Population Health Research Network (PHRN) Impact and Return on Investment

October 2017

LateralEconomics

CAPABLE, INNOVATIVE, RIGOROUS

I. Executive Summary

Context

Supporting a healthy, long-living and safe citizenry is one of the most important challenges facing Australia, or indeed any nation. Australia spends well over $165 billion annually, or more than a tenth of GDP, on health and related social issues.

Linking of public and other data collections (or data linkage) allows researchers to efficiently generate large, powerful 'big data' analyses that provide rich insight into influences on health, and the safety, quality and costs of interventions – often in ways not previously possible. This research can inform improvements to health and social sector policies, practices and technology, to increase national wellbeing through better health and/or lower health and social services costs (Figure 1).

The Population Health Research Network (PHRN) is a national network of organisations responsible for data linkage infrastructure across Australia. Since 2009, PHRN has been resolving the complex technical, institutional, legal and skill challenges that impede the effective production and use of linked data.

Pathways for Generating Impact

Supported by PHRN resources, systems and processes, jurisdictions are moving towards systematically linking more data collections on a routine basis, starting with more commonly used or requested data collections.

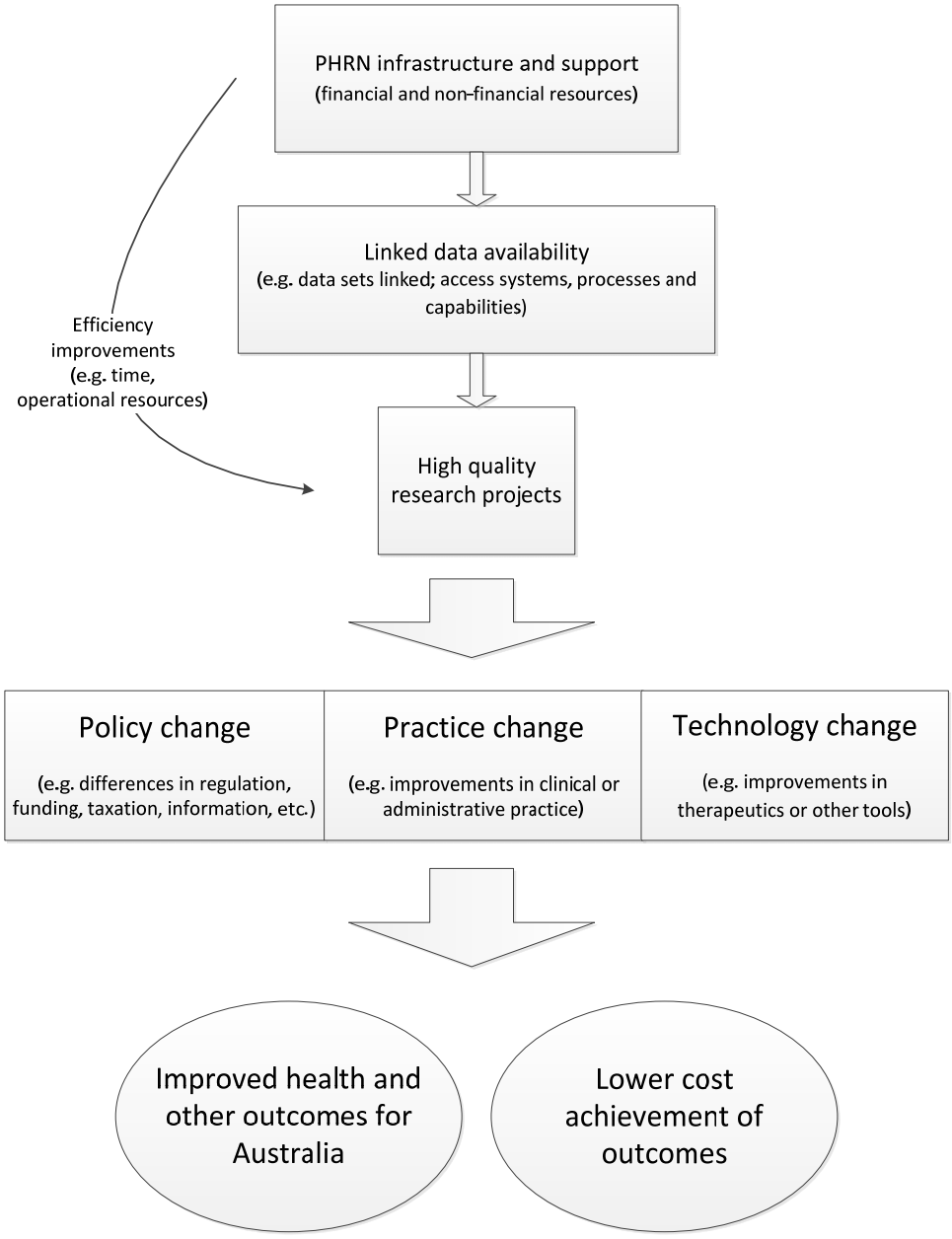
Routine data linkage allows linked data sets to be produced more efficiently, which leads to researchers and other users getting many months' faster access to data, and more timely research results.

There is a clear upward trend over recent years in researchers seeking and, following ethical and other approvals, accessing PHRN-related linked data. This is consistent with new availability of linked data and greater awareness across user groups. Cancer and cardiovascular disease appear to be two of the most significant subjects for formal research projects. The increased power that linked data creates means that the data is useful not just to population health researchers but also now to clinical and policy researchers. New uses for the data are also arising adjacent to pure health research as for instance with research into social issues such as homelessness and child development.

Over the five years 2010-11 to 2015-16, at least 390 peer-reviewed publications have emerged from PHRN-related formal research, with nearly half of these in 2015-16 alone. Government departments are beginning to utilise linked data for both published and unpublished or informal monitoring and analysis in a range of settings.

Stakeholders expect this kind of growth to continue, as further data collections are linked on a routine basis within and across jurisdictions, and as the benefits are further demonstrated. Arguably, we are at the 'tipping point' of moving beyond early adopter projects into consolidation as a mainstream approach, particularly as enduring linkage of rich Medicare Benefits Scheme (MBS) and pharmaceutical data is fully implemented and access lags reduced.

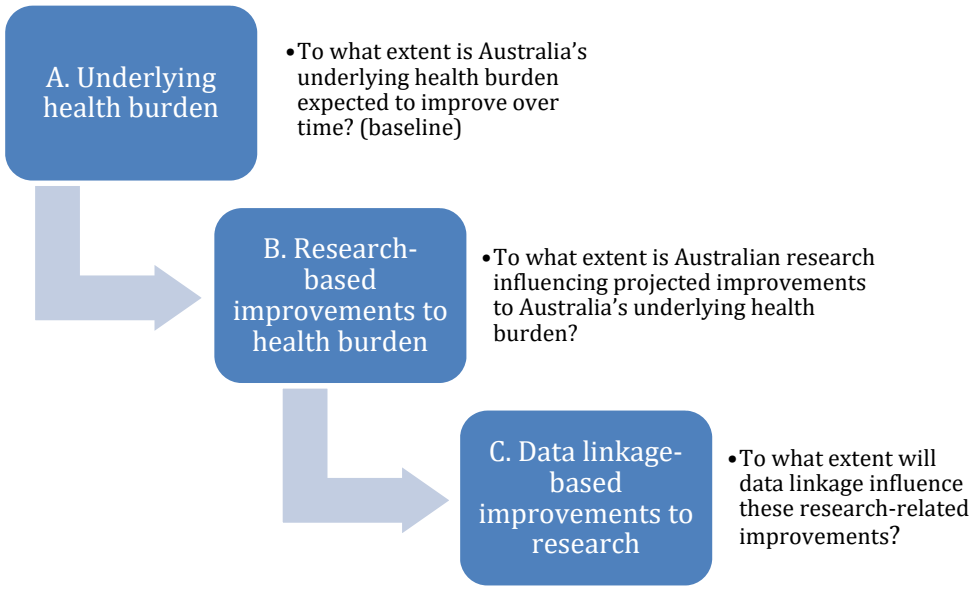
Figure 1 – Overview of 'value chain' of PHRN intended impacts over the longer term



Indicative Economic Analysis

The indicative economic analysis of PHRN's current and future impact considers the cumulative national effect of PHRN infrastructure on health outcomes over time, building on the framework and evidence above. Using a cost-benefit analysis approach, three-cascading questions (Figure 2) provide the structure.

Figure 2 – Cascading logic of economic assessment



To do this, the analysis considers four hypothetical scenarios for PHRN beyond 2018-19 (Table 1), considering costs and benefits to 2040-41. The scenarios reflect different levels of future maintenance and development of data linkage infrastructure, for example in drawing in further data collections and streamlining access. The timeframe allows for a ramp-up of impacts over time as research with linked data is conducted and, over time, influences decision-makers.

Table 1 – Hypothetical scenarios for PHRN resources from 2019

Scenario	Description
Scenario A	All annual resources slightly increase by 2.5% (in real terms) after 2019
Scenario B	All annual resources are maintained (in real terms) after 2019
Scenario C	Participant annual resources are maintained (in real terms) after 2019 but no further resources from the Australian Government through NCRIS or a similar initiative
Scenario D	No further resources for data linkage development after 2019.

Drawing assumptions from literature and stakeholder feedback, Table 2 summarises our best estimates of the net benefit of PHRN.

Table 2 – Summary of net benefit of PHRN by scenario (present value $2017)

Scenario	Benefits	Costs	Net benefit	Foregone net benefit relative to Scenario A	Benefit to cost ratio
Scenario A	$7,588 m	$460 m	$7,128 m	-	16.5 to 1
Scenario B	$6,718 m	$403 m	$6,316 m	$812 m	16.7 to 1
Scenario C	$5,403 m	$330 m	$5,072 m	$2,056 m	16.3 to 1
Scenario D	$2,637 m	$208 m	$2,429 m	$4,699 m	12.7 to 1

The general finding is that PHRN-related data linkage is expected to make the Australian community substantially better off.

Scenario A indicatively suggests a net economic benefit for Australian society from PHRN-related data linkage of over $7 billion, expressed as a present value in 2016-17 dollars.

These benefits reflect PHRN-related research contributing to policy, practice and technology improvements over time. For example, the analysis suggests that by 2034 over 0.53% of cancer burden reduction in Australia will be attributable to PHRN-related data linkage.

Scenario A has a benefit to cost ratio of 16.5, or over $16 in value for Australia for every $1 in cost.

Other scenarios also show strong net benefits. However, the extent of net benefit decreases through scenarios B, C and D, reflecting foregone net benefit relative to scenario A. Under Scenario A, Australians are well over $4.5 billion better off than Scenario D, in present value terms. The difference is less between the other scenarios. There is also arguably a higher degree of uncertainty regarding the realisation of benefits in scenarios that may not include strong coordination of national linkage development.

While we have attempted to be conservative, these indicative results (including the distinction between scenarios) are sensitive to the assumptions used and should be interpreted with care. Sensitivity testing suggests that, for Scenario A, the net benefits could be as high as $17 billion or as low as $1.2 billion under alternative reasonable assumptions. They also suggest that PHRN will generate substantially more benefits than costs to Australia even under systematically pessimistic assumptions for Scenario D which involves the least development of data linkage.

Data linkage is an emerging area of public management and the realisation of much future benefit is highly dependent on how data custodians, user groups in government, academia and the professions, and end-users of research insights, learn and evolve their understanding and practices.

While this indicative analysis provides confidence that PHRN is likely to be a good investment for Australia, future investment priorities and decisions are best made through focused assessment of specific investment proposals and how they seek to accelerate progress across the impact pathway.

Contents

II. Introduction

Background

The Population Health Research Network (PHRN) is a national network of organisations responsible for operating and expanding data linkage infrastructure across Australia. It commenced operations in 2009.

PHRN has a mission "to build a national data linkage infrastructure ...[supporting] research of national relevance which results in improved policy making and service delivery and demonstrates global best practice in maximising the benefits to the Australian community whilst preserving individual privacy."[1]

The network enables researchers across a range of institutions to access and analyse linked administrative and research data collections through proven privacy-protecting data linkage.

