default/files/publications/aust_govt_public_data_policy_statement_1.pdf
[47] Dept. of Finance. (2015). Australian Public Service Better Practice Guide for Big Data. Retrieved 3 August 2016 from http://finance.gov.au/sites/default/files/APS-Better-Practice-Guide-for-Big-Data.pdf
[48] AGD (2016). The Protective Security Framework (PSPF). Retrieved 3 August 2016 from https://www.protectivesecurity.gov.au/Pages/default.aspx
[49] https://www.finance.gov.au/files/2013/06/Draft-Big-Data-Strategy.pdf
[50] NHMRC. (2015). National Statement on Ethical Conduct in Human Research 2007. Retrieved 3 August 2016 from https://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72_national_statement_may_2015_150514_a.pdf
[51] https://www.oaic.gov.au/resources/information-policy/information-policy-resources/principles-on-open-public-sector-information.pdf
[52] Productivity Commission (2017) inquiry report – Data Availability and Use, Retrieved 26 May 2017 http://www.pc.gov.au/inquiries/completed/data-access/report
[53] National Heart Foundation of Australia. The National Heart Foundation of Australia's summary of the recommendations for cholesterol management. Canberra: NHF, 2014. http://www.heartfoundation.org.au/SiteCollectionDocuments/Heart%20Foundation%20summary%20of%20recommendations%20for%20cholesterol%20management%20-%20Catalyst%20-%20FINAL.pdf (accessed Mar 2014).
[54] Stone NJ, Robinson JG, Lichtenstein AH, et al; American College of Cardiology/American Heart Association Task Force on Practice Guidelines. 2013 ACC/AHA guideline on the treatment of blood cholesterol to reduce atherosclerotic cardiovascular risk in adults: a report of the American College of Cardiology/American Heart Association Task Force on Practice Guidelines. *J Am Coll Cardiol* 2014; 63: 2889-2934
[55] Australian Government Department of Health. Australian Statistics on Medicines 2011. Canberra: Department of Health, 2013. http://www.pbs.gov.au/statistics/asm/2011/australian-statistics-on-medicines-2011.pdf (accessed Jul 2014)
[56] Morgan TK, Williamson M, Pirotta M, et al. A national census of medicines use: a 24-hour snapshot of Australians aged 50 years and older. *Med J Aust* 2012; 196: 50-53. https://www.mja.com.au/journal/2012/196/1/national-census-medicines-use-24-hour-snapshot-australians-aged-50-years-and. (accessed May 2014). <MJA full text>
[57] Australian Broadcasting Corporation. Catalyst. Heart of the matter. Sydney: ABC, 2014. http://www.abc.net.au/catalyst/heartofthematter (accessed Apr 2015).
[58] Australian Broadcasting Corporation. Media Watch. Catalyst challenges the mainstream. Sydney: ABC, 2013. http://www.abc.net.au/mediawatch/transcripts/s3888657.htm (accessed Jan 2015
[59] Katterl R, Anikeeva O, Butler C, et al. Potentially avoidable hospitalisations in Australia: causes for hospitalisations and primary healthcare interventions. Primary Health Care Research and Information Service Policy Issue Review. Adelaide: PHCRIS, 2012. http://www.phcris.org.au/phplib/fi ledownload.php?fi le=/elib/lib/downloaded_fi les/publications/pdfs/news_8388.pdf
[60] Li SQ, Gray NJ, Guthridge SL, Pircher SL. Avoidable hospitalisations in Aboriginal and non-Aboriginal people in the Northern Territory. Med J Aust 2009; 190: 532-536. https://www.mja.com.au/journal/2009/190/10/avoidable-hospitalisation-aboriginal-and-non-aboriginal-people-northern
[61] NSA. (2016).About the Commonwealth arrangements. Retrieved 3 August 2016 from http://statistical-data-integration.govspace.gov.au/about-3
[62] Microdata are unit record data where each record represents observations for a person or an organisation. Microdata contain individual responses to questions on survey questionnaires, or administrative forms, including identifying information such as name, address, telephone number and age. Microdata are a valuable resource for researchers and policy makers. The challenge for data custodians is striking the right balance between fulfilling obligations to protect the identity of individuals and organisations, and maximising the information available for statistical and research purposes. This requires careful weighing of the identification risks and benefits.
[63] ABS (2016). 1015.0 - Information Paper: Transforming Statistics for the Future, Feb 2016. Retrieved 8 August 2016 from: http://www.abs.gov.au/AUSSTATS/abs@.nsf/be4aa82cd8cf7f07ca2570d60018da27/e4d483bab4e1ad93ca257f4c00170bb6!OpenDocument
[64] http://www.health.gov.au/internet/main/publishing.nsf/Content/Data-Access-Release-Policy August 2015
[65] Structured access in this policy refers to access to data via query and analytical tools in a controlled environment
[66] The Caldicott Committee (December 1997). "The Caldicott Report. Department of Health. Retrieved 2011-11-21.
[67] Review of Data Security, Consent and Opt-outs". The National Data Guardian. 2016-07-06. Retrieved 2016-07-12.
[68] https://www.oaic.gov.au/engage-with-us/consultations/guide-to-big-data-and-the-australian-privacy-principles/
[69] https://www.education.gov.au/national-collaborative-research-infrastructure-strategy-ncris
[70] NSW Health Analytics Framework Transformed health through data and insights. A five year vision for Analytics in NSW Health. January 2016 http://www.ehealth.nsw.gov.au/__data/assets/pdf_file/0020/303752/NSW_Health_Analytics_Framework.pdf