Flying Blind

Australian Researchers and Digital Health Australian Health Data Series: Volume 2

September 2018



CMCRC
ASSURING MARKET QUALITY

digital health crc

RESEARCH AUSTRALIA

© CMCRC and Digital Health Cooperative Research Centre.

The CMCRC and Digital Health Cooperative Research Centre provide though leadership, innovative research and break-through technology solutions for financial and health markets.

About this report:
This work has been produced by the ***Health Market Quality*** **Research Program** of CMCRC and the **Digital Health Cooperative Research Centre** as part of the *Health Data Series*. **Volume 2 of the Australian Health Data Series** focusses on Australian health researchers and their ability to access and use health and medical research data in a complex regulatory environment.

Authors: Uma Srinivasan, Divya Ramachandran, Cassie Quilty, Suresh Rao, Michael Nolan

Art Work and Graphics: Chris Taylor

Suggested Citation:
Srinivasan, U., Ramachandran, D., Quilty, C., Rao, S., and Nolan, M. 2018, *Flying Blind: Australian Researchers and Digital Health, Volume 2: Health Data Series,* Digital Health Cooperative Research Centre, Sydney.

ISBN: 978-0-646-99212-9

Contact:
Dr Uma Srinivasan
Digital Health Cooperative Research Centre
Email: umas@digitalhealthcrc.com

Acknowledgements:
The DHCRC acknowledges the support of the Australian Government's Cooperative Research Centres Program.

FlyingBlind

Australian Researchers and Digital Health
Australian Health Data Series: Volume 2

August 2018

Authors: Uma Srinivasan, Divya Ramachandran, Cassie Quilty, Suresh Rao, Michael Nolan

TABLE OF CONTENTS

EXECUTIVE SUMMARY 5

1 INTRODUCTION 6
1.1 BACKGROUND 6
1.2 THE AIMS AND OBJECTIVES OF VOLUME 2 6
1.3 OUR APPROACH 7

2 HEALTH RESEARCH IN THE AUSTRALIAN CONTEXT 8
2.1 AUSTRALIAN HEALTH AND MEDICAL RESEARCH 8
Australia's Health System 8
2.2 WHAT CAN AUSTRALIAN HMR LOOK LIKE? 9

3 BARRIERS TO AUSTRALIAN HEALTH AND MEDICAL RESEARCH 12
3.1 COMPLEX PROCESSES 12
Healthcare Agencies 12
Data Stewards and Data Custodians 13
Human Research Ethics Committees (HRECs) 14
Data Linkage Units and Secure Infrastructure Providers 14
Obtaining Consent for Research 15
Applying for Funding 15
3.2 DATA GOVERNANCE 16
Technology 16
Data Collection 16
Data Preparation 17
Data Incentives 18
Data After Research 18
3.3 LEGISLATION AND POLICIES REGARDING DATA USAGE 20
Privacy Legislation and Policies 20
Data Custodians: A Complex Policy Environment 21
Human Research Ethics Committee Policies 21
Funding Policies 21
CASE STUDY: MEDICATION SAFETY 23

4 HEALTH AND MEDICAL RESEARCH: LESSONS FROM OVERSEAS 25
4.1 THE UK 25
NHS Digital: A Data Curating Entity and a Digital Health Strategy with Teeth 25
A Research Coordinating Body 25
Major Datasets: The UK Biobank 26
4.2 THE USA 27
The Value of Data 27
HIPAA: Clear Legislation 27
The IRIS Registry: Best Practice and Research 28
4.3 OTHER COUNTRIES 29
Singapore 29
Finland and Estonia 29
4.4 LESSONS FOR AUSTRALIA 30
Governance Principles for Data Preparation, Management and Release 30
Privacy, Security, Confidentiality by Design 32
Data Access Protocols 33
Data after Research 33
Legislation and Policy 33

5 CONCLUSIONS AND CALL TO ACTION 34
CALL TO ACTION 35

6 APPENDICES 36
GLOSSARY OF TERMS 36
NATIONAL LEGISLATIONS 38

7 REFERENCES 41

EXECUTIVE SUMMARY

Australia holds high quality digital health data that are well coded and structured to be of incredible value to health and medical researchers. In addition to well organised structured data sets, there are vast amounts of untapped resources in the form of unstructured data such as texts, images, audios, and other digital data streams from a range of personal devices and monitors, much of which are unused.

In spite of the abundance of digital data, Australian health and medical researchers spend several months and even years to assemble data required for their research.

Research Australia's annual consumer surveys demonstrate that Australian consumers are willing to share their health data to support research. However, this is not reflected in the current restrictive environment where Australian health and medical researchers face a myriad of problems as they navigate a complex environment enmeshed in legislative, ethics and other barriers around data accessibility for research. Very often these obstacles result in long delays where research funding almost runs out, forcing many researchers to abandon linked data studies and make do with small data sets or seek overseas data banks to address their research questions.

Delays are caused by a range of factors. The fragmentation of health services delivery across primary, secondary, hospital and allied healthcare settings produces an equally fragmented data environment. Further barriers are embedded in a health and medical research ecosystem comprised of complex funding and ethics approval processes, ad hoc policies and data governance strategies that differ across state and federal data custodians. These processes and policies lack consistency and are often not transparent to researchers, causing inordinate delays in getting necessary approvals and access to health and medical research data.

The recent public response to the implementation of My Health Record (MHR) is an indication of a lack of consumer trust with the way the authorities plan to implement the MHR. This lack of trust should not be confused with an actual breakdown in technologies that deal with privacy and security of data, access controls and safeguards around data protection.

Ensuring transparency and clarity around data policies and processes is vital to building a trusted environment for the Health and Medical Research (HMR) sector to deliver real value to Australian healthcare consumers.

Consumer driven transformations that have occurred in other sectors have not occurred in health. We are flying blind. Clinicians practice without knowing the full context of their patients or their outcomes. Care is generic rather than personalised. Funders pay for interventions of unproven or negative value. Consumers lack access to solutions and information that are commonplace in other aspects of their lives.

Expediting access to HMR data is a real and urgent issue. Australia needs to embrace a nationally consistent, streamlined approach that embeds privacy, security and confidentiality by design. The privacy by design approach needs to encompass the use of data as well as data products such as analytical and predictive models that have the capacity to support preventive and precision healthcare.

As part of background research for this report, we reviewed the research environment and the best practices of several countries around the world including USA and UK. These have guided us in presenting a vision of a twenty-first century health and medical research environment. Our vision combines privacy protection, data security and streamlined governance processes to create a research environment capable of supporting our future health system. HMR has a vital role in helping shift the focus from reactive, acute care to proactive and preventive healthcare that can be translated into policy and practice without inordinate delays.

1 INTRODUCTION

1.1 Background

The Digital Health CRC builds on the foundational work of CMCRC's Health Market Quality R&D program, broadening scope to the entire digital health environment. The aim is to ensure Australia can realise the full potential of digital health to support connected and evidence-informed healthcare. This goal requires an environment that enables research excellence and efficiency, as well as best-practice use of the valuable data resources that Australia possesses.

The R&D program of the DHCRC will focus on enhancing the deployment and use of existing and emerging technologies on all forms of structured and unstructured data in a privacy-friendly, ethical manner. The aim is to improve health across all settings of care through providing, where possible, real-time decision support to consumers, service providers, clinicians and those charged with planning, regulating, funding and managing the system.

The R&D program will deliver national impact through four interlocking research programs:

- Enabling Information Discovery and Application
- Identifying and Managing Health Risk
- Better Value, Quality, Access & Safety
- Consumer Empowerment and Positive Behaviour.

1.2 The Aims and Objectives of Volume 2

Our first report, ***Flying Blind: Australian Consumers and Digital Health*** ('FB1') highlighted the fragmented nature of Australian consumer health data (1). The present volume, Australian Researchers and Digital Health ('FB2') turns its attention to Australian researchers and the access and use of data for health and medical research (HMR). HMR is vital to supporting and sustaining our health system. Australia has a wealth of health data resources, many of which are originally collected for other purposes such as administration or compliance. With appropriate access to these data and through data linkage, health researchers can generate new insights, uncover new trends and deepen our understanding of health and disease. In FB2, the aim is to understand how well these national data assets are used for research and where barriers may exist to more effective use.

This report explores:

- The nature of HMR data, as well as how it is collected, prepared and made available
- The processes researchers have to go through in their attempts to gain access to this data
- The policies and governance structures that make up Australia's wider HMR regime.

Australia's fragmented, piece-meal data environment creates impediments to important research that can enhance the quality of Australian healthcare and improve the health and wellbeing of every Australian. Australia needs to embrace a nationally consistent, streamlined approach to data preparation, access and release that provides clarity and simplicity for all involved, maximises opportunities for HMR, and, most importantly, minimises the time between research and translation into better policy and improved clinical practice.

1.3 Our Approach

In conducting this study, our approach has been to:

- Review the current HMR ecosystem and the possibilities afforded by data-informed HMR research
- Examine the policies and governance frameworks surrounding the collection, preparation and use of HMR data
- Investigate the practices and frameworks surrounding the linkage of HMR datasets
- Explore the methods adopted by ethics committees and data custodians in granting access to existing digital health data for HMR
- Discuss policies concerning research data reuse for subsequent research projects
- Present a vision of a twenty-first century HMR environment which combines privacy protection, data security and streamlined governance processes to enable Australian researchers to advance innovation and transform the way in which healthcare services are delivered.

In the process of writing this report, the team has also published a series of blog posts on the current HMR environment to spur discussion in the Australian research community (2). As such, this volume builds on the feedback generated by these posts as well as additional research to construct a series of recommendations aimed at supporting a national culture of HMR data confidence and abundance.

2 HEALTH RESEARCH IN THE AUSTRALIAN CONTEXT

This section serves as an introduction to the importance of research in Australia's health and medical ecosystem and in addition it also discusses the centrality of ***data*** in health and medical research and the opportunities afforded by better access to HMR datasets.

2.1 Australian Health and Medical Research

Australian health and medical research is world-class. We outperform in research output – ranking in the top 10 OECD countries – despite stagnant and below average expenditure on R&D as a percentage of GDP (3). As other countries increase their investment in HMR, Australian health researchers must do more with less just to keep up. This research delivers a range of benefits to our country such as new scientific discoveries, medicines, procedures, and improvements to service delivery, as well as life-changing innovations such as cochlear implants and cervical cancer vaccines. The true value of HMR is difficult to quantify. Each dollar spent on HMR produces direct and indirect benefits to the health and wellbeing of our population and the health our economy; one conservative estimate indicates at least $2.17 returned for every dollar invested (4). Despite this, there is scope for ***major*** improvement to Australia's HMR ecosystem, especially when it comes to how HMR data is made available for research use. Data is one of the most underutilised resources in Australian HMR and researchers face many barriers and delays along their research journey as they navigate policies and processes to access data they need.

Australia's Health System

Australia's health system faces significant challenges including an ageing population, increasing costs of medical interventions, changing expectations from the community and higher incidence of complex and chronic disease (5). Health spending as a proportion of GDP has been rising consistently over the past twenty years: in 2015-16 Australia spent more than 10% of GPD on health for the first time in history (6). Health consumers must also supplement costs out-of-pocket for many public and private health services, despite not having access to transparent information and decision-support about varying efficacy, quality and costs of these services. An example being instances of 'low-value' healthcare, defined as care that offers no benefit to patients or benefits outweighed by risks including being "inappropriate for a specific clinical indication in a specific population or an excessive frequency of services relative to expected benefit" (3). 'Low-value' care can be hard to define because patients have varying needs and preferences. While difficult to measure directly, the high prevalence of these services is well-documented in high-income countries and can be measured indirectly through 'unwarranted variations' in the geographical spread and intensity of procedures and services (4). Some estimates indicate up to 30% of all services delivered may fall in this category. As such, reducing the incidence of this 'low-value' care will see benefits to our health system through improved patient safety as well as rationalised costs.

Further challenges to our health system stem from its traditional focus on reactive, acute care in a hospital setting. Over time, our disease burden has shifted toward chronic conditions, often with multiple co-morbidities (9) that require out-of-hospital management. Over 10% of Australians live with asthma, 7% with high cholesterol and 5% with diabetes and the leading causes of death are heart disease, dementia and cancer (10). Moreover, 4 million Australians – nearly 20% of our population – have experience living with mental or behavioural conditions (11). As a result, there is a need to shift the health paradigm to a more proactive, prevention-focused system that can deliver quality care and management of health and disease in the community.

2.2 What Can Australian HMR Look Like?

Reducing the strain on Australia's healthcare services requires finding novel approaches to sustainable healthcare delivery. Key to this is investing pro-actively in initiatives that lessen demand by promoting wellbeing and by keeping people healthy for as long as possible. Australia needs a healthcare environment that predicts, prevents and delays the onset of chronic and long-term diseases, eliminates low-value care and other forms of waste, and constantly scrutinises pharmaceuticals and medical devices to ensure that they are functioning in the best interests of patients. It is urgent that we take steps toward building a health system based on ***prevention-based, precision-personal medicine***: one that places the promotion of the health and wellbeing of Australian citizens at the very centre of everything that it does. Research is vital to this endeavour.

The World Health Organisation argues that using evidence-based research to improve service delivery is one of the most important ways to strengthen health systems (5). What health research looks like has been changing over the past fifteen years. Advances in data analytics and computing power mean that large datasets are an important resource in the research process. These are often datasets that are ***routinely collected*** for another purpose – e.g. for healthcare claims administration – but contain enough detail for researchers to use for valuable ***secondary analysis (6)***. These datasets are often ***large***, ***pre-existing***, and, ***longitudinal***. As such, it is possible to use these datasets to understand trends, patterns and correlations at a large scale and investigate range of questions much more rapidly and cost-effectively than using methods such as surveys and clinical trials alone. It is also likely that ***unstructured*** data sets – e.g. images, free text, Internet of Things data and audio data – will prove of similar value in the future.

However, according to Australia's Population Health and Research Network, 'no single data collection' is sufficient to 'allow an understanding of the complex pathways that result in health or disease' or to determine 'whether Australia's health and social systems work in optimal ways' (7). The more datasets that researchers are able to ***link*** together, the more complex and ***valuable*** are the questions that they can hope to solve.

Linking data refers to process of matching different pieces of data that are thought to relate to the same person, family, place or event to form 'a new, richer dataset' (8). The Commonwealth Department of Health has noted that data linkage promises 'a long list of significant benefits to the health system', by making possible:

- Better information to inform government policy decisions
- A better understanding of what works, how well, for what cost, and in what circumstances
- A more efficient health system, by supporting the most cost-effective treatments, strategies and interventions on broad-based independent evidence (9).

Supporting data-informed research using routinely de-identified, ***pre-linked*** or ***linkage ready*** HMR datasets and unstructured data is an important part of improving Australia's health. Potential benefits can also be realised through linkage of survey and clinical trial data with non-HMR datasets such as those relating to education or labour status.

Linking numerous different sources of data can help identify the complex interplay of factors that lead to different diseases. This information can be used to predict the onset of morbidities in individuals and provide clinicians with the ability to perform targeted interventions. Such approaches can lead to better health, improving patient quality of life, and reducing demand for health services over their lifetime. ***Figure 1*** illustrates a virtuous research cycle that can be integrated into policy and services delivery. This can be particularly valuable in the context of chronic diseases which can develop ***slowly*** and without explicit symptoms with patients often unaware that their condition is developing until it is too late to prevent it.

Using datasets to predict chronic disease

The use of large scale data holdings to understand disease progression offers an alternative and considerably more cost-effective method for predicting the onset of chronic morbidities. To take one example, researchers from the University of Sydney and the CMCRC have used four years of private health insurance data to examine the progression of Type-2 Diabetes (T2D) and have been able to find a series of common comorbidities that indicated the onset of T2D before it had manifested: something allowing them to identify individuals at risk of T2D with 86% accuracy (10). Building on this research will give clinicians the means to make predictive interventions long before the initial symptoms of a chronic disease manifest and seriously improve quality of life.

Similar opportunities await in tracking the efficacy of different therapeutic techniques. Linking datasets will allow researchers to understand the value of care in ***real-time*** and better determine which interventions benefit patients and which are little better than placebos. For a specific procedure, you can even identify sub-groups of patients for whom it is likely to be less effective. For instance, knee arthroscopies are now commonly understood to offer no net befit to individuals over the age of 50 and are of questionable value in other circumstances – and yet the number of knee arthroscopies performed on individuals over 50 in private hospitals grew 14% between 2002 and 2014 (11). Clinicians can use this expanding evidence base to tailor interventions directly to patients in a form of ***precision medicine***.