PHRN is led by the University of Western Australia (UWA), with nine Participant organisations:

- Australian Institute of Health and Welfare (AIHW)
- NSW Ministry of Health – Centre for Health Record Linkage (CHeReL)[2]
- Victorian Department of Health and Human Services – Victorian Data Linkages (VDL)
- Queensland Department of Health – Data Linkage Queensland Linkage Group
- University of South Australia and others – SA-NT DataLink Consortium
- WA Department of Health – WA Data Linkage Branch
- Tasmanian Data Linkage Unit (Menzies Institute for Medical Research)
- Curtin University – Centre for Data Linkage
- Sax Institute

PHRN received Australian Government national research infrastructure funding of approximately $47 million over the period 2008-09 to 2016-17 inclusive[3], commencing with $20 million through the National Collaborative Research Infrastructure Strategy (NCRIS) programme. This has been complemented with significant cash and in-kind funding from PHRN Participant organisations.

---

[1] PHRN Annual Review 2014/2015

[2] Also incorporating the Australian Capital Territory (ACT)

[3] PHRN has also been allocated NCRIS funding for 2017-18 and 2018-19.

Purpose and scope of this document

Lateral Economics has been commissioned by the PHRN to assess the impact of and return on the PHRN investment since 2009. This includes the impacts to date as well as expected future impacts.

The analysis in this report was developed through synthesising extant documentation, reviewing economic and other literature, and seeking feedback from research, government and sector stakeholders including PHRN Participants, researchers and others.

The report is structured as follows:

- An Executive Summary of key points is Section I.
- Section II is this Introduction providing background and scope of the project.
- Section III highlights the strategic challenge which is the context for PHRN.
- Section IV sets out a framework for conceptualising the impact pathways of PHRN data linkage infrastructure, summarises the outputs it has delivered, and the nature and scale of usage by research groups, government agencies and more broadly.
- Section V furthers this by providing a quantitative-based indicative economic analysis of the society-wide benefits and costs of PHRN, including into the future.

III. Context

The strategic challenge – our growing national health and social burdens

Supporting a healthy, long-living and safe citizenry is one of the most important challenges facing Australia, or indeed any nation.

Diseases, conditions or injuries[4] lead to Australians dying early or living with ill health or disability (our 'disease burden', for short). In 2011, 4.5 million years of healthy life were lost through premature death or living with illness.[5] The five disease groups causing the greatest burden are cancer, cardiovascular diseases, mental disorders, musculoskeletal conditions and injuries, which underpins their priority within significant national and state health initiatives like the National Health Priority Areas.

Social and individual risk factors across childhood, adulthood and later life can influence a person's health status, as well as how diseases are identified, managed or treated.

As a community, we dedicate significant resources to reducing the national disease burden. Australian spending on health was $162 billion in 2014-15 (the most recent year available), reaching 10 per cent of gross domestic product (GDP) (see Figure 3). Over the last ten years, on average, real health spending has grown 4.6 per cent annually.[6] Expectations for high quality health services and treatment, and sometimes the underlying costs of services or treatment, continue to grow.[7]

This same general strategic challenge of growing burdens and public expenditure is common to a range of important social issues across Australia (many with a health interface), like homelessness, or child abuse or neglect.

---

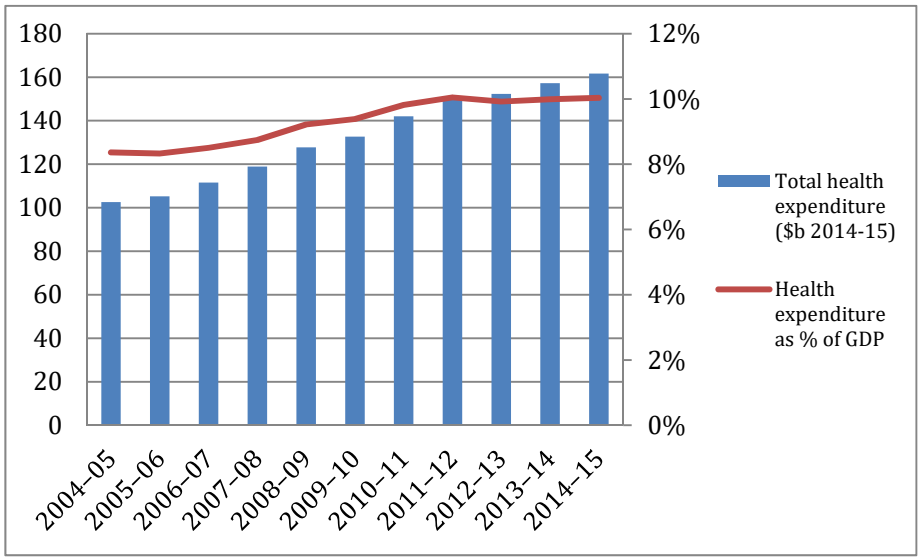
[4] Unless specifically stated, we generally use the word 'disease' to include all three of diseases, conditions or injuries.

[5] AIHW 2016, *Australian Burden of Disease Study: impact and causes of illness and death in Australia 2011*, Australian Burden of Disease Study series Number 3. The notion of disease burden is best illustrated with a simplified example. If a person died from cancer at age 20 in the year 2011, their associated disease burden might be 60-70 years (i.e. they would expect to live healthily to 80-90 if not for the cancer). The overall figure of 4.5 million years in 2011 is equivalent to losing 201 years of healthy life per 1,000 Australians.

[6] AIHW 2016, *Health expenditure Australia 2014–15*, 6 October

[7] For example, real wage growth increasing the cost of labour-intensive health services, or (in gross terms) expensive pharmaceuticals or medical devices.

Figure 3 – Australian health expenditure, constant value



Source: Adapted from AIHW 2016, *Health expenditure Australia 2014–15*, Supplementary tables and figures

Recurrent expenditure on child protection, out-of-home care, family support services and intensive family support services was $4.8 billion nationally in 2015-16 (a real increase of 7.7 per cent from 2014-15).[8] [9] Recurrent government expenditure on specialist homelessness services for 2015-16 was $763.6 million.[10]

Early childhood development education and care is intended, in part, to reduce vulnerability to social or health concerns, as children or later in life. Federal and state recurrent expenditure on early childhood education and care approached $9.0 billion in 2015-16.[11]

Knowledge to help reduce our disease burden

Reducing the disease burden (or other social burdens) – and the cost to government of that burden – can involve preventing disease

[8] Productivity Commission 2017a, *Report on Government Services (RoGS) 2017*, Volume F, Chapter 16 - Child protection services, http://www.pc.gov.au/research/ongoing/report-on-government-services/2017/community-services/child-protection/rogs-2017-volumef-chapter16.pdf

[9] Specifically, $1.2 billion child protection services, $2.7 billion out-of-home care services, $0.38 billion intensive family support services, $0.41 billion family support services (table 1.6A.1)

[10] Productivity Commission 2017a, *RoGS 2017*, Volume G, Chapter 19 - Homelessness Services, 19.3, http://www.pc.gov.au/research/ongoing/report-on-government-services/2017/housing-and-homelessness/homelessness-services/rogs-2017-volumeg-chapter19.pdf

[11] Productivity Commission 2017a, *RoGS 2017*, Volume B, Chapter 3 – Early childhood education and care, http://www.pc.gov.au/research/ongoing/report-on-government-services/2017/child-care-education-and-training/early-childhood-education-and-care

in the first place, or more effective treatment or management. Interventions for these purposes, not just limited to health interventions, can include policy settings, clinical or administrative practices, or technology.

To do this, we need to generate credible and influential knowledge on what factors tend to predict experiencing or avoiding a disease, and what leads to effective treatment or management. We also need to generate knowledge on what actions are most effective relative to cost, to better prioritise our health expenditure. Even a small improvement of the cost-effectiveness of government expenditure would have significant economic benefit.

High quality research of various kinds helps us achieve this. For example, Australian health R&D has been assessed to have an average benefit-to-cost ratio of around 2:1 (that is, $2 of benefit for each $1 of cost), with examples ranging from 0.6:1 to 6:1.[12]

Recurrent expenditure on Australian health research was approximately $5.1 billion in 2014-15 (see Table 3).[13] This is broadly equivalent to around 3 per cent of health expenditure. In addition, there is around a further billion in commercially-oriented health research in the private sector, typically pharmaceutical R&D.[14]

Table 3 – Health research expenditure, Australia, $m 2014-15

Source of funds	$m 2014-15
Australian Government	4,006
State and local government	773
Individuals	3
Other	286
Total	**5,068**
Average annual growth rate (constant prices), 2009-10 to 2014-15	0.7%

Source: AIHW 2016, *Health expenditure Australia 2014–15*, pp. 63 and 68

[12] Access Economics 2008, *Exceptional Returns: The Value of Investing in Health R&D in Australia II*, report for the Australian Society for Medical Research, Canberra.

[13] AIHW 2016, *Health expenditure Australia 2014–15*, October, Table A3, p.63. This includes research: "undertaken at tertiary institutions, in private non-profit organisations and in government facilities that has a health socioeconomic objective. Excludes commercially-oriented research funded by private business, the costs of which are assumed to be included in the prices charged for the goods and services…"

[14] 'Australian H&MR Research Facts, Research Australia (http://researchaustralia.org/australian-research-facts/), drawing from ABS data

The role of data linkage in meeting the strategic challenge

What is data linkage?

Various records about the same subject (e.g. a person) exist in multiple data sets. These are often routinely collected by government and non-government entities such as hospitals, health departments and other departments of state. Traditionally, these administrative data sets have been disconnected silos.

Data linkage is a process to link together – under certain strict conditions[15] – at least some of these records about people, places and events in different data collections. It can create a clearer and more complete picture of the experience of particular individuals or other subjects over time.

Why is it important?

While still in early days, the 'big data revolution' of insight from high-volume, diverse and timely data has the potential for creating rapid change and new discoveries.[16] One recent estimate by Lateral Economics, based on McKinsey analysis, suggested open access to health data in general could have a value to Australia of $5.9 billion per annum.[17]

Better research can result from improved access to data of many kinds, as recently recognised by the Productivity Commission. That data might be generated by governments or the private sector, or via the research sector.[18]

In general, more open access to data can:

- create opportunities for repurposing and re-using data
- stimulate new research networks and collaborations, including by creating greater opportunities for downstream research
- facilitate knowledge transfer to industry

[15] For example, personal identifiers are required for the linkage process, but are typically removed from the newly created dataset used by investigators.

[16] Groves P, Kayyali B, Knott D, Van Kuiken S 2013, "The 'big data' revolution in healthcare: Accelerating value and innovation", McKinsey & Company, Centre for US Health System Reform Business Technology Office, January

[17] Lateral Economics 2014, *Open for Business: How Open Data Can Help Achieve the G20 Growth Target*, June, p. 23. Based on McKinsey Global Institute 2013, *Open data: Unlocking innovation and performance with liquid information*, New York.

[18] Productivity Commission 2017b, *Data Availability and Use*, Inquiry Report No. 82, Canberra, 31 March, p.115

- allow for verification or correction of previous study findings.[19]

Data collections of particularly high value include those that are unique (or cannot be readily replicated), are of high quality, have a high degree of coverage in the relevant population, and are up-to-date or updated regularly.[20] Administrative data relating to health and related areas of social policy tend to be have these high value characteristics. Existing, regularly collected data about individuals is an under-utilised resource for researchers and other agents of influence.

With health-related data linkage, researchers can more easily access and use the large amount of administrative data already collected. This can be the basis for high quality population-level research and evaluation. (Examples given in Table 4)

Table 4 – Examples of research and types of data linkage

Research on issue	Through…
Relationships among personal, economic and lifestyle factors and health	Linking research studies and surveys with outcomes such as emergency department visits, admissions to hospital, cancer notifications and deaths
Societal and community influences on health	Linking health data with information from other agencies, such as education and community services
Safety, quality and costs of health care or other interventions	Linking data on different treatments or interventions with hospital and other outcomes

Source: Adapted from Centre for Health Record Linkage (CHeReL), http://www.cherel.org.au/about-us

Table 5 illustrates the various effects that data linkage can have on research, which variously span efficiency, quality and scope of research. For example, with larger sample groups and the ability to aggregate from specific individuals, such research can more thoroughly and more effectively answer research questions related to the strategic challenge above. Linked health and social data allows researchers to get a more complete picture of the disparate factors that contribute to the physical and mental health of a population.

Ultimately, this improved knowledge enables practitioners to improve the health and wellbeing of Australians and enhance the effectiveness and efficiency of health and social services systems.

[19] Houghton, J 2011, *Costs and Benefits of Data Provision: Report to the Australian National Data Service*, September, Victoria University, Melbourne

[20] Productivity Commission 2017b, p.281

Table 5 – Summary of how data linkage can increase quality and relevance of research

Potential effects	Justification
Improving data quality and integrity	Quality and integrity of data collections improved through the linkage process resolving duplication and other errors, and possibly encouraging more accurate recording of administrative data
Enabling hypothesis generation	Mining of big data can uncover potentially important and previously unconsidered relationships, to be further investigated
Opening new avenues for research to address knowledge gaps	Data linked from sources that are seemingly unrelated allows risk factors and outcomes from different health and social areas, including over time and throughout life, to be examined in the same cohort
Providing larger, more comprehensive samples	With larger samples across the population, rare events, subgroups (e.g. Indigenous Australians) or weaker but more pervasive relationships can be studied (including relating to individuals who have not received a given health service)
Reducing bias	Accessing routinely collected data can reduce selection, recruitment, participation and other biases which can lead to less accurate results
Lowering response burden	Using existing administrative data is less burdensome for citizens than dedicated surveys or other alternative methods to gather data
Improving data handling and confidentiality	Computerised records are generally more secure than paper records, and allow for removal of name and other identifiers

Source: Adapted from Council of Canadian Academies 2015, *Accessing Health and Health-Related Data in Canada: The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation*, pp.50-52; Centre for Big Data Research in Health

Professor Fiona Stanley has highlighted the significant gains could be made with the health budget if government appropriately harnessed linked health data, through reducing costly but ineffective clinical interventions, and detecting and preventing harmful health outcomes through early intervention (see text box over the page).[21] Just as one potential area of investigation, it is estimated that fewer than 5 per cent of the items on the Medicare Benefits Scheme

[21] As reported in Senate Select Committee on Health 2016, *Big health data: Australia's big potential* (sixth interim report), Commonwealth of Australia, 4 May. Professor Stanley is former Australian of the year for her work leading data linkage in Western Australia and using it in epidemiological studies which led directly to lowering of the disease burden in Aboriginal paediatrics.

(MBS) have been assessed for safety, effectiveness and cost-effectiveness against contemporary evidence.[22]

Essentially health-related applications arising from better research look to generate one or more of the following benefits to Australian society:

- reducing the incidence or the burden of particular diseases – through avoiding the problem or fixing it faster
- cost-effectiveness (or value-for money) within the health system, including through moving away from activities that are not effective and towards activities that are, and/or moving towards activities that are effective at a lower cost than alternatives.

[22] Submission from the Centre for Big Data Research in Health (University of NSW) to the Productivity Commission Inquiry into Data Availability and Use

Reflections on data linkage from Professor Fiona Stanley

Speaking in July 2017, noted public health researcher Professor Fiona Stanley said in an Australian Government Tech Talk:

> *"There's a huge number of ways that we can use these data. But the most exciting one that we haven't been able to do [yet] is the linkage of the pharmaceutical data to all of the health data. ..We would be able to say like 'that' [clicks fingers] whether a drug had an adverse effect. We'd be able to look at whether doctors were prescribing accurately. We'd be able to look at harmful effects of all interventions, not just drugs, and that would mean a huge cost saving for the nation and better outcomes for people. ...We'd have a much greater ability to look at cost-effectiveness of the whole system. It's really anguishing for me that we don't have that in place."*
>
> *"This is what's so exciting about ...opening these [Commonwealth] data sets. We could get State data, and we were hungry to link in not just the pharmaceutical data but the Medicare data. To have that for the whole nation means that you would have everybody monitored over time, trends in outcomes, whether services were effective, who was not responding to therapy, who was getting inappropriate investigations... The cost-effectiveness of it would be enormous. If you could scale that up to the nation, it would be a most innovative use of data that you've been collecting for a long time and not used appropriately."*
>
> *"There's some radical changes that would happen [with better linked data access], because ...there would be a real evidence base on which you'd make your decisions ... Everyone's always talking about evidence-based medicine but very few people are actually doing it. Once we get electronic health records, not only will patient care improve, but we as epidemiologists and public health researchers will be able to give you [governments] the very best evidence on what you can do to prevent these problems. Because that's where you're going to get you're biggest bang for your buck, and at the moment we're stopped from doing that. [Plus] getting people in the health sector, and in other areas of child protection, education and so on, to also have an evidence base, to know what they're doing is not harming people and ...is actually being good for the community. I think that's the most exciting way that you could use these data for the future."*

Source: https://www.pmc.gov.au/news-centre/public-data/data-tech-talk-series-launches-today

IV. Impact Pathways

Framework for understanding impact

Data linkage infrastructure is not a final product in itself, but an input into potentially diverse research and other processes. Therefore, to understand its impact,[23] we need to determine how PHRN-related infrastructure and support is used to generate usable linked data sets, how those data sets are utilised to generate new knowledge, and how that new knowledge is applied (or is likely to be applied) for impact in various contexts. In essence, value is driven by how information is generated, processed and used.

We use an 'impact pathway' to articulate this process, moving from inputs, activities and outputs (which can be controlled by PHRN) to outcomes which are directly or indirectly influenced by PHRN-related outputs, to broader impacts on society.[24]

This general 'value chain' of expected long-term impact – from PHRN infrastructure and support through to effects on the Australian community – is depicted in Figure 4.

In the following sections, we consider evidence of the extent to which each step in the 'value chain' is occurring. The last steps of research projects resulting in changes in the world (and associated community benefits) are essentially issues of knowledge translation. PHRN participants can in some ways influence this, but it is not in their direct control.

PHRN infrastructure and support

Certain types of basic health-related data linkage have been operating in limited parts of Australia[25] since the mid-1990s. Access to linked data, when made available following often complex approval processes, was generally through physical media.

Additional resources for data linkage in Australia through PHRN has allowed for establishment (where not already established), and significant expansion and improvement of data linkage infrastructure in all Australian jurisdictions, including the development of systems and processes to support current and future data linkage.[26]

---

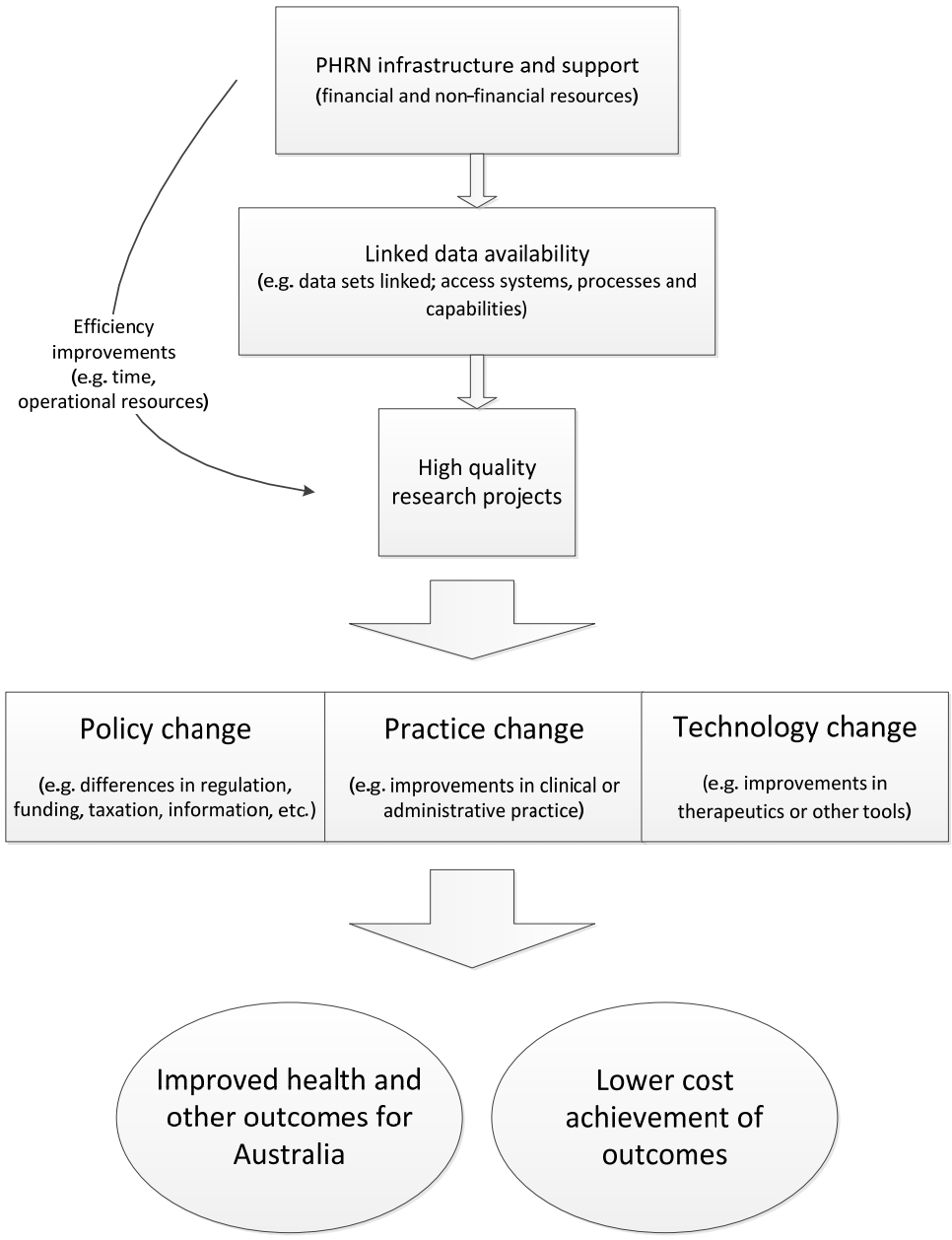
[23] We use language such as 'impact' or 'return' interchangeably; the general idea is the outcomes generated for Australian society as a whole.

[24] This is a common logic model used in, for example CSIRO 2015, *Impact Evaluation Guide*, November.

[25] Western Australia since the mid-1990s and in NSW & the ACT since 2006

[26] Specific details are available in PHRN Annual Reports on its website.

Figure 4 – Overview of 'value chain' of PHRN intended impacts over the longer term



The national PHRN effort has involved investment of resources of around $115 million over the period 2008-09 to 2016-17.[27] This incorporates approximately $47 million funding through various Australian Government initiatives, $23 million cash co-investment from PHRN participants, and $45 million worth of in-kind contributions (principally staff time) from those participants.

In summary, this data linkage infrastructure has incorporated:

[27] Direct PHRN funding from the Australian Government has been confirmed for 2017-18 and 2018-19 at a similar level to that for 2016-17.

- ICT infrastructure and equipment, including to allow for secure remote data access
- workforce training and skills development
- data management and data custodianship methodologies and processes (including to protect personal identity)
- cross-jurisdictional coordination and governance

PHRN nodes act as trustworthy conduits between those who hold data collections (data custodians) and researchers who, subject to approval, receive access to linked data sets for analysis.[28] The principles underpinning PHRN infrastructure development include a collaborative national focus and reducing barriers to access.

Table 6 summarises key elements and the proportion of (total) resources to end 2016-17 for each element.

Table 6 – Summary of PHRN activities/outputs to end 2016-17

Key elements	Description	% of resources
National coordination	• National Coordination (governance, management, communications, strategic priorities, coordination of online application system) through PHRN Program Office / Board	8% ($9.0m)
National Data Linkage Services	• PHRN Centre for Data Linkage (CDL) incorporating data linkage and IT specialists • Australian Institute of Health & Welfare (AIHW)	19% ($22.1m)
Secure Access	• Secure Unified Research Environment (SURE) through the Sax Institute, a remote-access data research laboratory for analysing linked data	12% ($13.6m)
Regional Data Linkage	• Data linkage units and managing nodes for each jurisdiction. Activities to streamline the structures and processes for the delivery of data from custodians through to researchers (e.g. development of jurisdictional linkage keys)	59% ($67.6m)
Network Projects	• Proof of Concept Collaboration • Data Delivery System - CDL • Confidentialisation Project • National Master Linkage Key	2% ($2.7m)

[28] PHRN Education Investment Fund Super Science Initiative (EIF-SSI) Project Plan – Attachment A

Linked data availability

Through this, PHRN has led to:

- data linkage facilities of international standard being established or expanded in Sydney, Melbourne, Adelaide, Brisbane, Hobart, Perth and Canberra; and
- Australian (and international) researchers having access to a range of linked health and health-related datasets across jurisdictions and sectors, complemented by user training and other support to encourage uptake.

Availability of linked data from health (e.g. disease registers, health system data) and other administrative collections has been progressively improving through PHRN-supported efforts.