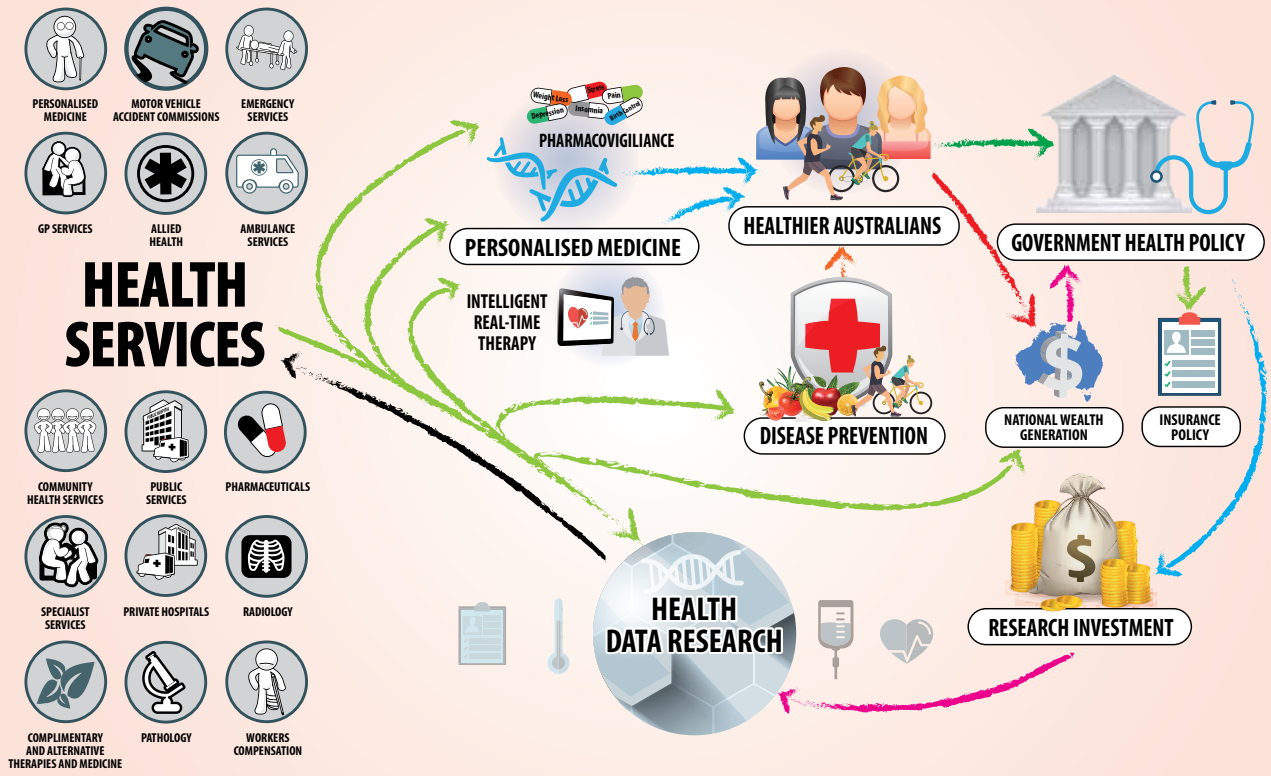


Figure 1: Virtuous Research Cycle

This process also promises to help disseminate best practice throughout the health system in ***real time***. The US National Institute of Medicine has stated that 'growing computational capabilities to generate, communicate and apply new knowledge' hold in them the 'potential to build a clinical data infrastructure to support continuous learning and improvement in health care' (12). The latest results of research and the real-time outcomes of different therapies will be instantly accessible for clinicians, allowing them to make informed decisions as to the best method of treatment for patients and discontinue therapies that have been shown to be ineffective. Policymakers will likewise be able to formulate healthcare policy based on real, 'living' treatment outcomes. The American Academy of Ophthalmology's IRIS Registry is a good example of movements in this direction (see Section 4).

Many more uses could be mentioned: research datasets hold important potential in dealing with outbreaks of acute disease or other urgent public health issues, identifying the sources of epidemics and best methods of treatment, and, for example, in anticipating and combatting antimicrobial resistance; they can also help end information asymmetries in Australia's health markets, empower consumers and reduce low-value care. Not only that, but it is likely that the functions associated with research data analysis will become increasingly automated over time. As technology advances and researchers continue to integrate methods from computer science and machine learning into their analytical approaches, they will develop algorithms that automatically trawl linked data for patterns, correlations and non-intuitive relationships for humans to investigate further. This can lead to evolutions in our understanding of how diseases progress and how effectively our health system is functioning.

The above represents a sampling of the opportunities afforded by better and more efficient access to large datasets. Data has a central role to play in moving Australian healthcare in the twenty-first century and beyond. The more sources of data that researchers are able to link to one another, the better. If we create an environment where research outcomes can seamlessly flow into health service delivery, it will change the way we deliver healthcare: away from expensive, ***curative*** care and toward ***predicting and preventing*** through precision medicine. Freeing up our data for research is of acute importance in the effort to radically improve our health environment – as well as in maintaining our research competitiveness, economic productivity, and, most importantly, if we are serious about improving the wellbeing of all Australians. Other countries have recognised these opportunities and are taking ***concrete*** steps to ensure that their health data is made accessible to researchers (see Section Four below). We cannot be left behind.

Advances in data analytics and computing power have precipitated a revolution in health and medical research. Large data sets that are often collected for another purpose are becoming increasingly important to the research process.

Routinely tapping into de-identified linked HMR datasets can go a long way toward improving Australian healthcare. The seamless flow of research outcomes into policy and practice will shift help shift from a curative to a preventative health model.

This is not an issue to be dealt with in the future: facilitating ease of access to linked datasets is a matter of national urgency.

3 BARRIERS TO AUSTRALIAN HEALTH AND MEDICAL RESEARCH

Australia is a federated entity consisting of multiple levels of government (Commonwealth, State/Territory and local) (13), and this impacts how both healthcare delivery and data are organised, funded, regulated and governed. In FB1 (1) we discussed how accidents of history led to the development of Australia's fragmented healthcare environment, in which primary, secondary and allied healthcare is split across public and private lines and overseen by health departments at different levels and across different geographies. As such, Australian healthcare has major coordination challenges. Responsibilities for health research are also dispersed across several different entities, including government departments, hospitals, universities, independent medical research institutions as well as for profit and not-for profit institutions.

This fragmentation has produced data ***silos*** which have major impacts on Australian HMR. Data access requests, data governance and policies also reflect the ***ad-hoc*** and ***organic*** manner in which Australia's healthcare environment has developed. There is a lack of ***consistency*** and ***clarity; elements*** required to provide confidence to all of those operating in HMR. Indeed, major challenges facing Australian HMR can be broadly categorised as:

- Inconsistent data access request protocols, duplication and wasted time
- Outdated data management, release and preparation frameworks slowing data release
- Complex legislative environments and funding policies affecting release of data and researcher incentives.

This section details the range of different processes and players encountered by researchers in their efforts to attain HMR data: processes that are neither uniform nor transparent and that lead to duplication, wasted time, and harm. While many of these barriers are ***systemic*** in nature, their impacts are felt by researchers daily.

3.1 Complex Processes

Inefficient and overly complex data access processes are common in Australia's current HMR environment. One submission to the Productivity Commission's recent ***Inquiry into Data Availability and Use***, describes the current state of Australia's HMR data access processes as an unnecessary 'barrier to [healthcare] progress' (14). This is because of inconsistencies, differing processes, duplication and wasted time. These processes require rationalisation, streamlining and harmonisation to better support HMR.

Researchers expend considerable amounts of time and energy attempting to find the datasets they need before applying for time-restricted grants from different funders. They must then negotiate with data custodians and data linkage units, who can be extremely cautious in providing researchers with access to health data. Researchers must also navigate numerous ethics committees while remaining mindful of Australia's similarly complex privacy and data-use legislation. Navigating these processes can take ***years*** and is an inefficient use of taxpayer-funded grant money. Even worse, it delays valuable research supporting and improving our health and wellbeing. Some of the challenges in the health research journey are illustrated in *Figure 2*.

Healthcare Agencies

A major barrier to data-enabled HMR is the fact that data is collected and held by a huge number of different agencies. Before applying for data, researchers must understand ***where*** the data they wish to source for their projects is located. Unfortunately, Australia has no single body coordinating the management and preparation of healthcare datasets. Thus, researchers often have to attempt to source data from:

- State or Federal Government health departments or ministries
- Local Health Districts and Population Health Networks
- Hospitals (both public and private)
- Clinical Registries
- Other contracted agencies performing health services.

Most of these agencies have their own separate processes surrounding the access and linkage of data by researchers working in HMR, and this fragmentation is one of ***the*** major sources of much of the wasted time and effort experienced by those attempting to conduct HMR. Each agency tends to come with its own Data Custodians, ethics committees and data linkage arrangements – and as such, its own differing but frequently duplicative data access request and release processes. Researchers must spend considerable amounts of time liaising and coordinating with these different parties in their efforts to secure data. Considering the skills and training of health researchers, this is a low value use of their time.



Figure 2: Researcher's Journey

Data Stewards and Data Custodians

The head of a healthcare agency is usually responsible for data collected by the organisation and has overall responsibility for protecting data and approving conditions or guidelines for its use and disclosure (15). These 'Data Stewards' may nominate one or more Data Custodians, who are responsible for day-to-day management, operations and support of data collections – a function which includes reviewing access requests for research. When applying for data access, researchers need to outline their research project and identify the datasets required for their study so that Data Custodians can assess the feasibility of releasing data. Once Custodians agree in principal, researchers usually have to obtain Ethics Committee approval. Custodians only will contemplate granting access to data when this approval is granted.

However, Data Custodians from different jurisdictions and different agencies have their own processes that need to be followed by researchers for making data access requests, and the lack of a nationally-standardised data access request protocol creates considerable delay for all parties involved.

Anecdotal evidence suggests that often the initial request to a Custodian becomes a lengthy and complicated endeavour: Custodians generally require researchers to describe every individual data item and develop comprehensive data dictionaries (listing of all fields in a dataset) with lists of variables. that researchers then need to justify the need for each individual data item (16). In the case of data linkage projects that involve multiple datasets and custodians, this process can add several months – if not years – to the research timeline.

Adding to this difficulty is the fact that Data Custodians are not incentivised to support research: while they have a range of rights and responsibilities, getting data to researchers is rarely a primary part of their role. Custodians are also presented with a complex and piece-meal legislative and regulatory framework that further reduces their willingness to release data. As we shall discuss below, making data available and maximising the research value of data holdings must become a core responsibility of data custodians and needs to be incentivised as such.

Human Research Ethics Committees (HRECs)

Ethics approvals are another core part of the data access and linkage request process. HRECs review research proposals that involve human participants to ensure they are ethically acceptable and in accordance with relevant standards and guidelines, such as the National Statement of Ethical Conduct in Human Research issued by the NHMRC (17). HRECs are established by a range of organisations – whether public, private, or not for profit – and are most commonly found in universities and hospitals. Federal and state-based health agencies have their own ethics committees to assess research proposals requiring access to their data too.

However, HRECs present researchers with some difficulties. Where research involves multiple datasets, researchers often need to apply ***individually*** to each HREC involved and obtain clearance from each one. Each HREC usually has its own request form, and these are notoriously long; they can be even longer if specific populations (such as Indigenous Australians) are involved. Not only that, but the vast majority of the information sought in each is duplicated: most ethics committees do not recognise approvals granted from other HRECs.

Attempts have been made to alleviate this problem. For instance, the NHMRC established the National Mutual Agreement (NMA), formerly known as Harmonisation of Multi-Centre Ethical Review in 2006 (18). The NMA aims to streamline approvals from multiple HRECs by stating that if an ethics application is approved by an NHMRC-certified HREC, this will satisfy HRECs at all other institutions in which a researcher is applying for data – and thus do away with the need for more ethics approvals. Unfortunately, a decade down the line progress has been slow. The NMA contains a large number of exemption clauses and exclusions that impact data-linkage studies; nor has acceptance of the NMA been mandated. Furthermore, not all HRECs are certified by the NHMRC. Other institutions such as the PHRN have also begun to offer specialised training about data linkage for HRECs, but this training is optional (19).

Ultimately, ethics approvals are designed to ensure that research does not harm participants – but when they lead to excessive delays in the provision of data for HMR, it may well be that they are acting ***against*** their mandate by delaying research that can actively benefit Australian health and wellbeing. In fact, researchers often find themselves in a paradox when seeking approvals from Data Custodians and HRECs: most HRECs demand data-related information before providing any sort of clearance, when at the same time Data Custodians are often unwilling to release the very same information without requisite ethics approvals. This process is only resolved after extensive back-and-forth between committee and custodian – with much wasted time on the part of researchers.

Data Linkage Units and Secure Infrastructure Providers

Should researchers wish to link data, they are then required to contact Data Linkage Units (DLUs) and Secure Infrastructure Providers (SIPs). DLUs are found in every state and territory and are responsible for creating linkage IDs or 'keys' that facilitate the connection of data across different health agencies and jurisdictions (20). Once a research project is approved, Data Custodians work with DLUs to determine which records and minimum content information is required for ***each specific project***. DLUs then create project specific linkage-IDs that are sent to Data Custodians who then extract required records, replace personal information with the project-specific linkage ID, and provide the data to the researchers. Researchers then use the project-specific linkage ID and content information from multiple custodians to merge data without access to any personal information.

SIPs are independent organisations who provide secure infrastructure to hold linked datasets (21, 22). Most of Data Custodians and DLUs work with preferred SIPs, and usually require researchers to use their preferred SIP for both storing and analysing data over the duration of the research project. There are only two Australian SIPs that offer remote-access data storage and analysis capabilities for researchers. The first is SURE (Secure Research Environment) which was established with funding from the Federal National Collaborative Research Infrastructure Strategy (NCRIS) as part of the PHRN and is currently run by the Sax Institute (21). The other, DataLab (22), is run by the Australian Bureau of Statistics – but virtual access is available solely to researchers from federal and state agencies. University researchers must apply for on-site access.

However, it is researchers who are generally responsible for coordinating data projects with DLUs and SIPs. Again, this is something that adds significant – and unnecessary – amounts of time and effort to the research process. Discussions with DLUs can only be initiated after approval from HRECs and Custodians has been granted, and DLUs tend to charge significant fees for providing project-specific linkage IDs. SIPs come with their own registration and training requirements as well as hefty access fees (23). The complexity of the processes involved also means that any change in the research project, such as the addition of a new team member, can often result in the need for amendments and/or fresh approvals from all HRECs, Data Custodians, DLUs and SIPs involved in the project. Likewise, the fact that data is not ***pre-prepared*** or made ***proactively linkage ready*** means that the process above must be ***repeated every time researchers apply for data*** – something explained in more detail below.

Obtaining Consent for Research

Obtaining consent for research represents another area of ambiguity for researchers. If researchers are using de-identified data, then they generally do not need to gain the consent of individuals being studied. If individually identifiable data is to be used in a research project, a signed consent form is usually required from all participants taking part in the study (although there are some situations in which this is not the case) and participants must be provided with adequate information about the research project and about how their data will be used.

However, if researchers wish to link individually identifiable data generated by a survey or other bespoke dataset to data from another source that pertains to the individuals being studied, they must receive consent for each data set they wish to link – a requirement which adds unnecessary complexity to the research process and can often leave participants confused and unduly concerned about how their data is going to be used. For instance, if a researcher wanted to link the survey data of 500 individuals with MBS and PBS records, they would need to attain consent 1500 times – one for original data, one for the MBS, and one for the PBS. Not only that, but if researchers are applying to more than one ethics committee the process can duplicate: they might have to ask study participants for their consent *again*.

Applying for Funding

Researchers are also required to apply for funding from different agencies, and in the process, compete with one another for a small number of time-limited grants. There is no question that this promotes considerable inefficiency in the research process. According to one researcher,

> "When at least one quarter of our jobs are spent writing competitive grants and less than 20 percent of those applications are being funded, it's clear that we need to cut some red tape, so we can free up time for researchers to get their actual work done (24)."