Data linkages can be created on an ad hoc basis (typically 'create and destroy'), or systemically within a Master Linkage Key.[29] Routine data linkage through a Master Linkage Key, once established, improves the quality and speed of linkage, and the amount of data able to be linked.

Supported by PHRN resources, jurisdictions are moving towards systematically linking more data collections on a routine basis, starting with more commonly used or requested data collections.

As an overview, Figure 5 summarises how many of the states and territories have routinely linked certain data collections within their data linkage unit's master linkage key (as at March 2017). Data collections counted are those routinely linked at least annually and available (subject to project approval) to any non-government researcher applicant. Data linkage units may have routinely linked other data collections for internal purposes.

To illustrate how different jurisdictions across Australia are likely to develop, the linkage system of WA (the largest and earliest in Australia) grew from 4 datasets in 1996 to 47 in 2015.[30]

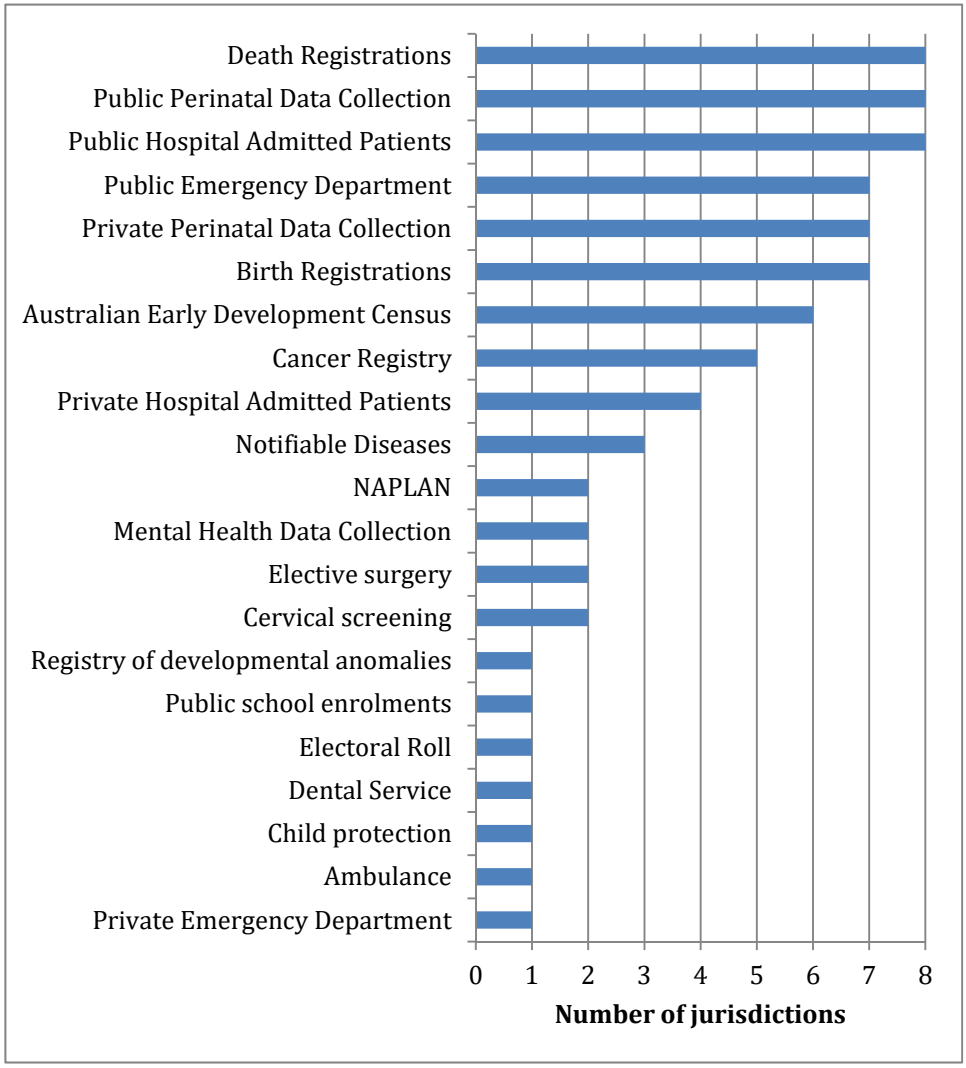
Data collections across jurisdictions are also being linked. For example, one PHRN proof-of-concept project linked patient-level hospital data across NSW, Queensland, WA and SA for the first time, to analysis cross-border hospital use and deaths.

---

[29] Data linkage can be project-based (ad hoc) or systematic. Project data linkage involves the linkage of two or more data sets for a specific project, and does not involve the maintenance of a master linkage file and master linkage key. Systematic data linkage involves the maintenance of a permanent and continuously updated master linkage file (involving linkage variables such as date of birth across multiple datasets) and a master linkage key.

[30] WA Department of Health submission to Productivity Commission Inquiry into Data Availability and Use

Figure 5 – Data collections that are routinely part of jurisdictional Master Linkage Keys (as at March 2017)



Source: Prepared by PHRN, 3 March 2017 based on information available on public websites. The extent to which historical data within these data collections are linked varies. Jurisdictions shown are Australian states and territories. Note: AIHW was not included as it, as of March 2017, had not yet established a routinely linked MLK.

The Commonwealth, through the AIHW, is in the process of establishing a routinely linked master linkage key for some of its data collections. It has, however, allowed access to data including PBS and (since mid-2015) MBS data through bespoke linkages. Resultant data sets, for example linking of immunisation and hospital data to analyse vaccination-based adverse events, were almost impossible to create prior to this.[31]

[31] Notably, a collaborative study linking HPV vaccinations with Victoria's Pap test Register was the first in the world to show a population-based HPV vaccination program resulting in a fall in cervical abnormalities.

Making data linkage routine allows linked data sets to be produced more efficiently, which leads to researchers getting faster access to data. For example, in the NSW experience of July 2012 to June 2014, the median time for full data access was six months faster for Master Linkage Key extracts than for new, bespoke linkages.[32] It allows research results to be produced in a timelier manner (less lag between the research results and the real-world experience the data relates to). It also allows more data linkage to be conducted, which is important given the rising level of user demand.

Researchers and other users sourcing linked data

Potential direct and indirect user groups of PHRN-related outputs are varied (see Figure 6). Users may be in formal public benefit-related research projects such as within university settings, or more informal monitoring or analysis within government.

Figure 6 – Potential PHRN user groups

Researchers or research collaborations	Other direct data users	Change agents (indirect uses)
•in universities (e.g. epidemiologists) •in federal and state government health departments or agencies •in hospitals •potentially, in commercial contexts (e.g. related to pharmaceuticals or medical devices)	•Government agencies or other institutions that undertake monitoring or other data-driven analysis	•Individuals or organisations that utilise (or commission) and apply analysis or conclusions produced by the above groups, including clinicians, health social system planners, social service delivers, or policy communities

Direct users of linked data for formal, public benefit research projects include academics in university settings, clinicians in public and private hospitals, and government agencies. Table 7 shows the institutions that have made applications to AIHW to integrate Commonwealth data from 2016 to mid-2017, as an illustration of typical kinds of users of linked data for formal research. Universities/ research institutes and public entities each account for roughly half of approved access applications nationally.[33] Specific research topics using linked data are wide-ranging (see following section).

[32] Irvine KA and Moore EA 2015, "Linkage of routinely collected data in practice: the Central for Health Record Linkage", *Public Health Research & Practice*, vol 25 no 4, September, p.2

[33] In 2015-16, 49% of approved applications by a 'government agency', 46% by 'universities/research institutes'. A small proportion (6%) are 'other'.

Table 7 – Illustration of applicants for approved Commonwealth data integration projects through AIHW from 2016 to mid-2017

Category	Applicant
Universities	Edith Cowan University Flinders University Griffith University James Cook University Monash University University of Melbourne University of Newcastle University of New South Wales University of Notre Dame University of Queensland University of South Australia University of Sydney University of Western Australia
Hospitals and other health research institutions	Baker IDI Heart and Diabetes Institute Cancer Council Victoria Cancer Institute NSW Concord Hospital (NSW) Doherty Institute Garvan Institute of Medical Research George Institute for Global Health Murdoch Children's Research Institute Royal Adelaide Hospital (SA) Royal Children's Hospital (Vic) Royal Melbourne Hospital (Vic) Royal North Shore Hospital (NSW) Royal Prince Alfred Hospital (NSW) Princess Alexandra Hospital (Qld) Prince Charles Hospital (Qld) Queensland Health QIMR Berghofer Medical Research Institute South Australian Health and Medical Research Institute (SAHMRI) Southern Metropolitan Health (WA) Sydney South West Area Health Service - Liverpool
Other	NSW Treasury

Source: AIHW, http://www.aihw.gov.au/data-integration/projects-2016/ and http://www.aihw.gov.au/data-integration/projects-2017/

Government departments in all jurisdictions are also utilising linkage data for a range of various purposes separate from formal research projects, most of which is not published (or at least is not a clear output of data linkage if published). The majority of internal government (policy) work using linked data is not necessarily transparent through specific publications. Examples include utilising data collections to produce regular, ongoing reporting on specific matters of interest, or internally conducted research on ad hoc or other specific issues.

Anecdotally, researcher and other user awareness is growing. The direct activities of PHRN nodes (e.g. in promoting the resource and user training) are one driver of this. Probably a more significant driver is the cumulative effect of linked data becoming more common in published local research. Other researchers are significant audiences for published local research, who become more familiar with the use of linked data through this channel.

There is a clear upward trend over recent years in researchers seeking and, following ethical and other approvals, accessing PHRN-related linked data. This is consistent with increased availability of linked data and greater awareness across user groups.

Nationally, there are now over 250 project applications annually.[34] For most jurisdictions, this reflects a substantial increase from negligible project applications, compared to very few (although not specifically counted) pre-PHRN.[35] Most are approved – for example, approximately 230 applications received all the necessary approvals in each year from 2013-14 to 2015-16.

Project complexity appears to be increasing.[36] For example, anecdotally where data linkage requests in the past may have been for 3 data sets over 3 years, now there are more requests for linkages of the scale of 20 data sets over 20 years. WA has publicly reported a 3.6-fold increase in highly complex projects over the period 2010 to 2015, against the backdrop of a 1.5-fold increase in

[34] In 2015-16: 254 applications; in 2014-15, 248 applications; in 2013-14, 285 applications. This does not include the WA node. Project applications include research projects as well as projects related to quality improvement, planning, monitoring, auditing and related activities (but not amendments to projects already approved). Source: PHRN Performance Indicator Dashboard.

[35] NSW is an exception, and WA is not counted.

[36] Complexity is a function of technical variables (i.e. new linkages required, data collections involved, cohort groups required, control groups required, manipulation of research data involved) and management variables (i.e. necessary ethics committee and data custodian approvals required, data custodians supplying data, cross-jurisdictional linkage involved). See PHRN's Project Complexity Reporting Guidelines for more detail.

requests received.[37] The proportion of projects across all reporting nodes with a complexity of 4 to 6 (on a scale of 1 to 6) increased from 20% in 2014-15 to 32% in 2015-16. It can be argued that the availability of more complex linked data can drive ambitious, insightful research that would not otherwise occur (given resource constraints). (See Table 8)

Table 8 – Approved PHRN projects by complexity, 2015-16

Project Complexity	Frequency	Proportion	Cumulative Proportion
Category 1 (most simple)	63	27%	27%
Category 2	8	3%	31%
Category 3	87	38%	68%
Category 4	21	9%	77%
Category 5	11	5%	82%
Category 6 (most complex)	42	18%	100%
All	232	100%	-

Note: See footnote 36 for description of project complexity

Anecdotally, users of linked data in population health research are increasingly being supplemented by users in policy and clinical research, as well as early uses outside the health sector. One factor beyond increased interest from these user groups is the increased currency of available data arising from routine linkage: for example, data that is 3 months old is much more useful than data that is 18 months old in many contexts.