As we shall see below, this does not simply slow the pace of research, but in the process, orients researchers to the kinds of shorter-term publishable projects that are of less benefit to the health and wellbeing of Australians than larger, longitudinal studies. The nature of the competitive grants process also means that researchers are often less willing to share their data and research with others – a flaw which slows the pace of HMR and effectively conflicts with the essential spirit of conducting scientific research. *Figure 3* illustrates the current complexities in funding.

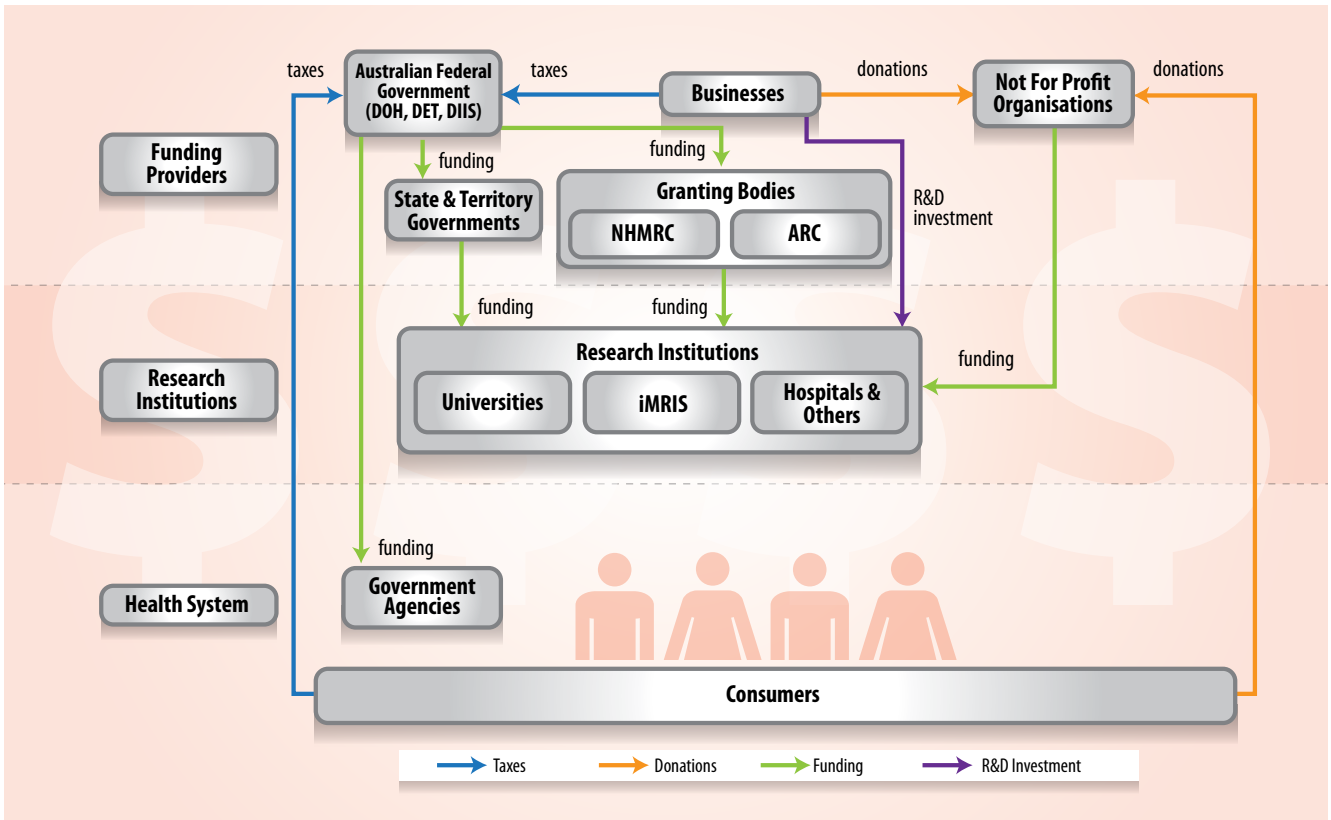


Figure 3: Complex funding Arrangements

3.2 Data Governance

In addition to the complexities of the research access request process we must also turn to the way in which data is collected, prepared, and governed by healthcare agencies. Indeed, while Australia is very good at ***collecting*** data, we often fail to ***use*** it (25). As discussed, this issue is linked to the fragmentation of the healthcare system; there are many data silos each supported by a range of approaches to data governance (1).

The absence of clear and consistent principles regarding the preparation of data for secondary uses provides considerable difficulty for all involved. Data is rarely made research or linkage ready prior to requests for usage, and nor do Custodians or healthcare agencies see these tasks as part of their core functions. The same extends to data ***after*** research: datasets are often deleted or left to lie defunct. Australia's haphazard approaches to data governance ensure a significant amount of unnecessarily wasted time, money, and effort – all of which could be spent saving lives and improving the health system. Taken together with the complexities of the research application process, these issues mean that Australian HMR researchers are often forced to plan their studies and answer questions based not on what is ***most needed*** or what they ***wish to understand***, but rather based on ***what data is available*** – and then, on top of this, conduct their studies with ***partially complete datasets*** of ***varying quality***. This is detrimental to the state of Australian HMR.

Technology

It is important to stress that the difficulties regarding data governance are rarely to do with technology. While there are some issues with data quality and interoperability, Australian health services data is generally well-coded and standardised, and governments have made substantial investments toward ensuring that the country is in a high state of technical readiness for data usage and linkage. Western Australia's Data Linkage Branch has contributed impressively to Australian research since its inception in 1995 (26), while the PHRN is now supported by a data-linkage agency in every state and territory (20). In fact, according to the OECD, Australia's health data collections are in a state of technological readiness for linkage and effective usage: it is the complex state of data governance that remains one of the major stumbling blocks preventing researchers from utilising the kinds of health data that can revolutionise health and medical care in this country (27).

Data Collection

Much of the root of Australia's data usage problems can be traced to the extreme fragmentation associated with the collection of health services data: an issue that was discussed at length in the first volume of Flying Blind. Multiple government agencies – and at times private, not-for-profit entities – are responsible for collecting the numerous different datasets generated because of the day-to-day activities of the health system. Many go to the trouble of creating their own unique sets of identifiers and come with their own rules and processes surrounding the collection of data. Likewise, while some agencies share their data routinely with other government departments or registries for the purposes of policy and evaluation, many are quite unwilling to do so with ***one another*** – let alone with researchers.

To take one example, cancer data is collected and held by a range of different public-sector organisations ***and*** NFPs across both Commonwealth and State/Territory levels, and each agency involved has its own collection processes as well as rules and regulations. At the same time, there is no overarching mandate for these organisations to share their data; access is only granted through piecemeal bilateral agreements (14).

Data Preparation

This does not simply mean that volumes of potentially useful health data remain unnecessarily siloed from one another, but also that, when researchers attempt to link data, various different collections are housed in different states of research and linkage readiness. The sheer number of organisations involved in collecting health services data brings with it inconsistency in the way in which data is *governed*, *housed*, and *prepared*.

Indeed, a common complaint amongst those involved in HMR is the fact that it is relatively unusual for Australian HMR data to be pro-actively made *research or linkage ready*. Datasets are rarely proactively de-identified and linked on a routine basis. Nor do many healthcare agencies have consistent rules about what data can be released and in what state of readiness data should remain. This means that when researchers or other interested parties request access to data, a whole range of ad-hoc governance processes regarding data access kick in at the same time. Unsurprisingly, these processes can differ quite drastically between healthcare agencies. In conjunction with HRECs, Data Custodians are charged with deciding *whether* data can be released, *what* data can be released, and in *what format* or at *what level* – for *every single request* that comes their way. It is only then that they begin preparing and, usually, de-identifying or aggregating data to release specifications, before working with DLU teams to create project specific linkage IDs.

The fact that this happens *each time* a request is granted and approved represents an unnecessary time burden on all parties involved, and, thanks to the inconsistency of health data governance rules, can mean that researchers receive data of differing quality and usefulness from different institutions. It also adds significant extra costs to all parties involved, as those responsible for data management are tasked with creating new datasets after access requests – often ones they have created before – from scratch or from very limited starting points.

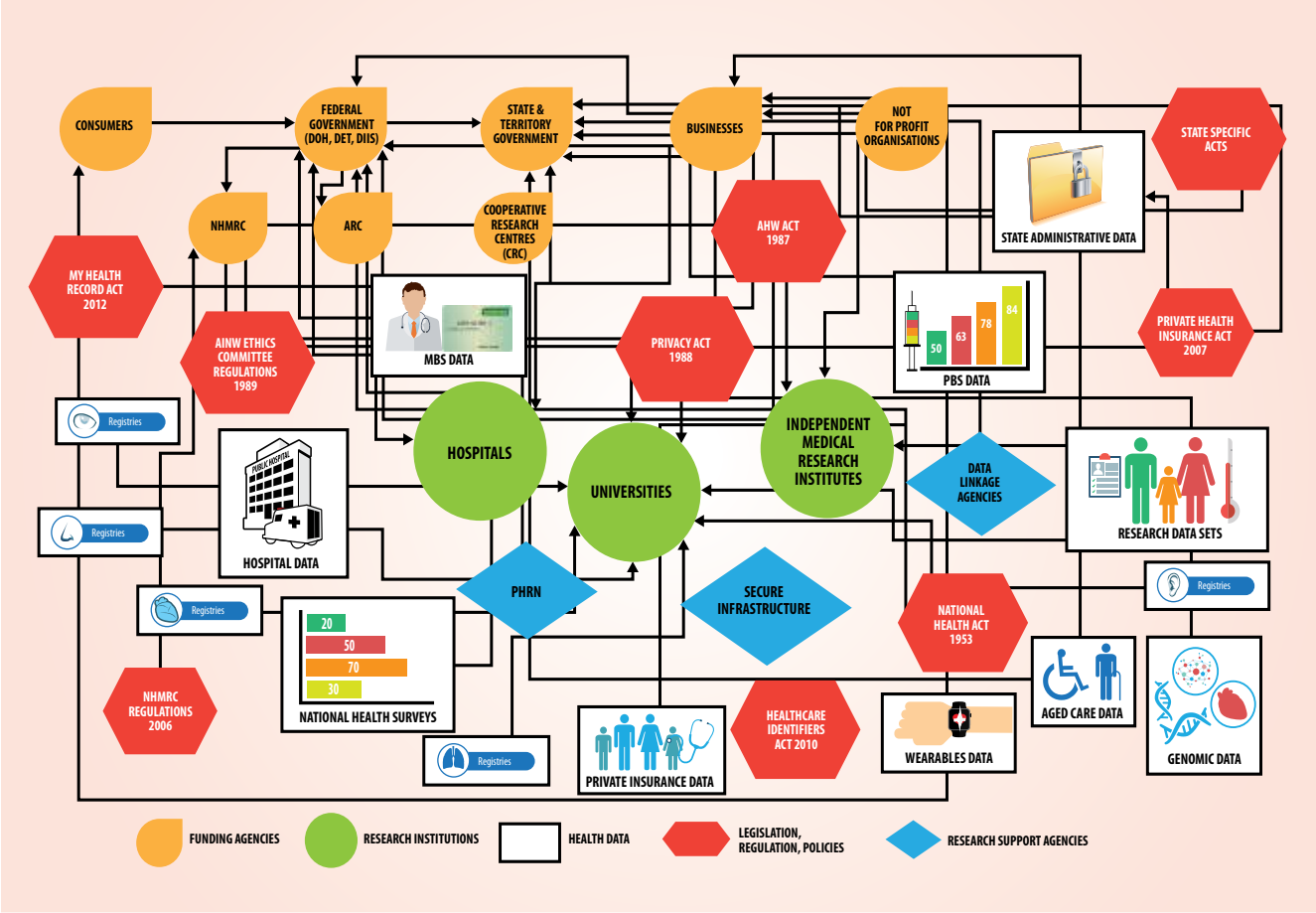


Figure 4: The Complex HMR Environment

As such, many agencies attempt to recover the costs of data preparation and linkage by charging researchers large amounts of money for their services, even when this simply amounts to the transfer of government funds.

De-identified datasets can be pre-linked or made ***research*** or ***linkage-ready*** and then held by agencies tasked with disseminating them for HMR and other useful endeavours, such as the real-time evaluation of policy. Doing so promises to rapidly expedite the speed of approvals and research, and, as such, the translation of findings into effective new directions in nearly every aspect of Australian healthcare. It is critical that Australia follows the example of forward-thinking countries overseas and adopts clear, streamlined approaches to data preparation that provide clarity to all involved and encourage researchers to spend more time on their core functions.

The complexity of the HMR environment is illustrated in *Figure 4*.

Case Study: Underutilised Registries

Clinical registries are defined by Monash University as databases that systematically collect health-related information within an overall governance and management structure on individuals who are (a) treated with a particular surgical procedure, (b) diagnosed with a particular illness or (c) managed via a specific healthcare resource (such as in an intensive care unit (28). Registries in many ways epitomise the dichotomy between Australia's excellent track record in collecting data and our relative failure to make effective use of it. Australian clinical registries represent an incredibly rich data source – and yet, due to the complexity of Australia's data governance arrangements, are not being used to their full potential.

As these data holdings are dispersed across so many different entities, researchers and other interested parties have to grapple with a wide array of processes, application protocols and ethics committee arrangements in attempting to gain access. Registries which are often funded through government grants sometimes charge large amounts of money for researchers wishing to access their data: something which means that researchers are paying for data which was created through tax payer funds.

Data Incentives

Despite the wealth of data collected by healthcare agencies, very few actively incentivise the release of datasets to researchers. Data Custodians are not judged or rewarded on the amount of data they have provided to researchers: in fact, few see enabling the provision of data as part of their core responsibilities. This, coupled with a complex legislative and policy environment (see **section 3.3**) contributes to a culture of data conservatism that leaves many Custodians disincentivised to release data or approve its linkage, even when the risks are clearly exceeded by the potential public benefits stemming from research. At the same time, Australian healthcare agencies are not required to budget or plan for the additional costs surrounding the preparation and linkage of data.

Data After Research

The lack of consistent policies regarding the re-use of linked data after research is unfortunate. Often the datasets researchers spend so much time and effort accessing and constructing remain unused and become ***defunct*** following the completion of the project. Several other developed nations encourage their researchers to make their de-identified datasets available for further research use, recognising the value of these data assets. However, it is still not mandatory for Australian researchers to make their datasets available to approved users at the completion of their projects (29). As many have spent lengthy amounts of time scraping and preparing these datasets, they are often understandably unwilling to share them with others. Likewise, whether linked datasets are retained varies between jurisdictions. Some state-based agencies retain linked datasets – or at the very least retain data linkage keys to facilitate easier re-linkage of data in the future – while datasets containing Commonwealth data are commonly destroyed at project completion. If a linked dataset uses federal administrative healthcare data (MBS or PBS), then this destruction is mandated (30, 31).

Thus, any researcher wishing to replicate a study or perform secondary analysis on a linked dataset will often find themselves having to navigate the same series of protocols and processes ***traversed by the original researchers***. This is deeply problematic: it drastically slows Australian research and limits the questions researchers can answer, and it limits the possibility of reproducing studies as researchers are already heavily incentivised to pursue 'new' research to establish careers and to succeed in the competitive grants process.

Traditional justifications for failing to promote the re-use of linked datasets is that individual privacy may be impinged upon if these datasets remain open to future use. This argument holds some merit if datasets hold identifiable data, but these can be subjected to de-identification to remove identifiable information. Modern privacy preserving techniques have advanced to the point where re-identification is difficult, and data security protocols provide an additional layer of protection. This is not to suggest that current de-identification and techniques are perfect. Rather, continued research and development in this area will serve to reduce the risk of re-dentification even further. Ultimately, these concerns need to be balanced against the multitude of public benefits that can flow from making linked datasets available and streamlining Australia's research environment more generally.

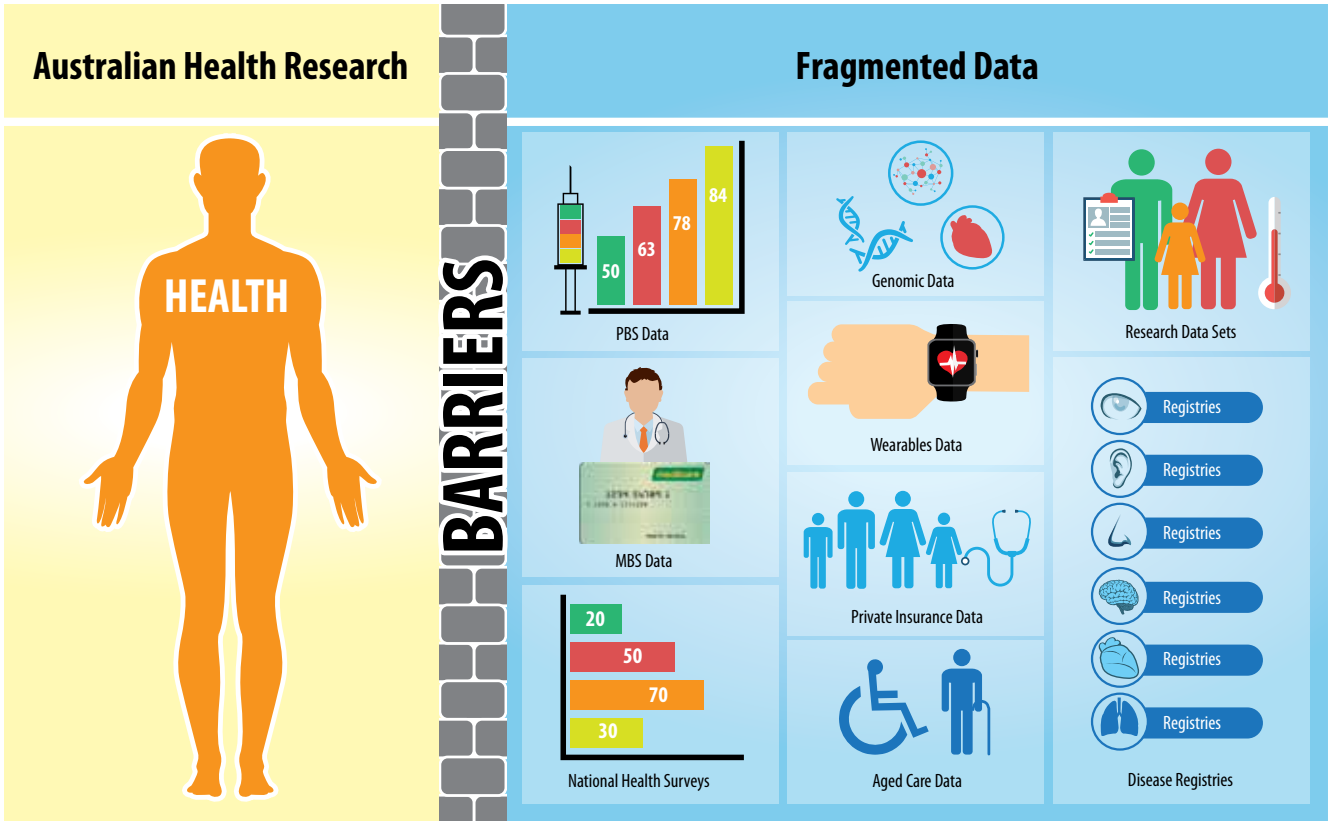


Figure 5: Structural Problems facing Australian Research

3.3 Legislation and Policies Regarding Data Usage

Amplifying the difficulties associated with accessing data for research is the fact that regulations, legislation and policies differ across levels of government, by jurisdictions, and between institutions – leading to serious inconsistency as to whether data will be made available, and at what level. Enabling data-driven HMR requires all stakeholders to ***unambiguously*** interpret and synthesise a range of policies around legislation around privacy, data management and ethical reviews.

Privacy Legislation and Policies

Legislation concerning privacy, data use and data collection is inconsistent between the States and Territories, which in turn differ from that of the Commonwealth. Likewise, a range of duties, exceptions and rules have been created by the intersection of privacy legislation, the statutes governing various government agencies and the common law (32). Set up originally in 1988 to protect personal information held by government agencies and to implement safeguards for collection and use of Tax File Numbers, the coverage and functions of the *Privacy Act* (Cth) have expanded to include the MBS and PBS, healthcare identifiers and more recently electronic health records (26). The patchwork of rules, regulations and guidelines around health data (*Figure 6*) is an immense challenge for anyone managing or working with data and has serious consequences for the release of data for research projects.

In fact, the recent public consultation paper regarding the secondary usage of My Health Record (MHR) data listed ***fifteen*** different pieces of Commonwealth legislation as relevant to the MHR and HMR more broadly (33).

They were:

- *My Health Records Act 2012*
- *Privacy Act 1988*
- *National Health Reform Act 2011*
- *Private Health Insurance Act 2007*
- *National Health Security Act 2007*
- *Australian Information Commissioner Act 2010*
- *National Health Act 1953*
- *Australian Bureau of Statistics Act 1975*
- *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*
- *Privacy Amendment (Private Sector) Act 2000.*

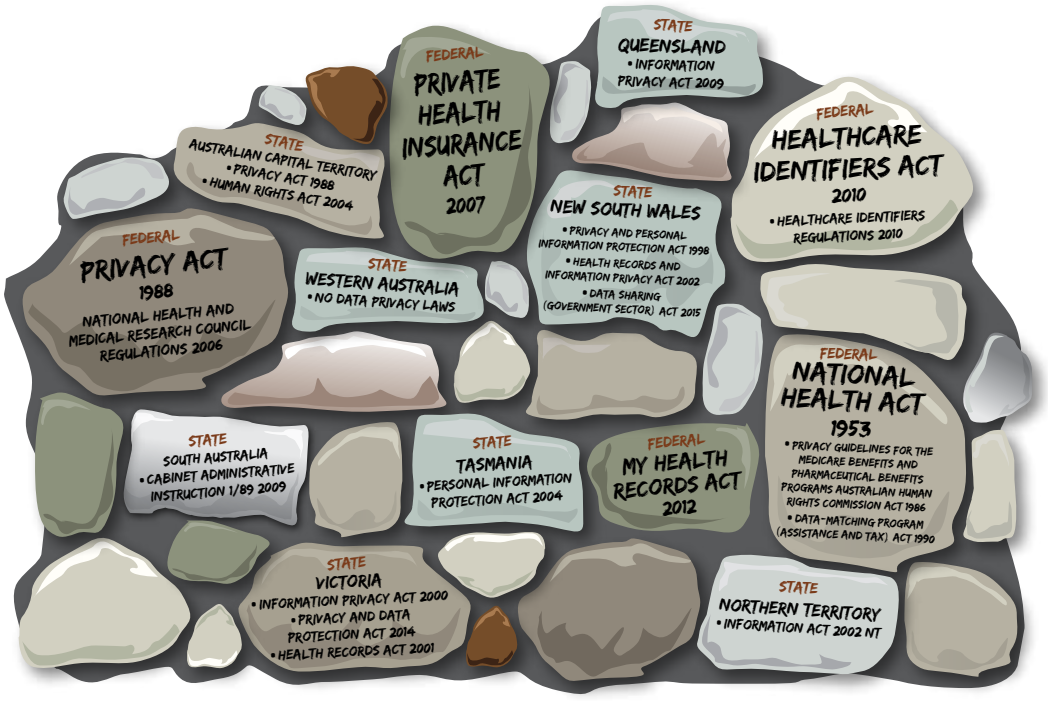


Figure 6: The Legislative Wall

It is important to stress that these pieces of legislation are at a *Commonwealth* level: there are a range of laws and regulations specific to each state as well. As a result, different data-holding agencies often have their own individual policies and codes of practice relating to the data they collect and share – a fact that drastically increases the level of duplication in the research application process and makes it incredibly difficult for researchers to know their chances of accessing the data they need before they make an application.

Data Custodians: A Complex Policy Environment

To truly enable data-driven HMR, stakeholders need to be able to interpret the legislative and policy environment in which they work in a manner that is both consistent and unambiguous (32). The Data Custodians who are primarily responsible for navigating the environment surrounding data release are personally accountable for the decisions they make, and it is no surprise then that the ambiguity and complexity involved in releasing data deters a considerable number from making data more available. Indeed, the recent Productivity Commission *Inquiry into Data Availability and Use* noted that:

> Too often public-sector data custodians and researchers are required to negotiate multiple pieces of legislation, many with criminal offences that apply strict liability for various actions… [this patchwork] forms a complex web, with intimidating consequences for missteps, that reduces the likelihood that data would be put to good use (34)

This problem cannot be overcome until the legislative and policy environment is standardised, harmonised and simplified as to provide clarity and confidence to those charged with its navigation.

Human Research Ethics Committee Policies

In a similar vein, inconsistent interpretation of data-sharing regulations and legislation on the part of HRECs leads to significant delays in the research project approval process. Ethics committees often have differing views when it comes to data sharing and how to handle data linkage studies. Many lack the experience with data usage or deep understandings of the possibilities inherent in data linkage – and this is amplified by the fact that, as we have seen, many HRECs do not accept approvals made by other committees.

Funding Policies

Several obstacles to efficient research are also related to existing funding models and mechanisms. The delays caused by the research application process outlined above are compounded by the fact that researchers are generally provided with time-limited grants, and so, when funding runs out, researchers must once again re-apply with no guarantee of continued funding – something that adds more unnecessary complexity to the research process as well as a great deal of stress for researchers. Researchers often also have to apply for support from multiple funders, many of whom have their own individualised application process. According to the 2016 *Medical Research Institute Sector Survey Report* (MRISSR) the complexity, time taken and bureaucracy involved in Australian HMR funding is 'holding everyone hostage' as researchers are forced to compete with one another from grant to grant (24). Of researchers surveyed, *93%* were of the opinion that current funding arrangements meant the sector was wasting significant amounts of time and money (24).

It is not just red tape which is the problem: but gaining access to data can be, in the words of the University of Melbourne, 'prohibitively expensive' – to the point that it 'consumes a disproportionate amount of research funding':

> One of the University's researchers in demographics and health reports allocating a significant proportion of ARC grant funds on accessing the base data for her research; this includes $60,000 in a single transaction to [a] Victorian Registry for access to 6000 births, deaths and marriages certificates. Over the past 17 years, the same researcher has allocated more than $250,000 in research funds toward purchasing registration certificates from Victoria and NSW that are vital to her work (35).

This situation is even more absurd considering the fact that this research is government funded – these transactions are nothing more than a transfer of public monies between different agencies.

Most critically, the issues surrounding funding mean that Australia's HMR environment is not as conducive as it could be to the kinds of large scale or long-term research – such as longitudinal or epidemiological studies – that are so important for healthcare. This is rendered even more problematic by the fact that in the present age of quantified research 'impact measurement' researchers are effectively required to 'publish or perish' if they wish to continue to access funding and advance their careers (36). In fact, *92%* of researchers responding to the MRISSR agreed or strongly agreed that the limited nature of funding seriously impacted their job security, and nearly *80%* said they had considered leaving the HMR sector because of this (24). In 2010 just 22.8% of applicants were successful in their attempts to gain NHMRC project grant funding, and by 2016, this had dropped to just 15.1%; difficulties securing funding mean that in 2016, one in four HMR scientists were unsure whether they would be employed in the following year (37). It is no wonder then that:

> The lack of funding is leading to an overall depressed feeling in institutes. Senior researchers are wondering what they will do, and students are wondering why they even started their PhD. Such an environment does not lead to best performance or enhanced creativity. *We are all just getting by and we can do so much better than that* (24).

This is detrimental to the quality of Australian HMR because researchers may err in favour of 'quick and easy' or 'safe' projects. It is generally only older, established researchers who can undertake the kinds of large-scale longitudinal studies that are so valuable to health and medical research. This can also limit opportunities for early career researchers and those with career interruptions. Funding constraints can diminish incentives for top talent to remain in health and medical research (12); many researchers opt to pursue research overseas or move to the private sector.

Data delays have real impacts: CT Scans (38)

The impacts of delays in the provision of research data are not abstract – they cost lives. In 2008 researchers at the University of Melbourne received funding from the National Health and Medical Research Council (NHMRC) to investigate whether CT scans increased the risk of cancer by linking cancer notifications (held by the States and Territories) with de-identified Medicare Benefits Schedule (MBS) data. However, the study was seriously delayed: Commonwealth approval took *three years* to gain. Having finally managed to link the data necessary, the researchers found that cancer risk was increased by an average of 24% for individuals exposed to CT scans before the age of twenty; for those exposed at 'very young ages', the risk was *200%* higher. The results were published in the *British Medical Journal* in 2013 (39) and led to the development of educational materials a year later aimed at making both radiologists and the public aware of the risks.

The delays encountered by the researchers may well have led to a number of unnecessary exposures to CT radiation, and, in the longer run, unnecessary cancers. The lead researcher on the project noted that:

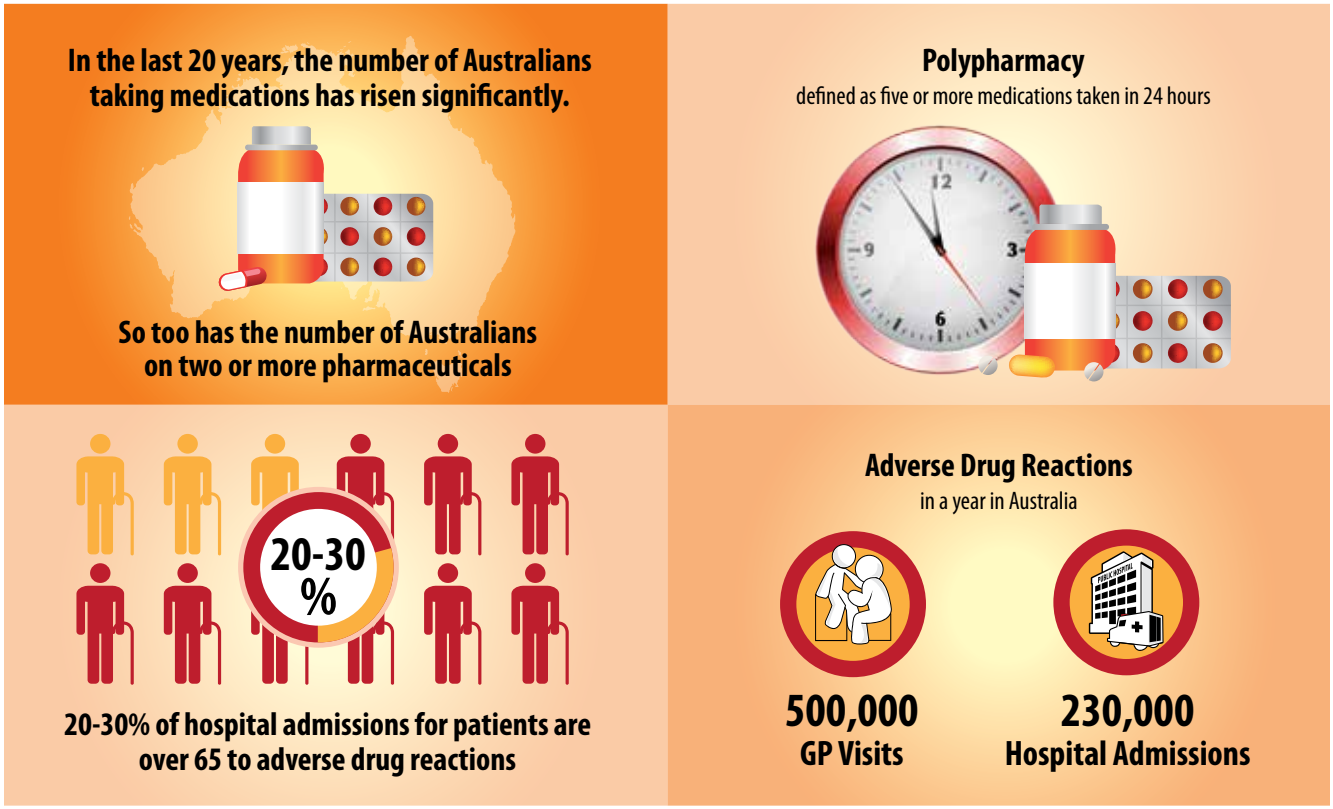
> Had our study been approved sooner, and been able to proceed at an earlier date, we would have had results sooner, with… benefits in terms of improved guidelines for CT usage, [fewer] exposures and fewer cancers.