Outcomes from utilising linked data

Over the five year period 2010-11 to 2015-16, at least 390 peer-reviewed publications have emerged from PHRN-related formal research. The number of publications is increasing in line with PHRN activity (see Figure 7): 43% of the publications over this time period were in 2015-16 alone.[38]

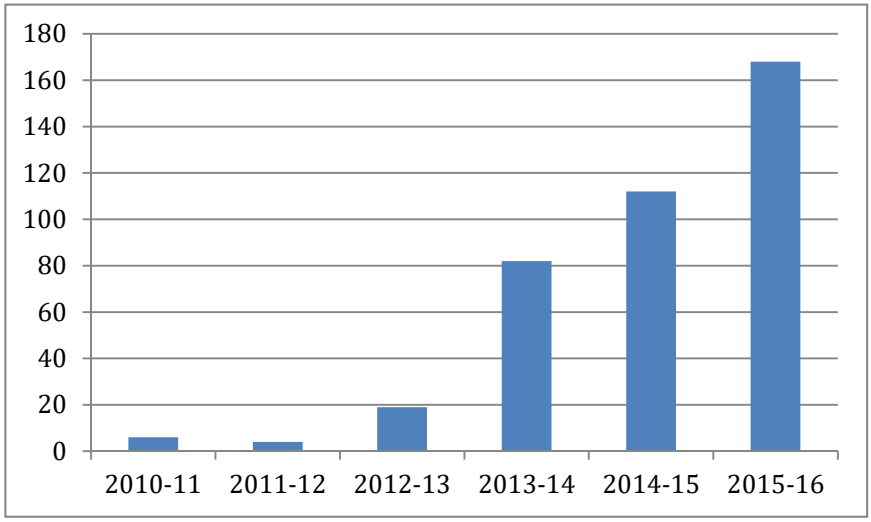
The subject matter covered by PHRN-related research includes:

- health determinants;
- the organisation and delivery of health services;
- health status and health outcomes; and
- non-health fields that affect health and vice versa.

[37] WA Department of Health submission to Productivity Commission Inquiry

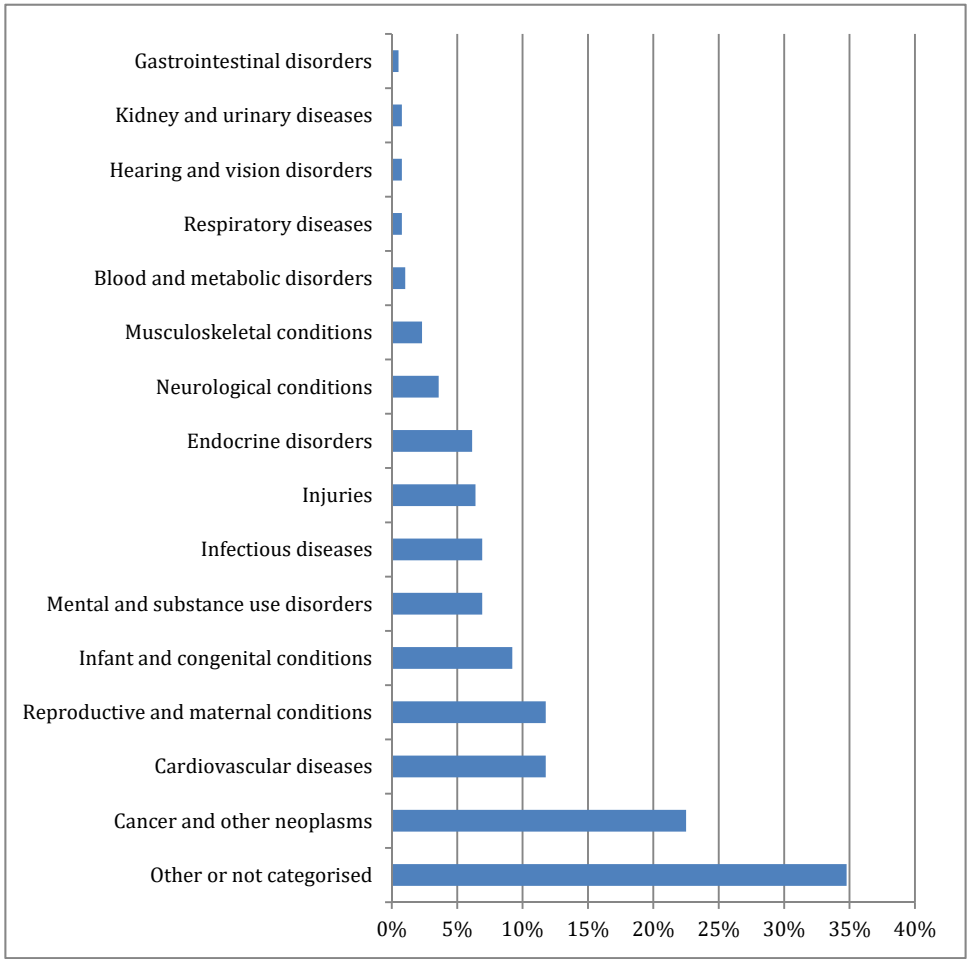
[38] This is only a partial set of PHRN-related publications, drawn from informal analysis of publications by PHRN. The growth pattern over time, rather than the specific count of publications, is the key point to note.

Figure 7 – Indication of growth in formal publications associated with PHRN-related data linkage, over recent years



Note: Only includes formal research projects as reported to PHRN.

Figure 8 – Indicative categorisation of PHRN-related research (2010-11 to 2015-16) to subject matter



Note: Only includes formal research projects as reported to PHRN. Allocation to diseases by Lateral Economics. Projects could alternatively be categorised by target group or other categorisation.

Cancer and cardiovascular disease appear to be two of the most significant subjects for formal research projects (see Figure 8). This aligns with their status as the two most significant contributors to Australia's disease burden. Cancer researchers have, in proportionate terms, been early adopters of PHRN-related data linkage. Recent years have seen fast uptake in cardiovascular and injuries research. Example A (see box at the end of this section) gives one example of the outcomes that can be achieved through linked data.

There is no consolidated list of uses of linked data for research, monitoring or other analysis by government bodies within jurisdictions, some of which is confidential. However, examples given in consultation range from child development to monitoring performance associated with social impact bonds. Example B gives an illustration of how linked data is supporting policy reform and service improvement in early childhood development and child protection.

A series of case studies of recent research using linked data can be accessed at: http://www.phrn.org.au/for-the-community/what-we-have-learnt/more-case-studies/ or https://goo.gl/bz3yCy.

Looking to the future

With the establishment of data linkage systems and processes across Australia with PHRN support, there are various opportunities to further develop data linkage as a tool. This includes the (routine) linkage of further, diverse data collections, and more diverse users and uses, to be able to provide rapid insights on strategic issues of importance.

Some in the sector suggest that we are at the 'tipping point' of moving beyond early adopter projects into consolidation as a mainstream approach, particularly as routine linkage of rich Medicare data through a national master linkage key is fully implemented (expected within the next two years) and further refined to reduce lag between recording of administrative and availability of that data (expected within the next five years).

A number of government stakeholders considered that awareness of the capabilities of linked data in policy and service design communities will mature over the next two to five years, based on work completed, in train and expected. A range of cultural and institutional factors may influence whether such awareness will translate into significantly higher demand for linked data-based analysis, including whether end-users (particularly in government) perceive value from data-rich research.

Possible future extensions of data linkage, building on the fundamentals that PHRN has supported, might include integrating a broader range of data sets, for example:

- linking other health-related data from private health insurers, pathology and medical imaging laboratories, general practices, and possibly genetic and genomic data;
- linking existing health and similar data sets with a broader range of administrative data (particularly those of the Commonwealth) such as social security (Centrelink) customer payment data, higher education and training student data, individual and business taxation data;[39]
- better data on family relationships particularly to consider cross-generational issues (which could be facilitated by Medicare records, for example);[40] and
- exploring ways in which individuals might be invited to vet data that exists on them and contribute to it in various ways, for instance with commentary or addition of their own data for instance biometric data from their smartphones, watches and other accessories and/or social media.

This may facilitate more and different sorts of research, across a range of disciplines.

For example, linked data has the potential to simplify and reduce the time and cost of certain types of clinical trials. Researchers can assess the comparative effects of pharmaceutical or other technology in patients who are already taking them, rather than through recruiting new participants. Further developments might also include greater automation in the collection and linking of data, which goes to more fundamental issues about how our society treats data as a resource and the role of health informatics.[41]

Such extensions would, of course, be subject to appropriate assessment of privacy and other considerations.

---

[39] The Multi-Agency Data Integration Project (MADIP), with the Australian Bureau of Statistics as the Accredited Integrating Authority, which involves integrating census data with other certain Australian Government administrative data sets including Medicare benefit claims, is one example of increased movements towards data linkage for research and analysis at a Commonwealth level.

[40] One example of an application might involve linking health outcomes of a child (e.g. cancer) to prior exposures of the child's mother or father.

[41] One stakeholder pointed to jurisdictions like Singapore which is adopting 'automatic' data collection with initiatives like a Patient managed Kiosk in hospitals to receive the patient and take bloods and observations 'automatically' when arriving at hospital, as part of a holistic health information management practice, to inform monitoring and evaluation of service delivery.

Example A – reducing cancer risk from better understanding of links between CT scans and cancer

Researchers at the University of Melbourne School of Population and Global Health requested access to linked Medicare service data and state-based cancer notification data (through the national cancer register), to study the relationship between certain medical services and cancer risk for young people.

Through AIHW, a linked data set with de-identified Medicare service data alongside cancer data was established.

Resultant research using this linked data showed a significant increase in cancer in young people receiving CT scans: "some 500 additional cases of cancer occurred in the CT cohort beyond what would have been expected from cancer risks in the age-matched cohort of unexposed Australians." The increase was particularly large for brain cancers in children exposed to CT at very young ages. The research also challenged conventional wisdom about the linearity of the dose response curve for cancers due to low-dose radiation, with the risk of cancer, per unit of dose, greater at low doses.

The research conclusions (published in April 2013) were of substantial importance, and become the most highly cited of any BMJ article in calendar year 2014. Current research is aiming to confirm and extend the earlier findings.

The researchers advised the Australian Government Department of Health, the Australian Radiation Protection and Nuclear Safety Agency (ARPANSA), the Royal Australian and New Zealand College of Radiologists (RANZCR) and the Australian Commission on Safety and Quality in Health Care (ACSQHC). The results influenced changes to medical professional guidelines and education about cancer risks and dosage and exposure practices for CT scans especially in children. It could be expected that fewer cancers will result, given these changes.

The lead researcher has stated that the study would be been 'almost impossible' without knowledge of what could be achieved by linking Medicare records to cancer registry records. However, he also noted the five year lag from initial request (prior to PHRN) to release of de-identified data may have deterred other researchers from undertaking the research at all.

This demonstrates the value of linked data in general, and the value of PHRN-supported activities that reduce barriers to linkage and help make data available in a more timely manner.

Source: Submission from Prof John Mathews, School of Population and Global Health, University of Melbourne to the Productivity Commission Inquiry.

Example B – understanding health and development of children to improve life outcomes

Since the establishment of SA-NT Data Linkage, South Australia has developed a rich administrative database of linked data for early childhood research, the SA Early Childhood Data Project (ECDP).

Over 7 million data records are linked for over 300,000 children (including around 12,000 Aboriginal children) born between 1999 and 2014. Records are from more than 30 government sources spanning hospital, child protection, youth justice, education, primary care and dental activities, amongst others. This joining up has never been done before in SA.

The process of data linkage has also improved data quality in relation to consistency between datasets. For example, consultation with the Aboriginal Health Council of SA has resulted in a set of decision rules that uses information from both the Birth and Perinatal data for people of Aboriginal or Torres Strait Islander origin.

The linked data platform is being used by researchers across various disciplines from both universities and SA Government departments. Such new opportunities for research can inform service provision and policy around child health, development and human capability formation from the perinatal period into adolescence, for the benefit of disadvantaged children in particular.

As of October 2016, there were 37 approved researchers, with projects being undertaken for the SA Department of the Premier and Cabinet, Child and Family Health Service, SA Health, Department for Education and Child Development, Families SA, Women's and Children's Health Network, Department for State Development, the Council for the Care of Children, Wardliparingga Aboriginal Research Unit, and the Aboriginal Health Council. A number of publications and reports have resulted.

Child protection-related research is a substantial forward focus following SA's Child Protection Systems Royal Commission which reported in August 2016. The SA Department of Premier and Cabinet's new Early Intervention Research Directorate (EIRD) is partnering with universities for research, including data analytics to better understand the experiences of children within the child protection system, and when, where and how it is most effective for support services to intervene. The newly linked data established through ECDP will facilitate this research. Over time, population-wide data linkage information systems could be used to help quantitatively evaluate service innovations on a routine, cost-effective and sustainable basis.

Sources: University of Adelaide 2016 'The Early Childhood Data Project', School of Public Health, BetterStart Research Series No. 4, October; SA Department of Child Protection, https://www.childprotection.sa.gov.au/department/fresh-start

V. Indicative Economic Analysis of Impacts and Benefits

Introduction

This economic assessment places indicative values on the cumulative effect of PHRN infrastructure across key impacts, building on the framework and evidence in the previous sections.

We use a cost benefit analysis (CBA) approach to quantify, to the extent feasible, the economic value to Australian society arising from PHRN and its utilisation, relative to counterfactuals of what would have been expected if PHRN (or a similar ongoing initiative) did not occur. This includes identifying and valuing the varying impacts of PHRN-related research – typically social cost savings relative to the counterfactual.

CBA is the standard approach in Australia for assessing the *ex ante* or *ex post* net benefit of activities with public funding. It involves a systematic evaluation of the impacts of an activity, accounting for all the effects (to the extent possible) on the community and economy. It provides an objective basis for comparing different impacts and impacts that occur in different periods, and converting impacts into present value dollar terms. The approach can incorporate non-market benefits, i.e. those impacts that do not have an effect on GDP as it is measured but can be expected to affect people's wellbeing. Overall, CBA provides a simple indicator of an activity's net contribution to society. [42]

Applying CBA techniques to an initiative like PHRN is indicative rather than precise, given various uncertainties including the timing and scale of downstream benefits into the future. Given this, we have sought to make conservative assumptions that err towards under- than over-estimation of benefit. With the limitations of economic analysis in this context, this CBA should be only one of a number of inputs that should influence future directions.

Framework

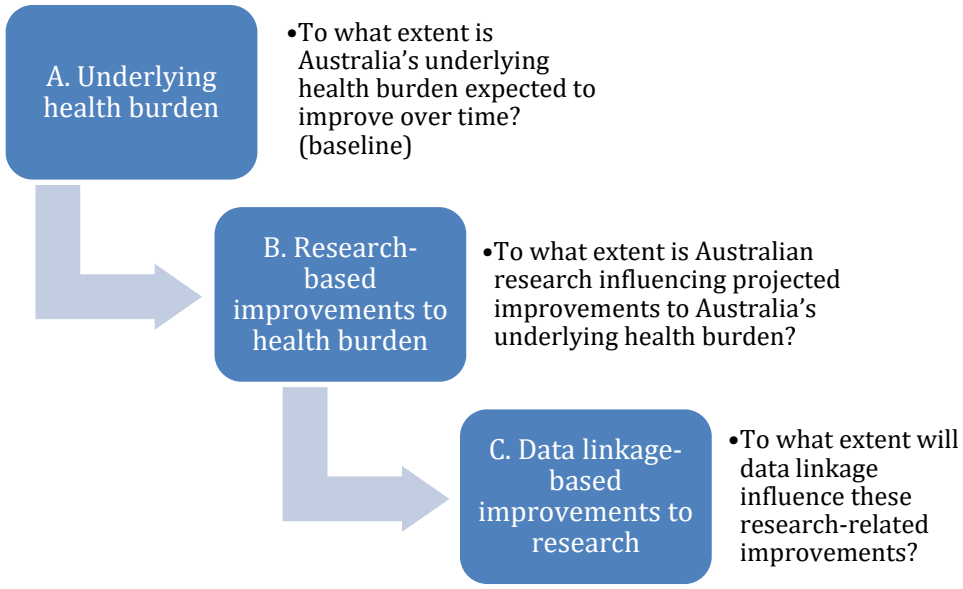
The overall approach to quantifying the benefits (or returns) of PHRN-related linked data is based on three-cascading questions, as illustrated in Figure 9.

While health is not the only focus for PHRN-related research, we focus the analysis on health burden for conceptual and practical reasons. Conceptually, as discussed earlier, there is a relationship between risk factors such as homelessness or a neglected childhood

[42] Australian Government Department of the Prime Minister and Cabinet 2016, *Guidance Note – Cost–benefit analysis*, Office of Best Practice Regulation, February, p.2

and a person's health status across a lifetime. Improvement in these areas would be expected to flow through to improved health and/or lower demand on the health system. Practically, there is good national data available on health outcomes across the community to use as a starting point for an empirical analysis.

Figure 9 – Cascading logic of economic assessment



Our general approach for calculating benefits is, for each major disease group, to:

- create a counterfactual or alternative scenario by projecting the annual disease burden faced by Australians into the future based on current trends;
- estimate an annual reduction in current and future disease burden (i.e. improvement in health outcomes) attributable to Australian health-related research, based on estimates of health improvements from research and Australia's contribution to worldwide health research;
- estimate the contribution of data linkage to that research impact; and
- estimate the economic value of the resultant impact, in present value terms.

A. Underlying health burden

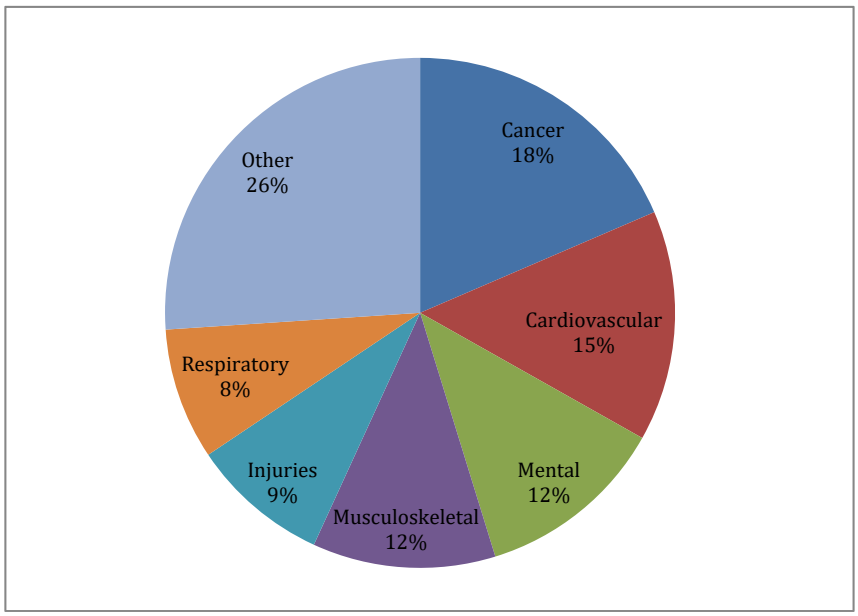
The first step involves projecting Australia's health burden into the future. This is a function of (a) change in underlying health burden of the community and (b) change in population size and structure.

Projecting future disease burden – estimating a counterfactual

AIHW 2016[43] reports Australia's annualised disease burden (as disability-adjusted life years, or DALYs) by major disease groups. DALYs are a consistent summary measure of dying early and living with illness. One DALY is equivalent to one year of healthy life lost from disease/injury.

Nearly three-quarters of Australia's disease burden relates to the six most significant disease groups (see Figure 10). We develop individual projections for five of these groups[44], and assume the remainder follows the trend of these five groups in total.

Figure 10 – Disease burden (in DALYs) by disease group, Australia, 2011



Source: Adapted from data in AIHW 2016

The calculation for cancer is used as an illustration. The projection for other disease groups also follows this method.

Table 9 shows the number of DALYs incurred per 1,000 persons as a result of cancer for 2003 and 2011, and the 1.3% annual reduction over this period. DALY rates are age-standardised by AIHW to remove the effect of changes in age structure over time, establishing the underlying burden independent of population change.

---

[43] AIHW 2016, *Australian Burden of Disease Study: impact and causes of illness and death in Australia 2011*, Australian Burden of Disease Study series Number 3

[44] Mental health is not individually calculated, and is included for the purposes of this analysis within 'Other'. The method used for the other groups is not appropriate for mental health, given the underlying burden of mental health in Australia is slightly increasing rather than decreasing.

Table 9 – DALY age standardised rate (ASR) per 1,000 persons, cancer

Type of burden	2003	2011	Annual change
Fatal (years life lost, YLL)	36.1	32.1	-1.5%
Non-fatal (years lived with disability, YLD)	2.0	2.1	0.6%
Total	38.1	34.2	-1.3%

Source: AIHW 2016, Tables 7.1, 7.2, 7.3

We assume the annual rate of change in underlying burden between 2003 and 2011 continues into the future as a long-term trend. This (overall) reduction is influenced by a range of factors including but not limited to health research.

We combine the calculated ASR in each year into the future with ABS population projections[45] to establish a projection of Australia's disease burden (as DALYs) into the future. This is done for different age and sex groups across the population.[46]

The projected disease burdens that result are shown in Figure 11.

- The upper curve (a stable underlying burden) is Australia's total cancer-related health burden (DALY) at each year in the future, at the age-standardised rate of burden in 2011.[47] It is the scenario if we stay the same at preventing or recovering from cancer after 2011 (e.g. no new early identification behaviours or new treatments). The overall burden increases because population is increasing and, on average, ageing.
- The lower curve is a projection of cancer-related health burden if we continue to improve at preventing or recovering from cancer. It assumes the standardised rate of improvement is the same as between 2003 and 2011 (i.e. 1.3% a year). The overall burden is still growing, but at a slower rate.[48]
- The gap between the two curves is the improvement in underlying burden of cancer over time in each year from 2011.

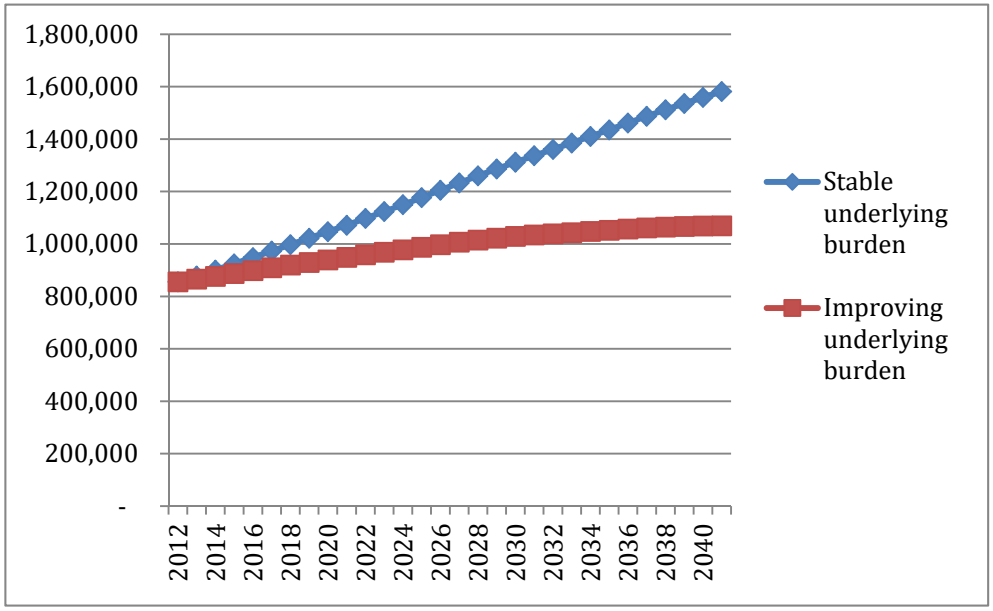
[45] ABS 3222.0, *Population Projections, Australia, 2012 (base) to 2101*, released 26 November 2013

[46] The age groups, for each of male and female, are under 5, 5–14, 15–24, 25–34, 35–44, 45–54, 55–64, 65–74, 75–84, 85–94 and 95+.

[47] The age standardised rate (ASR) provides a measure of health burden (DALY) against the size of the population (per 1,000 people), also taking into account the age structure of the population (in reference to the Australian 2001 Standard Population).

[48] The exact trend is uncertain given this reflects future health innovations and behavioural changes. We consider projecting forward the recent experience is a reasonable position to take for an indicative analysis.

Figure 11 – Projection of DALYs resulting from cancer, Australia



Timeframe

The basic infrastructure (processes, linkage keys, etc. established with PHRN support are long-lived, and are incrementally improved. Similarly, research conclusions may have currency for a reasonably long time period, either directly or through targeting follow-up work. This suggests a longer rather than shorter time horizon for the analysis, to ensure a proper depiction of impacts over time. We assume a timeframe to 2040-41, allowing for the ramp-up in impacts over around 25 years beyond the end of 2015-16.