The same researcher also stated that, as an established researcher 'in the twilight of my career' he had 'little to lose' and was not unduly deterred by 'long delays in obtaining approvals'. Had he been younger, he would have cut his losses and 'move[d] on to more immediately productive projects' – meaning it might well have been many more years before this research was undertaken. It is difficult to know how many potentially ground-breaking studies have been frustrated because of the same factors faced by these researchers – and the potential improvement to Australian health and wellbeing that has been delayed and lost. It is likely to be massive.

Case Study: Medication Safety

Australia's inability to utilise its vast reserves of data to improve post-market surveillance of pharmaceuticals epitomises the state of HMR data access in this country. As the number of Australians with chronic disease continues to grow, so too has the number of drugs consumed by individuals over long period of time. For instance, the number of Australians on cholesterol medication jumped from 1% in 1991 to 14% in 2011, while the number of individuals taking medication for diabetes tripled in the same time (40). This has resulted in number of Australians using two or more prescription drugs in addition to several un-prescribed off-the shelf supplements at the same time. As a consequence, medication safety resulting from multiple drug interactions has become an area of great concern in this country. It is estimated that about 500,000 visits to GPs and 230,000 hospitals admissions annually are due to adverse drug reactions – more hospitalisations than diabetes, asthma and heart failure ***combined.*** 12% of ***all medical admissions*** and 20-30% of all hospital admissions for patients over 65 are attributable to the same cause. These issues are likely to deepen unless action is taken ***now*** to understand the effects of multiple drug interactions in greater detail.

Post market surveillance to answer medication safety questions related to outcomes of multiple drugs on different population groups (some whom may not have been represented in the clinical trials) are rarely addressed. The issue is not the lack of PBS and MBS data or quality of data collections. This is often due to the fact that Australia's drug oversight mechanisms are failing to keep pace with the changing profile of pharmaceutical usage (41).



References: Australian Commission on Safety and Quality in Health Care (2013), Literature Review : Medication Safety in Australia. ACSQHC, Sydney
Morgan T, Williamson M, Pirotta M, Stewart K, Myers S, Barners J. A national census of medicines use: a 24 hour snapshot of Australians aged 50 years and older. The Medical Journal of Australia. 2011;196(1):50-53.

Figure 7: Medication Safety

Much ground could be made in tracking pharmaceutical efficacy if Australia had a consistent, systematic and data-backed post-market pharmaceutical surveillance initiative in place. We currently have no standardised system that links Australian Pharmaceutical Benefits Scheme data (PBS) to datasets that provide information on adverse events and patient outcomes: a fact that is particularly egregious considering that the information necessary to understand the phenomena above is being ***collected*** but ***not used***.

It is of no surprise that over the last 25 years fewer than 250 population-based studies have been conducted using PBS data: something which means that our understanding of pharmaceutical effectiveness is years behind where it could and ***should*** be (42).

Similar work has been successfully conducted in other contexts overseas. As long ago as ***2004*** researchers from the US healthcare provider Kaiser Permanente (KP) and the US Food and Drug Administration (FDA) analysed 1.4 million KP health records and found that individuals taking Vioxx (rofecoxib), a drug for rheumatoid arthritis, tripled the risk of heart attacks and cardiac deaths compared to similar pharmaceuticals. According to senior FDA investigator Dr David Graham, Vioxx was probably responsible for 140,000 heart attacks and ***60,000*** deaths in America alone (43).

While scientists at Merck – the company producing the drug – may have been aware that there were problems with Vioxx (44), it was Kaiser's database which led to the 'first real pickup' that there was a serious risk associated with its use (45). Merck voluntarily withdrew the product on Sep. 30 of that year and faced a congressional grilling and a number of lawsuits as a result. KP has since performed similar research on other drugs: for instance, a recent study on oral contraceptives (OCs) using 835,000 KP health records showed that certain OCs lead to a much higher risk of venous thromboembolic events (46). Another, non-KP study has also demonstrated that the analysis of electronic health records would have allowed researchers to determine in just eighteen months that a common diabetes medication raised the risk of heart attack for patients – when it took 7-8 years using traditional methods for concerns to be raised (47).

It is quite clear then that linking data and performing routine analysis on drugs can elicit a great deal of information regarding pharmaceutical efficacy and make a significant contribution to improving the health and wellbeing of Australians. It will reduce demand on the healthcare system and help prevent more public health tragedies along the lines of Vioxx – but it also promises to help democratise understandings of prescription drugs and make both patients and GPs more aware of the risks and benefits associated. As we continue to stress, the time to act is ***now***: we have the data. We need the will.

Researchers do their best to chart their way through a byzantine mix of inconsistent policies and processes surrounding data access and use. Likewise, Australian research suffers from the fact that data is rarely pre-prepared and made analysis-ready prior to requests for usage, and nor are there clear policies regarding research data after research.

Funding arrangements and legislative inconsistency further exacerbate an already complex environment, leaving data custodians unwilling to grant data and researchers focusing on shorter-term research projects. The confluence of these factors leads to a much slower pace of research. As a consequence, the research outcomes are often not able to influence service delivery in a timely manner.

4 HEALTH AND MEDICAL RESEARCH: LESSONS FROM OVERSEAS

The failure of successive Australian governments to take meaningful action around health and medical research data stands in stark contrast to the examples of several other developed nations. When it comes to the use of HMR data, Australia is falling behind countries already taking proactive steps to ensure that their data resources are used to their full potential. Thus, in this section we will briefly examine some key international best practice examples of arrangements surrounding HMR data. If anything, the experiences of these countries demonstrate that effective frameworks regarding the collection, sharing and use of data for health and medical research can be developed given a combination of political backbone and sufficient advocacy on the part of the research sector.

4.1 The UK

The UK stands as a prime example of what can be achieved when there is both *vision* and *sufficient will* on the part of healthcare leaders, researchers and politicians to grapple with the incredible opportunities afforded by data-enabled health research. In recent years the National Health System (NHS) has established several bodies and initiatives aimed at standardising data access processes and making large amounts of data available to health researchers with a minimum of time and effort – all while placing an important emphasis on the protection of patient confidentiality (48-53). The section below focuses on three key components of this shift: the establishment of a body concerned with the collection and curation of HMR data, one for coordinating and streamlining research, and finally, a massive, integrated healthcare dataset for researchers to use with clear access protocols.

NHS Digital: A Data Curating Entity and a Digital Health Strategy with Teeth

The English component of the NHS established NHS Digital in 2013, an agency responsible for collecting, analysing and disseminating England's health and social care data. NHS Digital has a succinct and well-articulated strategy in place which contains an ambitious goal for 2020:

> Our overarching objective is that by 2020 *we will have revolutionised the way* technology, data and information are used to transform the delivery of England's health and social care services (48)

The use of data to inform improvement in healthcare is identified by the strategy as of critical importance in 'support [ing] the achievement of better and health and wellbeing in the population' by 'enhancing the ability of citizens, providers, commissioners, national bodies and researchers to use data to make better decisions on how to improve health and healthcare'. As part of the roadmap to digital health, the strategy has identified four key goals (among others) related to data access:

- **Ease of access:** ensure that data are readily and easily accessible
- **Transparency:** have a clear and transparent process for accessing data
- **Innovative access:** develop new and innovative ways of accessing data
- **Greater reach:** increase the number of customers of the data so that more data is being used to improve health and care with an appropriate legal basis.

The implementation of a data-centric strategy supported by clearly articulated and achievable objectives – such as the ones above – can go a long way toward promoting a culture of data sharing and use. Indeed, alongside NHS Digital sits the National 'Health Data Finder', which aims to provide researchers with a comprehensive list of all research-available health datasets as well as their metadata and data dictionaries.

A Research Coordinating Body

Likewise, the NHS also has a Health Research Authority (HRA). The HRA is an arm's length body of the UK Department of Health that was established in 2015 with the intention of providing a single entity for the coordination of HMR. It is explicitly aimed at making the UK 'a great place to do health research' and 'to build confidence and participation in health research [so as to] improve the nation's health' (49). The organisation is responsible for the functions of streamlining and regulating research while protecting the interests and confidentiality of patients, and this dual responsibility means that neither efficiency in research nor patient interests are compromised.

Key to this has been ensuring the creation of an open, consistent and transparent research application process. Researchers can use a publicly accessible decision-making tool as a guide regarding what sorts of approvals are required for their project – something that minimises potential ambiguity and wasted time. Likewise, the HRA (51) is a part of the UK's

Integrated Research Application System (IRAS) (50), which lets researchers utilise a ***single system*** when it comes to applying for permissions and approvals for HMR: something which stands in clear relief to Australia's incredibly fragmented system of ethics approvals and application processes. In fact, the ***status*** of all reviews by ethics committees and other research approval bodies are open so that researchers have a good idea as to if and or when their project will be approved (51). Researchers are also encouraged to register their projects in a publicly accessible database.

Major Datasets: The UK Biobank

The UK has also pioneered the creation of large-scale datasets aimed specifically at researchers operating in the HMR sector. The UK Biobank (53) is a dataset with complete health and wellbeing data of 500,000 volunteer participants that aims to help researchers improve prevention, diagnosis and treatment of a wide range of diseases – both acute and chronic – such as cancer, heart disease, stroke, diabetes, arthritis, osteoporosis, eye disorders, depression and dementia. Participants aged between 40-69 were recruited in 2006-2010 and have provided the dataset with a wide variety of measurements, including:

- Diet, cognitive function, work history, digestive health
- Blood samples and saliva samples
- Genomic information.

Encouragingly, the Biobank demonstrates that individual privacy and public benefit need not be mutually exclusive. Consultations regarding the ethics and governance frameworks surrounding the Biobank were held between 2000-2004, data collected between 2006-2010, and access procedure consultations held in 2011; researcher access was granted in the following year. In the words of its website, the Biobank is open to 'any bona fide researcher undertaking health related research, anywhere in the world', whether they are from academia or industry (54). Access is provided through a simple online form, and, as per the Ethics and Governance Framework, researchers are required to place findings derived from Biobank data (whether positive or negative) in the public domain. The Biobank's *Access Procedures* document states that

> The objective of these Access Procedures is to facilitate access to the samples and data so that they get the ***widest possible usage*** while ensuring that such access and usage is consistent with the undertaking given to the participants and the wider public interest (including being lawful and compatible with respect for human rights) (55).

This is not mere rhetoric: the Biobank's Annual Review of 2016/17 states that 4000 researchers have been approved to use Biobank data, and the number using the resource has grown significantly each year (52). The number of papers published using Biobank data doubled year-on-year between 2013-6; as of 2016, this number was 160. The Biobank also embarked on an ambitious project in 2015: to image 100,000 participants' brains, hearts, abdomens, bones and carotid arteries and to link these data to a wide range of other data points – currently cancer, death, hospital episodes and general practice – and is developing algorithms to accurately identify diseases and their subsets. It is hardly surprising then that such an eminent researcher as Prof Stephen MacMahon AO, Professor of Cardiovascular Medicine at the University of Sydney and Principal Director of The George Institute for Global Health has stated that:

> The UK Biobank project is widely regarded as a ground-breaking study that will deliver a wealth of new information about both environmental and genetic determinants of common diseases. In the last century it was the US-based Framingham Study that made fundamental discoveries such as cholesterol as a cause of heart disease and blood pressure as a cause of stroke. This century, the UK Biobank looks currently set to be the talking point – on diseases as diverse as Alzheimer's and prostate cancer (56).

In fact, the UK is so far in front of Australia when it comes to data usage that, as noted by the Productivity Commission, 'some of our best health researchers use UK health datasets, as ours are unavailable to them' – so that they do not have to wait as long as eight years for access (25).

4.2 The USA

The Value of Data

In 2013 President Obama identified data as a key component in 'fuel[ing] entrepreneurship, innovation and scientific discovery that improves Americans' lives and contributes significantly to job creation' (57). To this end, he issued an executive order mandating that data resources held by federal agencies should by default be made open and machine readable – including those in the health domain. The framework (58) developed by the US Chief Information Officer in response to this order contained several principles for open data, of which the key tenets are:

- **Public:** 'agencies must adopt a presumption in favor of openness to the extent permitted by law and subject to privacy, confidentiality, or other valid restrictions'.
- **Accessible:** data needs to be made available (to the extent permitted legally) in machine readable formats that assist automated processing.
- **Described:** data should be fully described such that users are able to understand the strengths and limitations of datasets and are able to determine how best to use them.
- **Complete:** data should be held, described and made available in its primary form – i.e. as collected at the transactional level.
- **Timely:** data must be made available for secondary use as close to the time of collection as possible.

Another significant aspect of the framework is the fact that agencies are also required to ***budget*** for the use of data – something which means that agencies must plan for alternative and secondary uses of the data they generate upfront. Taken together, these principles are highly important in promoting a culture of data sharing and ***usage*** as opposed to one centred on risk-averse protection as in Australia. Indeed, by January 2017, the US Department of Health and Human Services had made more than 3000 health datasets routinely available for research and innovation, while a number of technology companies have used this data to develop innovative applications and products (59).

HIPAA: Clear Legislation

Central to this effort is a clear legislative environment that underpins the broader usage and re-usage of healthcare data. HIPAA, the American ***Health Insurance Portability and Accountability Act*** (1996) (60), is a piece of legislation that outlines the provisions and security standards required for safeguarding individually identifiable healthcare data. HIPAA applies to all entities involved in the exchange of health information and is an example of a clear, well-considered piece of law-making: one which, in clear contrast to our existing legal environment, acts as an ***enabling*** framework rather than one that restricts healthcare data use. Indeed, HIPAA supports the US Government's broader movements toward Open Data by providing clarity to those charged with preparing datasets for wider and for secondary usage – and thus helps expedite the research access process. Nor does it place any restrictions on the use or disclosure of de-identified healthcare data.

Of special importance is the fact that HIPAA supports consistent standards for research data preparation. In 2012 the Office for Civil Rights (OCR) developed clear guidelines on the methods used for the de-identification of protected health information in accordance with the HIPAA Privacy Rule. There are two options available: (a) Safe Harbour: the removal of eighteen specific identifiers regarding the individual and their relatives, and (b), Expert Determination, which defines a formal set of statistical principles to be applied by an expert statistician as well as a range of other conditions – such as what constitutes an expert statistician, the acceptable level of identification risk, principles to be used by statisticians and the duration for which determination is valid. Most importantly, the OCR has also described the workflow required for data managers and administrators working with the expert statistician to assess the risk of re-identification.

The IRIS Registry: Best Practice and Research

American clinicians and researchers have also been making great strides in bridging the gap between research and best practice. IRIS – 'Intelligent Research In-Sight' – is a clinical registry maintained by the American Academy of Ophthalmologists (AAO) and has records for 13,500 ophthalmologists, about two-thirds of all those practicing in the US (61). IRIS was developed in order to allow participating ophthalmologists to benchmark their performance against their peers; the registry works by extracting data directly from practice electronic health records. Yet, IRIS also stands at the confluence between research and clinical improvement: it is a powerful resource for answering research questions and feeding this information back to clinicians and policymakers in real time.

This is best demonstrated by recent discussions surrounding the most effective way to treat wet age-related macular degeneration (Wet AMD) – one of the leading causes of blindness in the US (62). In 2015 an FDA Committee enquired into whether compound injection therapy used to treat Wet AMD led to higher complication rates than single injections. A member of the committee had financial interests in a company creating single injection drugs and argued that compound injections led to higher rates of complications and should be heavily regulated. In response, researchers made a query into IRIS, found that there was no higher rate of complication, and were able to confirm that both types of therapy were equally effective. Consequently, the FDA did not regulate against compound injections (63).

This was an issue which would have taken several years and considerable amounts of money to evaluate through traditional clinical trials. Yet, IRIS allowed the researchers to access data on 827,000 Wet AMD injections and have results in ***thirty-six hours*** at a fraction of the cost. Had compound therapies been found to cause higher rates of complication, this insight would have been rapidly disseminated back to clinicians – thus facilitating the rapid adoption of best practice treatments.

IRIS has already led to other advances – such as the suggestion of new therapeutic techniques for anterior vitrectomies performed during cataract surgery – and offers great potential in a whole host of other areas, such as drug surveillance and drug monitoring (64). A number of papers have also been published using IRIS data. It is anticipated that that these investigative functions will become automated over time: indeed, AAO Trustee George Williams has compared the current state of the registry to the 'first iPhone' and expects that future iterations will be exponentially more powerful and capable of performing far more complex tasks (65). Ultimately, this might mean that IRIS is able to continuously achieve what AAO Chairman David Parke II has termed the 'holy grail' of healthcare – the enhancement of outcomes and reduction of 'unnecessary or ineffective care' (66). A number of other American specialist colleges – such as Neurology and Rheumatology – are now following suit and established their own clinical quality registries.

4.3 Other Countries

Singapore

In 2015, the National Medical Research Council (NMRC) of Singapore directed that all research grant applications greater than SGD 250,000 (AU$245,000) must include a data sharing plan that covers the sharing of both research findings and data assets (67). The costs of data and research sharing are funded by the NMRC as indirect research costs separately from the research project budget.

Finland and Estonia

Finland and Estonia are both remarkable in their adoption of e-health records (68, 69): each country has complete Electronic Health Records for their entire population. Not only that, but they have moved to share these records between the two countries so that both can enable better healthcare for their respective populations (70).

Estonia has led the European Union resolution to adopt an open-data-approach for new public sector IT-systems or when rebuilding existing systems (71).

Each citizen of Estonia who has visited a doctor has an online e-health record which is part of a nationwide system that integrates data from Estonia's many healthcare providers. "By default, medical specialists can access data, but any patient can choose to deny access to any case related data, to any, or all care providers; including one's own general practitioner/ family physician." (72).

A lesson from the Estonian experience is that "… the Citizen needs to be confident in the government's ability to keep their data safe -- in terms of confidentiality, integrity and availability – establishing a strong link between privacy and information security." (72)

4.4 Lessons for Australia

The international examples discussed provide some insights for addressing the barriers to data-enabled HMR in the Australian context. As identified in Section 3, the major barriers are:

- Inconsistent data access request protocols, duplication and wasted time
- Outdated data management, release and preparation frameworks slowing data release
- Complex legislative environments and funding policies affecting release of data and researcher incentives.

To realise the full potential of HMR it is necessary, but not sufficient, to collect and securely store data. In order to be safe and useful, data must be made available and accessible in an efficient and predictable manner that creates trust through prioritising data security and the preservation of privacy and confidentiality. This requires a shift from the current reactive, ad-hoc data practices to more proactive, transparent and standardised protocols for data collection, storage, preparation, extraction, transfer and release. In Australia, there is a need to address current arrangements for funding and ethics approvals to create more streamlined processes reflective of modern research techniques. These data, ethics and funding arrangements must also be underpinned by a legislative and policy environment that supports a culture of data confidence for all stakeholders. This would enable us to move from a protective and closed culture to one of data sharing and release to authorized researchers.

Through an integrated approach that incorporates lessons learnt overseas with the particulars of our Australian context we can support the HMR ecosystem and realise the full potential of data-enabled HMR.

Governance Principles for Data Preparation, Management and Release

Australia needs a harmonised set of HMR data governance arrangements that are distributed and layered: arrangements that apply ***proactively*** at each stage of the data journey that includes data preparation, data management and data release for access by researchers. These should provide two things: first, a consistent ***path*** for data flows from collection to researcher, and second, clear guidelines outlining the steps necessary to ensure that privacy is preserved, confidentiality is maintained, and data security principles are applied ready to be used to its full potential by every legitimate stakeholder operating in HMR.

To ensure that data flows routinely and consistently across Australia's healthcare environment, clearly articulated ***routes*** must be established. For this to happen, we suggest that data needs to move from the primary point of collection – e.g. from care settings through to Data Holding Organisations (DHOs), such as government agencies, research institutes and technology firms, who can prepare the data for wider use. Once the data is prepared, it must be managed by an accredited agency in a secure environment before the data can be released to authorized trusted researchers.

We envision that DHOs would be primarily responsible for the collation, storage, preparation and maintenance of primary datasets. Besides data captured in structured databases, it is extremely important that DHOs collect and maintain data that are ***semi-structured*** or ***unstructured***. It is estimated that "for every hour spent with the patient, providers spend two hours documenting patient care" essentially for input into a structured database. If this can remain in unstructured form for downstream analysis and interpretation, it would free up valuable provider time for direct patient care. Currently this type of data is largely ignored.

Unstructured data is important

In addition to the health data held in structured databases, significant proportion of all the digital data generated by the healthcare sector is unstructured and is largely untapped. Some estimates put this number at 80%.

Big Data technologies and tools have advanced sufficiently to effectively exploit the massive amounts of unstructured data in the form of free texts, images, audio recordings, that are generated across the health system. For example, a recent proof-of-concept project in the U.K has demonstrated that it is possible to derive meaningful insights from unstructured clinician notes using automated technologies to deliver better patient outcomes (73).

Analysis of unstructured clinical data together with structured information can lead to targeted patient-centred care and better overall health outcomes. We cannot, any longer, afford to consign analysis of unstructured data to the too-hard basket.

Accredited Release Authorities (ARA), as defined by the Productivity Commission (25), are 'trusted independent entities responsible for acquiring, developing and integrating datasets with a view to sharing and release for research'. We see ARAs taking on the role of managing the secure infrastructure necessary for data received from one or more DHOs and implement policies and protocols necessary for release of data for research. ARAs would build data collections from structured and unstructured data flows from DHOs, using standardised, privacy preserving ingestion tools. They would ensure that all data sets, including those from MBS and PBS, are kept linkage ready, along with appropriate meta data dictionaries and exploratory tools to enable trusted researchers to explore and test data from different data collections.

ARAs would also be responsible for accreditation of trusted researchers, standardize transparent access protocols before data is released.

At all times the flows outlined above must operate within nationally consistent and ***unambiguous*** governance regimes that specify how data is to be governed at each stage of the data journey. As soon as data is collected it should be cleaned, prepared and made machine-readable by DHOs with appropriate privacy and security controls.

When data flows to ARAs, standard, high-value datasets should be de-identified and pre-linked to facilitate efficient access for trusted users, while other datasets must be made research or ***linkage-ready*** to provide for bespoke research projects in the shortest possible timeframe.

We note that privacy, confidentiality and security are distinct from each other. While all three concepts need to be considered at every stage of the data research data flow, an engineered design approach involves a focus on privacy upfront at the initial data holding stage, a confidentiality focus at the data release stage through domain hubs, with robust data security underpinning at all stages of data transmission and especially wherever data is stored. A layered governance approach to privacy, security and confidentiality is illustrated in *Figure 8*.

Privacy vs Confidentiality vs Security

Although often used interchangeably, privacy, confidentiality and security are distinct concepts (74). Privacy is about the person. It is essentially the right of individual to keep their personal data private. Confidentiality is about sharing of the data. It primarily relates to the release of data and related data products and information about an individual to third parties without express consent of individuals. Security is about data storage. It is about keeping stored information, whether on paper or electronic media, safe from unauthorised access by individuals and organisations.

Robust de-identification techniques and privacy preserving frameworks are necessary to protect ***privacy*** of personal information; and appropriate policies, procedures and risk-based frameworks are key aspects to maintain ***confidentiality***. Security is largely a technology issue although safe practices and policies are essential to support security technologies such as encryption, multi-factor authentication etc.

Governance strategies need to be targeted to address each of these distinct concepts.

Privacy, Security, Confidentiality by Design

We believe that it is of utmost importance to implement an approach that combines what we perceive as three cardinal principles of privacy, confidentiality and security, right from the initial movement of data from the primary-point of collection through to release for secondary use through ARAs. Protecting privacy is not simply a matter of regulatory compliance that can be 'bolted on' or performed last minute by entities involved in data preparation, but rather that privacy preserving techniques are applied as early as possible in the data flow. In the US, privacy-protective legislation such as HIPAA (60), and in Europe, data confidentiality protection regulation such as GDPR (75), are in place to provide clear guidelines that are ***mandatory*** for all organisations dealing with personal data. While movements have been made in this direction in Australia, progress has been regrettably slow.

Privacy protecting techniques, such as de-identification (76), need to be robust enough to drastically lower the potential risk of re-identification. While current techniques have come a long way over the years, continued research and development is necessary to further lower the risk, given the high sensitivity of health data.

Why is a "privacy by design" approach (77) important? Often the process of building of privacy protection and data extraction for specific data is initiated only after all individual approvals from the various data holding agencies, data linkage organisations and the human research ethics committees are obtained. If any of the approval agencies raise a concern or a query at any stage the researcher must re-commence the cycle of investigation, data discovery and communication with data custodians and data holding agencies. Minor changes can lead to the whole cycle repeating. This is an extraordinarily time consuming, reactive and cumbersome process that often adds very little value to the integrity of research conduct.

A governance approach that builds these privacy protections (based on legislative frameworks such as the proposed Data Sharing and Release Act (25)) upstream as a matter of course would allow for a consistent, proactive and efficient approach to preserving data privacy and security. This would free up more resources for research and analysis rather than researchers' time being dominated by low-value bureaucratic functions.

The Five Safes Framework (78) also proposed by the Productivity Commission is a step in the right direction for data sharing and release.

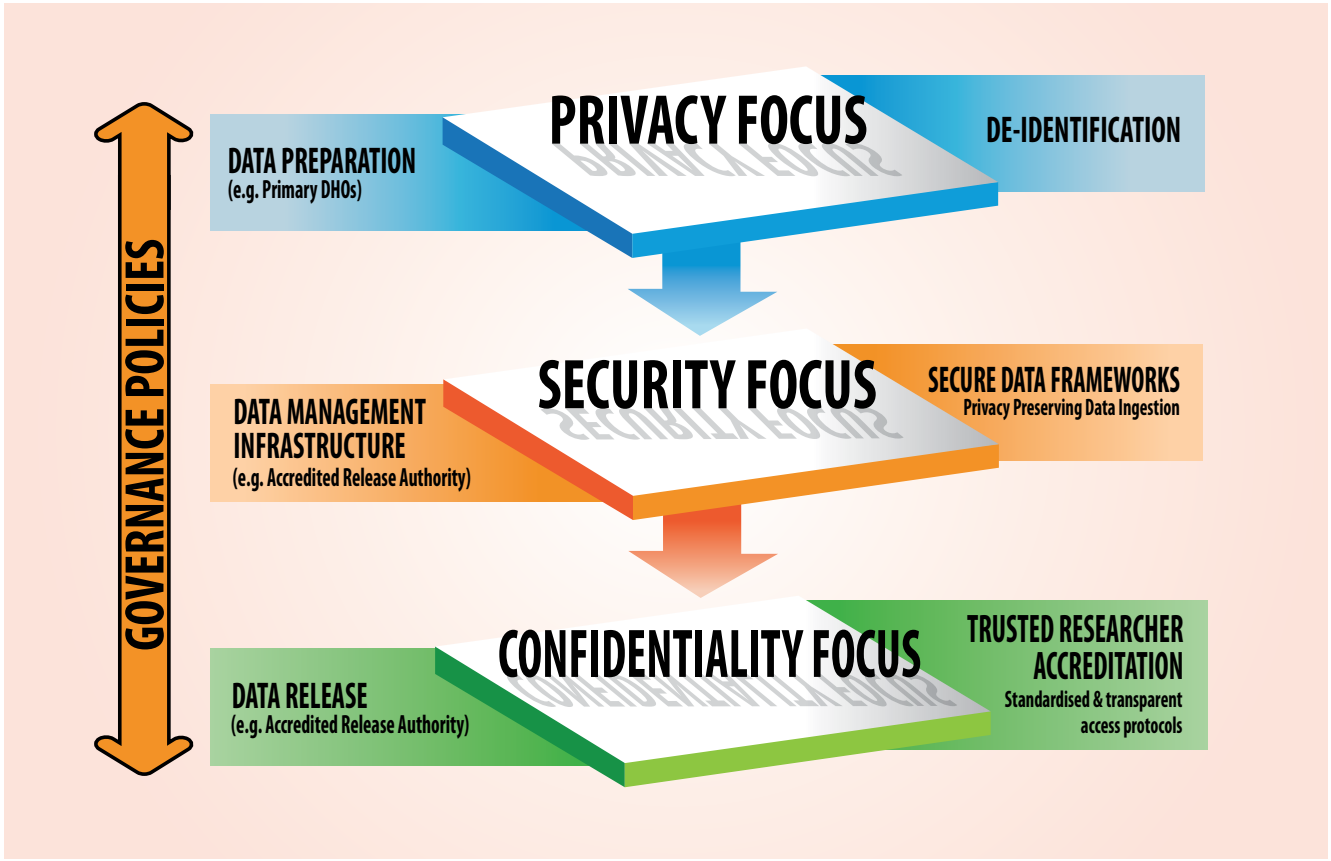


Figure 8: Layered Governance

Data Access Protocols

It is of critical importance that Australia implements standardised data access protocols to minimise inefficiencies in the application process. This requires the creation of streamlined, cross-national data access rules that are ***consistent*** across HMR: protocols that are recognised and accepted across all public agencies involved in data holding, data release and research approval, e.g. DHOs, ARAs, HRECS.

Data access protocols must:

- Be transparent and ***openly published***
- Provide specific timelines for approval and linkage processes as well as information about charges involved
- Detail ***where*** to apply, ***what*** to expect, and with ***whom*** researchers must liaise
- Provide clear rules around consent
- Outline guidelines regarding tools researchers should use for discovery, exploration and analysis of data.

Alongside this, it is important that these access rules ensure that researchers are able at ***all times*** to check the status of their data requests and be provided with realistic updates and timeframes. The protocol should also detail exactly what steps researchers should take with the datasets they have created to ensure that they are usable and accessible for future analysis. Finally, the establishment of a single national ethics application, or at the very least mandating that ethics applications are recognised across different HRECs also promises to streamline the process without unduly compromising ethical considerations.

Data after Research

Consistent policies mandating the retention of linked data from publicly-funded research need to be implemented to further increase the efficiency of the research process. There are movements in this direction: The Productivity Commission has recommended that the Federal Government abolish its requirement to destroy linked datasets at the completion of research projects using MBS or PBS data – and that data and metadata used in publicly funded research projects should be made available to other researchers to use with appropriate approval at project completion (25). These steps cannot be neglected if Australian HMR is to thrive.

Legislation and Policy

Finally, it is important to clarify and simplify legislative arrangements surrounding the use of health data and data more generally. The Productivity Commission's proposed ***Data Sharing and Release Act*** (DRSA) is a positive step forward (79). The DRSA is an 'umbrella' piece of legislation that promises to apply to all digital data, and can, in turn, help Custodians and others view data as a ***valuable asset*** and not as a risk or overhead. This is of the utmost importance in ensuring a culture of data confidence: one that encourages data holders to actively release non-sensitive data.

When it comes to funding, it is also important that a variety of funding is available, including longer-term, funding – with appropriate incentives in place – to promote valuable longitudinal studies and combat the culture of 'publish or perish'.

Australia is well behind other developed nations who have recognised the opportunities afforded by coherent data policies. The UK has taken important steps forward to maximise health data usage and is rightly seen as a world leader for its clear, consistent and streamlined approach to research data preparation and access.

Australia cannot afford to fall further behind. HMR data needs to flow consistently across the entirely of our healthcare environment. Data preparation and access processes and procedures need to be standardised and rationalised. Governance arrangements need to apply at each stage of the data journey to improve the efficiency of research access requests.

It is imperative that clear pathways are provided for research to translate into better policy, practice and service delivery.

5 CONCLUSIONS AND CALL TO ACTION

Australian health and medical research (HMR) is world-class, with over 340 research organisations involved in conducting health and medical research with a collective annual research funding exceeding $6 billion. Australian researchers excel in areas including disease causation, service delivery, development and monitoring of medications and other interventions, public health and policy as well as medical technologies and devices just to name a few. Yet many of the research findings are not integrated into policy or practice in a timely manner, as unnecessary delays in getting access to data coupled with administrative complexities prevent timely translation of research outcomes into best practice clinical care and public policy.

Australia routinely collects and stores comprehensive health services data that is well coded and structured. The AIHW manages several national data collections and prepares annual reports to inform policy decisions. As AIHW itself notes in its 2018 report (80), despite the breadth of health information available there are gaps in our knowledge and opportunities to make better use of linked data.

Australia's health data landscape is changing rapidly. New devices and applications are adding many types of new data that include text, video, audios and digital streams that are mainly consumer focused. The emerging area of genomics is slowly delivering important new data that could lead to targeted personalised care. Yet the capacity of health researchers to access and use this data is heavily compromised. This problem does not stem from consumer willingness to participate in health research or a lack of technological capacity. The Australian Digital Health Agency's recent National Health Digital Strategy (81) points out that technological developments and improvements in data analytic capabilities mean that data can now be used in ways that were not envisaged even 10 years ago. This is particularly relevant in the digital health context where the linkage, aggregation and sharing of health data can lead to significant new insights, more efficiently (and potentially more accurately) than ever before.