B. Research-based improvements to health burden

Estimating an annual reduction in current and future disease burden resulting from Australian health research

We firstly attribute the proportion of DALY change over time (i.e. improvement in underlying burden) to changes resulting from global health-related research, relative to other non-research influences.

Health burdens can be reduced by a range of factors: for example, changes in exposure to the risk factors, improved prevention, and advancements in treatment and management including more cost-effective approaches, and better management of services. For example, there may be improvements in medical or health system practice or technology (e.g. diagnostics/screening, surgical procedures, vaccines/pharmaceuticals), and behavioural changes in the population. Each can be influenced by research, in varying ways.

Research knowledge translation and reduced health burden

Pathways to health-related research having an impact on the health and related outcomes of real-world Australians are dependent on 'knowledge translation'. Knowledge translation refers to "the use of knowledge in practice and decision making by the public, patients, health care professionals, managers, and policy makers".[49]

Measuring the extent of knowledge transition is a challenge for many reasons, including attribution issues, time lags between investments and realising impact, and that it is typically not a simple or linear process. A range of cumulative evidence over a period of time (supported by engagement methods of researchers) typically affects the changing of policy or clinical practice, or the experience of one jurisdiction affects changes over time in approaches in another. Moreover, impact on decision-makers is more typically generated from a group of related projects rather than a single project.

Nevertheless, one recent study looking at 36 impact studies of multi-project health research programs worldwide, found that:[50]

- 64% of projects reported some impact of a combined category such as policy and clinical impact (range 60-67%)
    - 35% of projects reported some policy impact (range 5-100%)
    - 32% of projects reported some impact on clinical behaviour or informing clinical practice (range 10-69%)
- 27% of projects reported some impact on wider health impacts (e.g. health gain, patient benefit, improved care or other benefits to the healthcare system) (range 6-48%)

Literature suggests positive average returns to health-related research in the Australian context.

A 2016 assessment for the Australian Academy of Science (AAS) drew on the expert opinion of eminent Australian doctors and medical scientists. This estimated that without the application of knowledge from recent advances in the biological sciences, the burden of disease would have been higher by between 18% (in the low case) and 34% (in the high case)[51], or a mid-point of 26%. Table

---

[49] Straus SE, Tetroe, JM, Graham, ID 2011, "Knowledge translation is the use of knowledge in health care decision making", *Journal of Clinical Epidemiology*, vol 64, issue 1, January 2011, pp.6-10

[50] Hanney S, Greenhalgh T, Blatch-Jones A, Glover M and Raftery J 2017, "The impact on healthcare, policy and practice from 36 multi-project research programmes: findings from two reviews", *Health Research Policy and Systems*

[51] Australian Academy of Science (AAS) 2016, *The importance of advanced biological sciences to the Australian economy*, prepared for the Office of the Chief Scientist and the Australian Academy of Science by the Centre for International Economics, p.28

10 breaks this down by specific disease groups. Note that advanced biological sciences would only be a subset of health-related research.

Table 10 – Estimated impact of recent advances in biological sciences on the burden of disease in Australia

Disease group	Lower bound	Upper bound	Mid-point	Mid-point ratio [52]
Cancer (malignant neoplasms)	27%	54%	41%	1.6
Cardiovascular disease	35%	40%	38%	1.4
Mental disorders	0%	28%	14%	0.5
Nervous system and sense organ disorders	15%	26%	21%	0.8
Chronic respiratory disease	16%	32%	24%	0.9
Diabetes mellitus	2%	4%	3%	0.1
Other diseases and conditions	14%	28%	21%	0.8
All diseases and conditions	18%	34%	26%	1.0

Source: First two columns from AAS 2016 p.29, drawing on previous research and expert opinion

This is broadly consistent with findings of other literature (as context even if not directly comparable);

- that the number of years of potential life lost before age 80 in 2011 would have been 22% higher if no new items had been listed in the PBS between 1989 and 2002[53]
    - noting that medical advances are more than pharmaceuticals alone
- that about one third of reduction in mortality from cardiovascular disease is due to invasive treatments, one third from pharmaceuticals, and the remaining third from behavioural change[54]
    - which Deloitte Access Economics in various reports use as a foundation for the assumption that 50% of

[52] The mid-point ratio is the ratio of the mid-point of the disease group to the mid-point of 'all diseases and conditions'. For example, 41% (cancer) divided by 26% (all) equals 1.6.

[53] Lichtenberg FR 2017, "The impact of pharmaceutical innovation on premature mortality, hospital separations, and cancer survival in Australia", *Economic Record*, vol 93 issue 302, September (as reported in 2015 version in AAS 2016)

[54] Cutler and Kadiyala 2003. Deloitte Access Economics' 2016 report *Australia's health and medical research workforce* for the Australian Society for Medical Research is one example of where this is discussed.

improvements in healthy lifespan is attributable to health and medical R&D.

Taking into account the literature above, we attribute 50% of health improvements on average to global research across disciplines. We then adjust this for individual disease groups consistent with the mid-points of the per-disease assessment for the AAS. The impact for some diseases is below 50%, and some above. This is summarised in Table 11.

Table 11 – Attribution of health improvements to global research

Disease group	Adjustment ratio	Attribution of improvements to research
Cancer	1.6	78%
Cardiovascular	1.4	72%
Respiratory	0.9	46%
Musculoskeletal	0.8	39%
Injuries	1.0	50%
Others	1.0	50%

Proportion of research impact in Australia from Australia

Based on bibliometric analysis, Deloitte Access Economics (2016) estimates that Australia produced 3.8% of world health and medical research output in 2012 (an increase from 2.5% in 2002).[55]

Although economic models often assume that information flows freely, we know that in many areas that is not the case. Indeed, even where the transactions costs of alternative investments are relatively equal in different countries, it is remarkable how much home market bias there is in our equity markets. Anyone from around the world can invest in CSR, Cochlear or Telstra or one of our banks and good research on all those companies is available to anyone wishing to invest the time and/or money. Simple portfolio theory suggests that Australian firms are unusually attractive to firms in other countries for the extent to which they can diversify risks for investors most exposed to other countries. And yet most of the investors in these companies are Australian.

Accordingly we would expect to see an analogous 'home market bias' in the use of research. Australian research is likely to have substantially more relevance to Australian clinical practice and policy

[55] Deloitte Access Economics 2016, *Australia's health and medical research workforce: expert people providing exceptional returns*, report for the Australian Society for Medical Research, 19 October 2016, p.10

actions than overseas research. Looking at 19 Australian cancer clinical guidelines and 4,700 cited references to journal literature over the period 1981 to 2008, Penn and Webster (2009) found that Australian-authored papers played a disproportionately significant role in these Australian guidelines. Australian-authored research was cited roughly 4.8 times its proportion of world research. [56] Given that Australia's proportion of world research is 3.8% (according to Deloitte analysis), we assume the proportional impact of Australian research on Australian practitioners is around 18% (i.e. 4.8 * 3.8% = 18%).

We expect there to be some home market bias making local research in technology development and adoption more influential in Australia than foreign research. Even where it is not technically more relevant, the proximity of researchers is likely to somewhat improve local take up of new locally won knowledge. In the absence of specific evidence on the size of home market bias, we have assumed conservatively that the Australian impact of Australian pharmaceutical and technology research is proportional to Australia's share of world health and medical research (3.8%).

We therefore assume that Australia's overall contribution to the impact of world research in Australia (as discussed in the previous section) is equivalent to 10.9%, based on Deloitte's analysis (see previous section) that invasive treatments and pharmaceuticals have a roughly equal contribution to health improvement.[57]

Therefore, on average, 5.5% (10.9% multiplied by 50% from the section above) of Australia's change in disease burden is attributable to Australian health-related research, although it varies by disease (as per Table 12). For context, the assumption means that 19 out of 20 improvements in Australian health outcomes result from overseas research (for example, a new pharmaceutical product developed overseas) or factors other than research. We think this is a conservative approach.

[56] Webster BM and Penn D 2009, "What is the evidence in evidence-based practice? Citation analysis of papers referenced in Australian cancer clinical practice guidelines", presentation to Positioning the Profession: the Tenth International Congress on Medical Librarianship, September, see https://espace.library.uq.edu.au/view/UQ:184760

[57] The mid-point of 18% and 3.8% is 10.9%.

Table 12 – Proportion of DALY improvement attributable to Australian research

Disease group	Proportion of improvement
Cancer	8.5%
Cardiovascular	7.9%
Respiratory	5.0%
Musculoskeletal	4.3%
Injuries	5.5%
Others	5.5%

C. Data-linkage based improvements to research

Estimating the contribution of data linkage to Australian research impact

The next issue is the relative importance of linked data, over time, on Australian research impacts.

Data linkage-supported research, by its nature, focuses on Australians' actual experiences with the health system (in terms of who gets care, when, how and with what results) — and, in many cases, how this can be continually improved. Generating relevant local knowledge in this way is a clear pre-condition for knowledge translation.

The main areas of economic benefit of PHRN relate to:

- resource efficiencies (e.g. reduced time/cost of accessing and using data) in research that would otherwise occur; and
- more significantly, the returns (benefits) achieved from additional research conducted.

Resource efficiencies in research

The key resource efficiency benefit of linked data is the saved time and cost when re-using standardised data that is originally collected for administrative or operational reasons.[58]

Calendar time is certainly being saved, as shown in part IV. Less time to create a data set means the period from research idea to research results is shortened. This also increases the currency and relevance of results to stakeholders.

[58] Especially where many of the costs of making data linked are only incurred once.

Labour effort (and cost) is also saved, probably most through linked data being an alternative to methods like individual longitudinal field studies, if those methods are even practically feasible with given resource constraints.[59]

Typically, resource efficiencies in research will get re-invested into further research, as the aggregate funding envelope for health-related research is typically externally determined. For example, a researcher might undertake more complex projects (that lead to higher quality results) – evidenced, in part, by more complex data set requests described in part IV. Or they may undertake further research to extend earlier findings. Therefore, to avoid double-counting, we do not estimate this effect separately to the returns achieved from additional research.

Impact of data linkage on research returns

In assuming the influence of PHRN-related data linkage on health outcomes, there are a number of considerations:

- the pattern of adoption of linked data in research over time, taking into account
  - how we expect researchers will incorporate linked data into research methodologies into the future (data linkage adoption)
  - the proportion of health-related Australian research using data linkage when data linkage adoption is mature;
- the level and timing of change in research effectiveness, taking into account:
  - the pattern of adoption (from above)
  - the level of improvement in research effectiveness/quality when adoption is mature
  - the time lag from undertaking research to that research having a real-world effect on health outcomes.

Scenarios

A key issue is what scenario for PHRN activities in the future should be assumed.

PHRN's strategic planning highlights the need to make Australia's data linkage infrastructure more efficient, linked data more

[59] The Sax Institute's long-term '45 and Up Study' has recruited more than a quarter of a million people aged over 45, and is the largest ongoing study of healthy ageing in the Southern Hemisphere. While '45 and Up' is broader than just health issues, it illustrates the kind of data collection process that could be obviated in part of full by more efficient access to linked public health data.

accessible and to meet the emerging needs of the research community, governments and industry. This includes, for example, working collaboratively with stakeholders to further build system capacity (including progressing enduring linkage of Commonwealth and Commonwealth/state/territory data, and including new data collections in national and regional linkage systems) and grow demand.[60] The achievement of these aspirations will depend to a large degree on available resources.

We identify four broad hypothetical scenarios for the future of PHRN, each with different scale of activity and associated costs (Table 13). We do not suggest these are the only scenarios available for the future of PHRN, or make any comment as to their desirability.

Table 13 – Hypothetical scenarios for PHRN beyond 2019

Scenario	Description
Scenario A	All annual resources slightly increase by 2.5% (in real terms) after 2019
Scenario B	All annual resources are maintained (in real terms) after 2019
Scenario C	Participant annual resources are maintained (in real terms) after 2019 but no further resources from the Australian Government through NCRIS or a similar initiative
Scenario D	No further resources for data linkage development after 2019.

Benefits will naturally vary between these different scenarios, given the varying activities involved and resultant effects for users.