To capture the new opportunities afforded by data-driven HMR, we must shift the paradigm from risk aversion and fragmentation to confidence borne from rationalised and transparent policy and process. The opportunity cost that we miss out by maintaining the status-quo is significant. Allowing secondary use of MHR as outlined in ADHA's National Digital Health Strategy paper is only part of the solution and not a complete solution to the data availability and access problem faced by HMR. Until we release data for research we are holding back on active industry involvement and technological innovation. Concerns related to data security privacy and confidentiality are not peculiar to Australia. Other developed countries such as USA and UK have developed data governance frameworks and policies that enable efficient data release for research and research translation to policy and practice.

Our conservative 15-year projections show that cost savings and ROI in HMR are in the range of $3 billion in 15 years, providing almost 15% savings by 2033.

Only through embracing the expanded role of data in the health and medical research landscape and the support of a national mandate at the highest level can we achieve our shared goals of supporting the health system, delivering world class healthcare and ensuring economic sustainability and success.

Call to Action

For Australian researchers to have easy access to HMR data, Australia needs to move from ad-hoc data practices to ***standardised*** methods of data preparation, de-identification and release that reflect modern research techniques that are streamlined, transparent and consistent. Expediting access to routinely collected health data is not an abstract issue to be dealt with at some point in the future. The time to act is ***now***.

To resolve the current situation, we need to recognise the following:

- Consumers today are willing to share their health data to support research. Research Australia's surveys over the years consistently show that consumers are increasingly ready to share their data for research. ADHA's study too revealed that consumers trust in Australian Researchers accessing de-identified data for research. Therefore, the government policy in regard to data release should reflect consumer sentiments.
- It is essential to formulate and articulate a well documented governance framework that is transparent to researchers and enables them to get access to de-identified HMR datasets to address their specific research questions.
- Several countries around the world including UK and USA have been agile in improving access to health data for their researchers. For this to happen in Australia, an explicit commitment from the highest authorities at a national level is essential to streamline secure data flow across state borders and jurisdictional boundaries for research.
- The problem of fragmented data custodianship has resulted in opaque and complex processes in generating longitudinal linked datasets for HMR. There is a need to formulate policies that incentivise data custodians to prepare data for research readiness and promote an environment of data abundance for HMR.
- The current culture that polices data flow rather than enable data flow has got to change. For this to happen, data agencies need to be supported with appropriate standardised methodologies and technologies that can support a secure research environment that preserves privacy and confidentiality.
- We need to formulate a strategy for harnessing the massive amount of unstructured health data that is being generated all across the health system for research.
- Currently in the health care sector there is a significant lag before research findings are translated into policy and practice. We need to streamline processes to allow real-time (de-identified) service data to flow into the research environment while simultaneously enabling research findings to flow back to influence pro-active policy formulation and support evidence-based real-time service delivery.
- A vibrant HMR environment is essential to support a world-class health medical technology sector that can deliver innovative solutions that can benefit Australia and Australian healthcare consumers.

At a detailed implementation level, some of the actions that will be required are as follows:

- Formulate policies that enable HMR data flows continually and routinely across all points of the health system, from clinical points of service all the way to researchers without compromising individual privacy. The policy should enable a single window approach for researcher's access to both structured and unstructured data that is collected routinely.
- Design and articulate clearly defined risk-based frameworks so that standard, high-value linked datasets held by agencies such as AIHW, are available in de-identified form to trusted researchers without the current complex ethics approval processes.
- Ensure that the National Mutual Agreements for ethics approvals are respected and implemented uniformly across universities and ALL state and federal jurisdictions, with no fine print exemptions.
- Design a robust and well-articulated national security and privacy framework that specifies not only requirements for secure data management, but also state-of-the-art and standardised methodologies to ensure data privacy and confidentiality.
- Expedite the implementation of the Data Sharing and Release Act for the health sector. While data from several sectors are to be released, sadly the health sector is once again left out due to the unpreparedness of governments to establish trust and adopt a tech savy approach to tackle the debilitating privacy concerns not necessarily shared by the Australian public, as evidenced by several consumer surveys.

6 APPENDICES

Glossary of Terms

Term	Description
ABS	Australian Bureau of Statistics
AIHW	Australian Institute of Health and Welfare
Allied Health Services	Allied health includes services provided by health professionals, other than doctors, nurses, and dental professionals. Allied Health Professionals include physiotherapists, occupational therapists, dietitians etc.
Ancillary Health Services	Health services provided by health professionals, but which are not classed as Medical or Hospital, and are not covered by Medicare. Ancillary services include physiotherapy, dental services, speech therapy, ambulance travel, home nursing and spectacles. May also include some medicines that are not on the Pharmaceutical Benefits Scheme (PBS). Also known as General Treatment. (PrivateHealth.gov). The term "Ancillary" is typically used by Private Health Insurers.
ARA	Accredited Release Authority
Australian health system	Australia's health system is a 'web': a web of services, providers, recipients and organisational structures. (AIHW). It is a complex maze of private and public health services, funded by the public sector, private funders, and the consumers.
CMCRC	Capital Market Cooperative Research Centre
DHA	Digital Health Agency
DHO	Data Holding Organisation
DoH	Department of Health
DTO	Digital Transformation Organization
DVA	Department of Veteran Affairs
FB1	Flying Blind: Australian Consumers and Digital Health
GP	General Practice / General Practitioner
HCP	Hospital Casemix Protocol
HMQ	Health Market Quality
HREC	Human Research Ethics Committee
MBS	Medicare Billing Schedule
Medicare	Federal health insurance that provides Australian residents access to healthcare. Medicare aims to ensure that all Australians have access to free or low-cost medical, optometry, midwifery and hospital care and in special circumstances, allied health. (DoH)
NHMRC	National Health and Medical Research Council
PBS	Pharmaceutical Benefits Schedule

Term	Description
PhD	Doctor of Philosophy
PHI	Private Health Insurance
PHN	Primary Health Networks
PHRN	Population Health Research Network
POS	Point of Service. Any point where healthcare services are provided to the consumer. E.g. GP, Hospitals, CHC, Dental, Allied Health Service etc.
POSD	Point of Service Data
Primary Care	A person's first point of contact with the health system usually delivered via general medical and dental practitioners, nurses, Indigenous health workers, pharmacists and other allied health professionals such as physiotherapists, dietitians and chiropractors.
Private Hospital	Private hospitals are mainly owned and managed by private organisations; either for-profit companies, or not-for-profit non-government organisations. They include day hospitals that provide services on a day-only basis, and hospitals that provide overnight care.
Public Hospital	Hospitals mainly owned and managed by the state and territory governments. Public acute hospitals mainly provide 'acute care' for short periods, although some provide longer-term care, such as for rehabilitation. Public psychiatric hospitals specialise in the care of people with mental health problems, sometimes for long periods.
S95 Guidelines	NHMRC's Guidelines under Section 95 of the Privacy Act 1988 – These guidelines provide a framework in which medical research involving personal information obtained by Commonwealth agencies should be conducted, to ensure that such information is protected against unauthorised collection or disclosure.
s95a Guidelines	NHMRC's Guidelines under Section 95 and 95a Guidelines approved under Section 95A of the Privacy Act 1988 provide a framework to ensure privacy protection of health information that is collected, used or disclosed in the conduct of research and the compilation or analysis of statistics, relevant to public health or public safety, and in the conduct of health service management activities.
Secondary Care	Secondary care is medical care provided by a specialist or facility upon referral by a primary care physician (Nicholson 2012). It includes services provided by hospitals and specialist medical practices. Secondary healthcare can also refer to ongoing services not necessarily provided in the hospital, such as psychiatrists, physiotherapists and occupational therapists. (Health Issues Centre)
State DoH	State Department of Health
The ATS	The Australasian Triage Scale – A triage system is the basic structure in which all incoming patients are categorised into groups using a standard urgency rating scale or structure.
Trusted user	An organisation or individual accredited by an Accredited Release Authority for access to Data Hubs managed by the ARA.

National Legislations

Legislation	Description
The Privacy Act 1988 (Privacy Act)	**Purpose:** To regulate how personal information is must be handled, used and managed by most Australian and Norfolk Island Government agencies, all private sector and not-for-profit organisations with an annual turnover of more than $3 million, all private health service providers and some small businesses. **Relevance to data sharing:** As health information is regarded as one of the most sensitive types of personal information the Privacy Act provides extra protections around its handling. For example, individual's consent required before an organisation can collect their health information. **References:** https://www.legislation.gov.au/Details/C2018C00034 https://www.oaic.gov.au/privacy-law/privacy-act/ https://www.oaic.gov.au/privacy-law/privacy-act/australian-privacy-principles
NHMRC's Guidelines under Section 95 of the Privacy Act 1988 (s95guidelines) and **NHMRC's Guidelines under Section 95A of the Privacy Act 1988 (s95A Guidelines)**	**Purpose:** Provide guidelines on exceptions to the Privacy Act on handling of health information and personal information for health and medical research purposes, where obtain individuals' consent to use is often impracticable. These guidelines were issued by NHMRC and are approved by the Privacy Commissioner and legally binding. **Relevance to data sharing:** Guidelines under Section 95 of the Privacy Act 1988, set out procedures that HRECs and researchers must follow when personal information is disclosed from a Commonwealth agency for medical research purposes. Guidelines under Section 95A of the Privacy Act 1988, provide a framework for HRECs to assess proposals to handle health information held by organisations for health research (without individuals' consent). They ensure that the public interest in the research activities substantially outweighs the public interest in the protection of privacy. **References:** https://www.legislation.gov.au/Details/F2014L01500 https://www.legislation.gov.au/Details/F2014L00243
Census and Statistics Act 1905 (CSA)	**Purpose:** The CSA provides the Australian Statistician with the authority to conduct statistical collections, including the Census of Population and Housing, and, when necessary, to direct a person to provide statistical information. The CSA requires the ABS to publish and disseminate compilations and analyses of statistical information and to maintain the confidentiality of information collected under the Act. **References:** https://www.legislation.gov.au/Details/C2016C01005

Legislation	Description
National Health Act 1953 **Section 135AA of the National Health Act 1953**	**Purpose:** An Act relating to the provision of pharmaceutical, sickness and hospital benefits, and of medical and dental services. It was used to determine who could access health concessions and concession cards ('concessional beneficiary'). **Relevance to data sharing:** In 2008, Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programs was issued by the Privacy Commissioner under section 135AA of the National Health Act 1953. Among other things, section 135AA(5) of the National Health Act requires that the Guidelines prohibit agencies from storing claims information obtained under the Medicare Benefits Program and the Pharmaceutical Benefits Program on the same database and establish the circumstances under which this information may be linked. The Guidelines also prescribe periods of time for which claims information may be retained in various forms. **References:** https://www.legislation.gov.au/Details/C2017C00250 https://www.legislation.gov.au/Details/F2008L00706 https://www.alrc.gov.au/publications/61.%20Electronic%20Health%20Information%20Systems/medicare-and-pharmaceutical-benefits-database#_ftnref38 https://www.oaic.gov.au/images/documents/migrated/migrated/healthreview.pdf
Health Insurance Act 1973	**Purpose:** The Health Insurance Act of 1973 introduced Medicare, a universal healthcare plan in Australia **References:** https://www.legislation.gov.au/Details/C2017C00255
Australian Bureau of Statistics Act 1975	**Purpose:** The Australian Bureau of Statistics Act 1975 establishes the ABS as an independent statutory authority, defines the functions of the ABS, establishes the office of Australian Statistician and describes the terms under which the Australian Statistician can be appointed to, and removed from, office. The Australian Bureau of Statistics Act 1975 also provides for the appointment of the staff of the ABS and establishes the Australian Statistics Advisory Council. **References:** https://www.legislation.gov.au/Details/C2014C00618
The Freedom of Information Act 1982 (FOI Act)	**Purpose:** The Freedom of Information Act 1982 (FOI Act) Provides a legally enforceable right of access to individuals to request access to documents from Australian Government ministers and most agencies, although the obligations of agencies and ministers are different. **References:** https://www.legislation.gov.au/Series/C2004A02562

Legislation	Description
Australian Institute of Health and Welfare (AIHW) Act 1987.	**Purpose:** The Australian Institute of Health and Welfare (AIHW) is Australia's national agency for information and statistics on Australia's health and welfare., established as an Australian Government statutory agency under the Australian Institute of Health and Welfare Act 1987. **Relevance to data sharing:** The Act contains very strong confidentiality protections for all data held and requires the AIHW to publish two key biennial reports in alternate years: Australia's health and Australia's welfare. Numerous other reports are produced each year, all of which are available free of charge on the AIHW website.[3] Privacy Obligations of AIHW are laid out under Section 29 of Section of the Australian Institute of Health and Welfare Act 1987. NHMRC's s95 and s9A guidelines apply for exceptions. **References:** https://www.legislation.gov.au/Details/C2016C01008 https://www.aihw.gov.au/privacy-policy

7 REFERENCES

1. Srinivasan, U., Rao, S., Ramachandran, D., & Jonas, D. 2016, Flying Blind: Australian Consumers and Digital Health, Volume 1: Australian Health Data Series, Health Market Quality Research Program, CMCRC, Sydney.
2. Volume 2: Researchers and the Health Data Maze | Flying Blind [Internet]. Flyingblind.cmcrc.com. 2018 [cited 17 May 2018]. Available from: https://flyingblind.cmcrc.com/researchers-health-data.
3. Chalmers K, Pearson S-A, Elshaug AG. Quantifying low-value care: a patient-centric versus service-centric lens. BMJ Quality & Safety. 2017;26(10):855-8.
4. Brownlee SM, Chalkidou KMD, Doust JP, Elshaug AGP, Glasziou PP, Heath IF, et al. Evidence for overuse of medical services around the world. Lancet, The. 2017;390(10090):156-68.
5. Ghaffar A, Langlois EV, Rasanathan K, Peterson S, Adedokun L, Tran NT. Strengthening health systems through embedded research. Bulletin of the World Health Organization. 2017;95(2):87-.
6. National Academies of Sciences a, Division H, Health B, Threats F. Case Studies in Big Data and Analysis [Internet]. Ncbi.nlm.nih.gov. 2018 [cited 17 May 2018]. Available from: https://www.ncbi.nlm.nih.gov/books/NBK401950/.
7. PHRN. Submission to the Productivity Commission Data Availability and Use Inquiry from the Population Health Research Network. July 2016 [Internet]. Pc.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.phrn.org.au/media/80973/phrn-response-to-the-data-availability-and-use-inquiry_v10_29072016.pdf
8. Australian National Statistical Service [Internet] What is data linking? [cited 17 May 2018]. Available from: http://www.nss.gov.au/nss/home.nsf/533222ebfd5ac03aca25711000044c9e/91242a5a14b12e26ca257ba8007b0819/$FILE/data%20linking%20w.pdf.
9. Sixth Interim Report (Big health data: Australia's big potential) – Parliament of Australia [Internet]. Aph.gov.au. 2018 [cited 17 May 2018]. Available from: https://www.aph.gov.au/Parliamentary_Business/Committees/Senate/Health/Health/Sixth_Interim_Report.
10. Khan A, Uddin S, Srinivasan U. Comorbidity network for chronic disease: A novel approach to understand type 2 diabetes progression. International Journal of Medical Informatics. 2018;115:1-9.
11. Bureau of Health Information – 2016 – How does NSW compare? [Internet]. Bhi.nsw.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.bhi.nsw.gov.au/BHI_reports/healthcare_in_focus/2016.
12. Medicine Io. Best Care at Lower Cost: The Path to Continuously Learning Health Care in America. Smith M, Saunders R, Stuckhardt L, McGinnis JM, editors. Washington, DC: The National Academies Press; 2013. 436 p.
13. Federation | australia.gov.au [Internet]. Australia.gov.au. 2018 [cited 13 August 2018]. Available from: https://www.australia.gov.au/about-government/how-government-works/federation.
14. Cancer Australia. Cancer Australia Submission Productivity Commission Inquiry – Data Availability and Use. July 2016 [Internet]. Pc.gov.au. 2018 [cited 13 August 2018]. Available from: https://www.pc.gov.au/__data/assets/pdf_file/0019/203653/sub104-data-access.pdf.
15. WA Department of Health. WA Department of Health, Data Stewardship and Custodianship Policy (2016) [Internet]. Health.wa.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.health.wa.gov.au/circularsnew/pdfs/13318.pdf.
16. The Ethics Quagmire: Case Studies | Flying Blind [Internet]. Flyingblind.cmcrc.com. 2018 [cited 28 February 2018]. Available from: https://flyingblind.cmcrc.com/researchers/ethics-quagmire-case-studies
17. National Statement on Ethical Conduct in Human Research (2007) – Updated May 2015 [Available from: https://www.nhmrc.gov.au/guidelines-publications/e72.
18. National Approach to Single Ethical Review of Multi-Centre Research | National Health and Medical Research Council [Internet]. Nhmrc.gov.au. 2018 [cited 13 August 2018]. Available from: https://www.nhmrc.gov.au/health-ethics/national-approach-single-ethical-review-multi-centre-research.
19. PHRN: Training [Internet]. Phrn.org.au. 2018 [cited 28 February 2018]. Available from:http://www.phrn.org.au/news-events/training/.
20. PHRN: Population Health Research Network. Data Linkage Units [Internet]. Phrn.org.au. 2018 [cited 28 February 2018]. Available from: http://www.phrn.org.au/about-us/who-is-involved/australian-data-linkage-units/
21. SURE – Sax Institute [Internet]. Sax Institute. 2018 [cited 13 August 2018]. Available from: https://www.saxinstitute.org.au/our-work/sure/.
22. About the DataLab [Internet]. Abs.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.abs.gov.au/websitedbs/D3310114.nsf/home/CURF:+About+the+ABS+Data+Laboratory+(ABSDL).
23. SURE Users – Summary Steps [Internet]. Saxinstitute.org.au. 2018 [cited 13 August 2018]. Available from: https://www.saxinstitute.org.au/wp-content/uploads/SURE_flow_diagram.pdf.
24. Best and Brightest. Advancing Medical Research 2016. Medical Research Institute Sector Survey Report [Internet]. Professionalsaustralia.org.au. 2018 [cited 17 May 2018]. Available from: http://www.professionalsaustralia.org.au/mri/wp-content/uploads/sites/70/2016/04/Best-and-Brightest-Advancing-Medical-Research.pdf.
25. Productivity Commission. Data Availability and Use, Report No. 82, Canberra. 2017. [Internet]. Pc.gov.au. 2018 [cited 13 August 2018]. Available from: https://www.pc.gov.au/inquiries/completed/data-access/report/data-access.pdf.
26. Bass A, Rosman DL, Smith MB, Semmens JB, Glasson EJ, Brook EL, et al. A Decade of Data Linkage in Western Australia: Strategic Design, Applications and Benefits of the WA Data Linkage System. Australian Health Review. 2008;32(4):766-77.
27. OECD Recommendation on Health Data Governance. 17 Jan 2017 [Internet]. Oecd.org. 2018 [cited 13 August 2018]. Available from: http://www.oecd.org/health/health-systems/Presentation-Health-Data-Governance-Recommendation.pdf.
28. What are clinical registries? [Internet]. Public Health and Preventive Medicine. 2018 [cited 13 August 2018]. Available from: https://www.monash.edu/medicine/sphpm/registries/what-are-clinical-registries.
29. National Health and Medical Research Council Open Access Policy: Open Access Policy: 15 January 2018 [Internet]. Nhmrc.gov.au. 2018 [cited 14 August 2018]. Available from: https://www.nhmrc.gov.au/_files_nhmrc/file/research/nhmrc_open_access_policy_15_january_2018_v2.pdf.

30. Cross Portfolio Statistical Integration Committee (CPSIC). High Level Principles for Data Integration Involving Commonwealth Data for Statistical and Research Purposes.Endorsed by Portfolio Secretaries 3 February 2010 [Internet]. Nss.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.nss.gov.au/nss/home.NSF/533222ebfd5ac03aca25711000044c9e/7afdd165e21f34fdca2577e400195826/$FILE/High%20Level%20Principles%20for%20Data%20Integration%20Involving%20Commonwealth%20Data%20for%20Statistical%20and%20Research%20Purposes.pdf.

31. Privacy Guidelines for the Medicare Benefits and Pharmaceutical Benefits Programsl Office of the Australian Information Commissioner – OAIC [Internet]. Oaic.gov.au. 2018 [cited 13 August 2018]. Available from: https://www.oaic.gov.au/agencies-and-organisations/legally-binding-guidelines-and-rules/privacy-guidelines-for-the-medicare-benefits-and-pharmaceutical-benefits-programs-issued-march-2008-effective-from-1-july-2008.

32. Allen J. and Adams C. Submission to the Produtivity Commission Issues Paper: Data Availability and Use. July 2016 [Internet]. Pc.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.pc.gov.au/__data/assets/pdf_file/0003/203655/sub106-data-access.pdf

33. Development of a Framework for Secondary Use of My Health Record Data: Public Consultation Paper. Commonwealth Department of Health; 2017.

34. Data Availability and Use. Public Inquiry. May 2017. [Internet]. Pc.gov.au. 2018 [cited 13 August 2018]. Available from: https://www.pc.gov.au/inquiries/completed/data-access#report.

35. The University of Melbourne. University of Melbourne submission to the Productivity Commission Inquiry into Data Availability and Use 29. July 2016. [Internet]. Pc.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.pc.gov.au/__data/assets/pdf_file/0007/204559/sub148-data-access.pdf

36. Gunn A, Mintrom M. Measuring research impact in Australia. Australian Universities' Review, The. 2018;60(1):9-15.

37. What's the state of medical research funding in Australia? [Internet]. ABC News. 2018 [cited 13 August 2018]. Available from: http://www.abc.net.au/news/2017-01-26/australias-medical-research-funding-by-the-numbers/8215016.

38. John D Mathews: Submission 36 to Productivity Commission – Data Availability and Use; 2018 [Internet]. Pc.gov.au. 2018 [cited 14 August 2018]. Available from: https://www.pc.gov.au/__data/assets/pdf_file/0003/202989/sub036-data-access.pdf.

39. Mathews J, Forsythe A, Brady Z, Butler M, Goergen S, Byrnes G et al. Cancer risk in 680 000 people exposed to computed tomography scans in childhood or adolescence: data linkage study of 11 million Australians. BMJ. 2013;346(may21 1):f2360-f2360.

40. Libby Roughead. Medication safety; new challenges in the 21st century. lecture delivered at University of South Australia 2013. accessed at https://www.youtube.com/watch?v=gexXGsUXOq0.

41. Pearson S. The big data revolution: What does it mean for Australian pharmacoepidemiology : research? [Internet]. Flyingblind.cmcrc.com. 2018 [cited 17 May 2018]. Available from: https://flyingblind.cmcrc.com/files/researcher-resources/sallie_pearson_-the_big_data_revolution_-what_does_it_mean_for_australian_pharmacoepidemiology_research_.pdf.

42. Pearson S-A, Pesa N, Langton JM, Drew A, Faedo M, Robertson J. Studies using Australia's Pharmaceutical Benefits Scheme data for pharmacoepidemiological research: a systematic review of the published literature (1987-2013): AUSTRALIAN PHARMACOEPIDEMIOLOGICAL RESEARCH. Pharmacoepidemiology and Drug Safety. 2015;24(5):447-55.

43. David Graham On The Vioxx Verdict [Internet]. Forbes.com. 2018 [cited 13 August 2018]. Available from: https://www.forbes.com/2005/08/19/merck-vioxx-graham_cx_mh_0819graham.html#524ee9245698.

44. Krumholz H, Ross J, Presler A, Egilman D. What have we learnt from Vioxx?. BMJ. 2007;334(7585):120-123.

45. What Went Wrong with Vioxx? [Internet]. 2018 [cited 13 August 2018]. Available from: http://www.abc.net.au/radionational/programs/backgroundbriefing/what-went-wrong-with-vioxx/3441410.

46. Sidney S, Cheetham T, Connell F, Ouellet-Hellstrom R, Graham D, Davis D et al. Recent combined hormonal contraceptives (CHCs) and the risk of thromboembolism and other cardiovascular events in new users. Contraception. 2013;87(1):93-100.

47. Brownstein J, Murphy S, Goldfine A, Grant R, Sordo M, Gainer V et al. Rapid Identification of Myocardial Infarction Risk Associated With Diabetes Medications Using Electronic Medical Records. Diabetes Care. 2009;33(3):526-531.

48. Our strategy – NHS Digital [Internet]. NHS Digital. 2018 [cited 17 May 2018]. Available from: https://digital.nhs.uk/about-nhs-digital/corporate-information-and-documents/our-strategy.

49. NHS Health Research Authority (UK) About us [Internet]. Health Research Authority. 2018 [cited 17 May 2018]. Available from: https://www.hra.nhs.uk/about-us/.

50. Integrated Research Application System [Internet]. Health Research Authority. 2018 [cited 13 August 2018]. Available from: https://www.hra.nhs.uk/about-us/committees-and-services/integrated-research-application-system/.

51. HRA Assessment Review Portal (HARP) [Internet]. Health Research Authority. 2018 [cited 13 August 2018]. Available from: https://www.hra.nhs.uk/about-us/committees-and-services/harp/.

52. UK Biobank Ethics and Governance Council. Annual Review 2016/17 [Internet]. Egcukbiobank.org.uk. 2018 [cited 17 May 2018]. Available from: https://egcukbiobank.org.uk/sites/default/files/UKBEGC_Review2016_2017.pdf.

53. UK Biobank [Internet]. Ukbiobank.ac.uk. 2018 [cited 17 May 2018]. Available from: http://www.ukbiobank.ac.uk/wp-content/uploads/2012/09/Access-Procedures-2011.pdf.

54. About UK Biobank | UK Biobank [Internet]. Ukbiobank.ac.uk. 2018 [cited 13 August 2018]. Available from: http://www.ukbiobank.ac.uk/about-biobank-uk/.

55. ACCESS PROCEDURES: Application and review procedures for access to the UK Biobank Resource [Internet]. Ukbiobank.ac.uk. 2018 [cited 13 August 2018]. Available from: http://www.ukbiobank.ac.uk/wp-content/uploads/2012/09/Access-Procedures-2011.pdf

56. Scientific Community | UK Biobank [Internet]. Ukbiobank.ac.uk. 2018 [cited 17 May 2018]. Available from: http://www.ukbiobank.ac.uk/scientific-community/.

57. Executive Order -- Making Open and Machine Readable the New Default for Government Information [Internet]. whitehouse.gov. 2018 [cited 17 May 2018]. Available from: https://obamawhitehouse.archives.gov/the-press-office/2013/05/09/executive-order-making-open-and-machine-readable-new-default-government-.
58. Principles – Project Open Data [Internet]. Project-open-data.cio.gov. 2018 [cited 13 August 2018]. Available from: https://project-open-data.cio.gov/principles/.
59. HealthData.gov Reaches 3,000 Datasets: A Look Back and What's Ahead | HealthData.gov [Internet]. Healthdata.gov. 2018 [cited 13 August 2018]. Available from: https://www.healthdata.gov/blog/healthdatagov-reaches-3000-datasets-look-back-and-what%E2%80%99s-ahead.
60. Summary of the HIPAA Privacy Rule [Internet]. HHS.gov. 2018 [cited 13 August 2018]. Available from: https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html.
61. Patients U. Niva Dorell C, Physician R. Retinal Physician – Using IRIS Data to Set Benchmarks for Retina Patients [Internet]. Retinal Physician. 2018 [cited 17 May 2018]. Available from: https://www.retinalphysician.com/issues/2016/september-2016/using-iris-data-to-set-benchmarks-for-retina-patie.
62. Gebhart F. IRIS Registry already transforming time between science, best practices: infusion of big data in medicine making recognizable difference in how patients treated. Advanstar Communications, Inc; 2015. p. 26.
63. Christopher Kent S. The IRIS Registry: The First 18 Months [Internet]. Reviewofophthalmology.com. 2018 [cited 17 May 2018]. Available from: https://www.reviewofophthalmology.com/article/the-iris-registry-the-first-18-months.
64. Ophthalmology A. Ophthalmology's Data Science Initiative Yields Important Insights on Rare Complications of Common Eye Procedures [Internet]. Prnewswire.com. 2018 [cited 17 May 2018]. Available from: https://www.prnewswire.com/news-releases/ophthalmologys-data-science-initiative-yields-important-insights-on-rare-complications-of-common-eye-procedures-300178878.html.
65. Parke Ii DW, Lum F, Rich WL. The IRIS® Registry: Purpose and perspectives. Der Ophthalmologe. 2017;114(S1):1-6.
66. American Academy of Ophthalmology, 'IRIS Registry Question and Answer', 21 Jul., 2016, accessed at https://www.youtube.com/watch?v=CPVuH1P06hU.
67. Data Sharing – National Medical Research Council [Internet]. Nmrc.gov.sg. 2018 [cited 13 August 2018]. Available from: http://www.nmrc.gov.sg/content/nmrc_internet/home/top_nav/policy---guideline/data-sharing.html.
68. Finland's national health data service benefits citizens, healthcare professionals, pharmacies and the society – FinlandHealth – Finland Health [Internet]. Finlandhealth.fi. 2018 [cited 13 August 2018]. Available from: http://www.finlandhealth.fi/-/finland-s-national-health-data-service-benefits-citizens-healthcare-professionals-pharmacies-and-the-society.
69. e-Health Records — e-Estonia [Internet]. e-Estonia. 2018 [cited 13 August 2018]. Available from: https://e-estonia.com/solutions/healthcare/e-health-record/.
70. Estonia and Finland to start sharing patient data. And that's just the start! — e-Estonia [Internet]. e-Estonia. 2018 [cited 14 August 2018]. Available from: https://e-estonia.com/estonia-and-finland-to-start-sharing-patient-data-and-thats-just-the-start/.
71. Tallinn Declaration on eGovernment at the ministerial meeting during Estonian Presidency of the Council of the EU on 6 October 2017 [Internet]. Ec.europa.eu. 2018 [cited 14 August 2018]. Available from: http://ec.europa.eu/newsroom/document.cfm?doc_id=47559.
72. Priisalu J, Ottis R. Personal control of privacy and data: Estonian experience. Health and Technology. 2017;7(4):441-451.
73. Intel. White Paper: Data Analytics for Unstructured Clinical Case Notes [Internet]. Intel.com. 2018 [cited 13 August 2018]. Available from: https://www.intel.com/content/dam/www/public/us/en/documents/white-papers/clinical-data-analytics-paper.pdf.
74. Harman B.L et.al. Electronic Health Records: Privacy, Confidentiality, and Security. Virtual Mentor. 2012;14(9):712-719.
75. EU GDPR Information Portal [Internet]. EU GDPR Portal. 2018 [cited 13 August 2018]. Available from: https://www.eugdpr.org/.
76. De-identification and the Privacy Actl Office of the Australian Information Commissioner – OAIC [Internet]. Oaic.gov.au. 2018 [cited 13 August 2018]. Available from: https://www.oaic.gov.au/agencies-and-organisations/guides/de-identification-and-the-privacy-act.
77. Cavoukian A. Privacy by design: the definitive workshop. A foreword by Ann Cavoukian, Ph.D. Identity in the Information Society. 2010;3(2):247-51.
78. 1160.0 – ABS Confidentiality Series, Aug 2017 [Internet]. Abs.gov.au. 2018 [cited 13 August 2018]. Available from: http://www.abs.gov.au/ausstats/abs@.nsf/Latestproducts/1160.0Main%20Features4Aug%202017?opendocument&tabname=S#FIVESAFES.
79. CMCRC. Response to Productivity Commission Inquiry into Data Availability and Use: Australian Government Productivity Commission; [cited 2016 6 September]. Available from: http://www.pc.gov.au/__data/assets/pdf_file/0014/203315/sub076-data-access.pdf.
80. Australia's health system: Aihw.gov.au; [cited 2016 7 March]. Available from: http://www.aihw.gov.au/australias-health/2014/health-system/.
81. Digitalhealth.gov.au. (2018). Available at: https://www.digitalhealth.gov.au/about-the-agency/publications/australias-national-digital-health-strategy/ADHA-strategy-doc-(2ndAug).pdf [Accessed 28 Feb. 2018].