Adoption of linked data by researchers

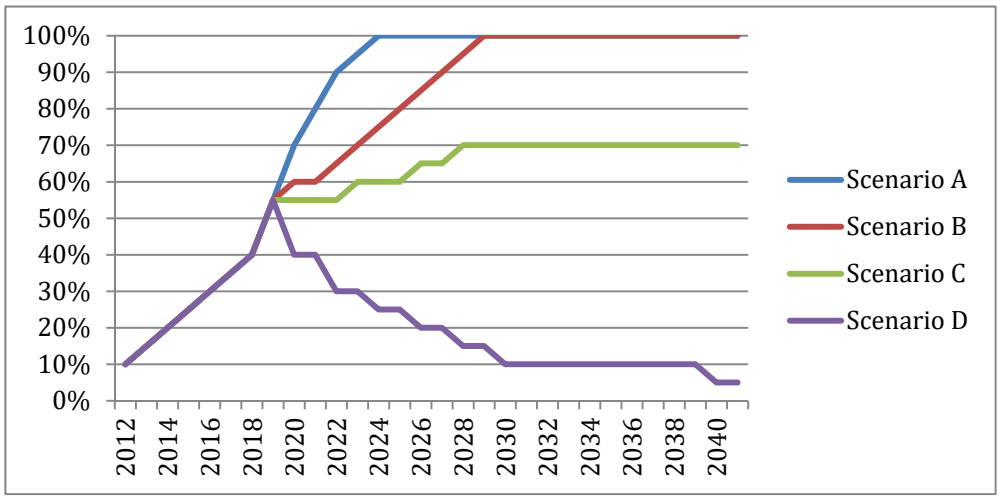
We know that levels of research with linked data is growing year on year (part IV), and there is clearly large potential for future use of data linkage for a range of research-driven purposes.

The ramp-up to a mature position as further linked data becomes available and researchers become more competent in its use, if that is achieved, varies by scenario. Scenarios vary through the assumed level and rate of 'data linkage research maturity'. In Scenarios A and B, resources are available to maintain and increase data linkage (for example, to incorporate further data collections, streamline processes and update technology) to maturity (with maturity modelled as 100%). The assumed growth paths in Figure 12 are consistent with growth patterns to date and in line with stakeholder feedback that they expected the usage of linked data in Australian research to more than double over the next 5-10 years, subject to continuing improvement effort, and that 'peak usage' of linked data by Australian health researchers (based on awareness and capability) is some years away. In Scenarios C and D, we

[60] *PHRN Strategic Plan 2017-2026* (June 2016)

expect some diminishing over the longer term in adoption as existing capabilities become less relevant or useful in line with diminishing resources to support the process of data linkage (e.g. ethical approvals, maintenance of systems).

Figure 12 – Expected adoption path of linked data in research towards 'maturity'



Based on stakeholder feedback, we assume 25% of Australian health-related research data will utilise PHRN-related linked data when data linkage is 'mature' (i.e. evolves into a mainstream methodology utilised across a range of research questions in university and government contexts).[61]

Expected improvement in impact

Data linkage can result in higher quality research projects – those with more ambition and greater specificity. A number of factors influence this, summarised in Table 5 in part III. For example, analysis and conclusions based on population-wide data rather than samples or other observations can be more robust, and more relevant to decision-makers.

Also, as discussed previously, assuming any time saved would be used to do more research, we can add the returns to the additional research – effectively, treating the efficiency saving as additional expenditure on additional research.

The size of this overall marginal improvement from linked data is somewhat uncertain, and there is little established literature to guide assumptions. For example, Canada's Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation "found many examples of compelling and

[61] This might be considered linked data-based research across Australia minus types of data linkage projects that may have otherwise occurred in a counterfactual (e.g. particularly for government purposes within WA and NSW).

pertinent research around the world" from linked data, but found it "difficult to quantify the resulting benefits."[62] We assume, on average, the adoption of linked data gives a 20% impact improvement over alternatives, which represents a mix of quality, scale and timing effects. This assumption draws on a recent economic evaluation of Australian clinical quality registries which examined the effects of better information on specific measures of performance compared to baselines. The impacts from various registries, under conservative assumptions, ranged from a 12% reduction in the incidence of cancer returning following surgery (positive surgical margins rate),[63] a 21% reduction in active intervention in low risk patients, a 23% reduction in average length of stay, and a 15% reduction in dialysis mortality rate.[64]

We adjust this by disease group, based on the observed intensity of use of linked data to date in research regarding particular diseases (see Table 14). For example, there is a higher degree of improvement in cancer (given more research using data linkage) and a lower degree in musculoskeletal conditions.[65]

Table 14 – Adjustments to 'peak' impact improvement from data linkage

Disease group	Proportion of PHRN-related publications compared to proportion of disease burden	'Peak' improvement from data linkage
Cancer	More	25%
Cardiovascular	Similar	20%
Respiratory	Less	15%
Musculoskeletal	Less	15%
Injuries	Less	20%
General (all other disease groups)		20%

[62] Council of Canadian Academies 2015, *Accessing Health and Health-Related Data in Canada*, The Expert Panel on Timely Access to Health and Social Data for Health Research and Health System Innovation, Ottawa

[63] Positive surgical margin is a technical term referring to margins of a tumour showing cancer cells.

[64] Australian Commission on Safety and Quality in Health Care 2016, *Economic evaluation of clinical quality registries – Final report*, prepared by Monash University and Health Outcomes Australia, November

[65] We compare, for individual disease groups, the proportion of PHRN-related research publications related to that group (to 2015-16) with the proportion of national disease burden from that disease group. These proportions were discussed in section IV. If there is proportionally a higher degree of research, we increase the expected level of improvement related to data linkage for that disease group.

Timing lag from research impacts to impacts

Real-world impacts do not occur immediately. There is a lag between research conclusions and influence on decision-makers, and a lag between changed policy, practice or technology affecting individuals.

De Oliveira *et al* 2013 estimate a mean time lag of 12.8 years (with a standard deviation of 4.0 years) for health-based academic research.[66]

That being said, practical research commissioned by decision-makers (e.g. government) to help guide decisions regarding a specific policy issue might have a more immediate impact. For the purposes of this conservative analysis, we assume a lag of 10 years before PHRN-related research has any impact on Australian health outcomes.

Overseas impacts

As this analysis is framed as the benefit for Australia, it does not consider how other countries may benefit from Australian research utilising linked data.

However, clearly there would be some degree of reputational benefit to Australia (with possible further positive effects on attracting and retaining human capital locally) if Australian research-driven innovation has adoption worldwide, or if Australia participates in international collaborations in the development and utilisation of data linkage infrastructure.

Gross benefit results

The combined effect of the above drivers suggests that, for cancer under Scenario A, by 2034 over 0.53% of the improvement in health burden associated with cancer in Australia will be attributable to PHRN-related data linkage. This is illustrated in Figure 13, with more detail on the other scenarios in Table 15.

[66] de Oliveira C, Nguyen HV, Wijeysundera HC, Wong WW, Woo G, Grootendorst P, Liu PP, Krahn MD 2013, "Estimating the payoffs from cardiovascular disease research in Canada: an economic analysis", *Canadian Medical Association Journal* (CMAJ) Open, July

Figure 13 – Combined effect of data linkage on decrease in Australian DALY burden from cancer (scenario A)

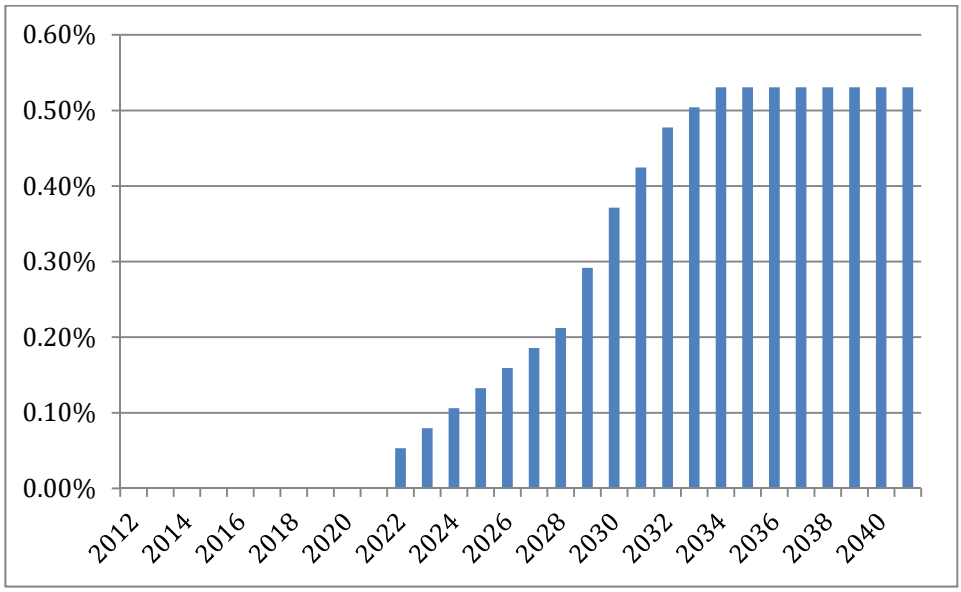
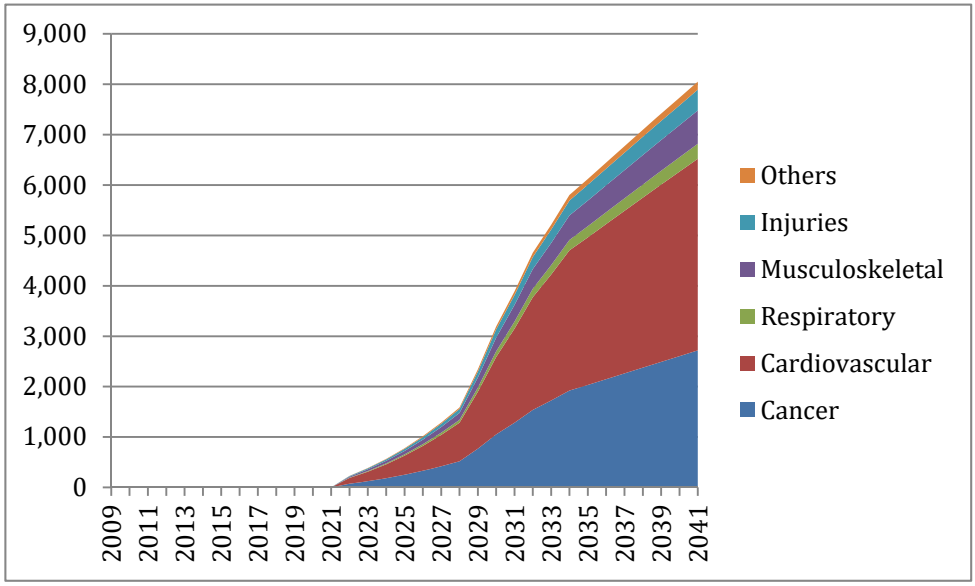


Table 15 – Illustration of highest level of influence on cancer improvement and year reached

Scenario	Highest level of influence	Year reached
Scenario A	0.53%	2034
Scenario B	0.53%	2039
Scenario C	0.37%	2038
Scenario D	0.29%	2029

In terms of the DALYs avoided, the pattern (for scenario A) by disease group over time is shown in Figure 14.

Figure 14 – Avoided DALY from PHRN-related data linkage (scenario A)



Monetising the resultant impact in present value terms

Value of a statistical life year

To translate a DALY into an economic value, we need the economic value of a statistical life year (VSLY). A VSLY can be interpreted as the value of a year of life free of injury, disease and disability.

Drawing from literature, we assume a constant VSLY of $191,000 in 2016-17 dollars. This uses the advice in DPMC 2014[67] which recommends that Dr Peter Abelson's (2008)[68] empirical estimate of $151,000 (in 2007 dollars) be used for a VSLY, and adjusts that value for inflation (the Consumer Price Index) since 2007. The size of this value is arguably very low compared to some more contemporary assessments (particularly from the United States), however we take the conservative approach of consistency with Australian Government guidance.

Some literature suggests that a VSLY should vary over a lifetime (that is, the annual value of a year of a younger person's life should have a different value to that of an older person's). Abelson 2008 argues that, while arguments are inconclusive and evidence is thin, it seems preferable to use a constant VSLY instead of a constant

[67] Australian Government Department of Prime Minister and Cabinet 2014, *Best Practice Regulation Guidance Note – Value of statistical life*, December

[68] Abelson, P 2008, "Establishing a Monetary Value for Lives Saved: Issues and Controversies", WP 2008-02, Working papers in cost-benefit analysis, Office of Best Practice Regulation, Australian Government Department of Finance and Deregulation

value of a statistical life (VSL) (a constant VSL implicitly has a VSLY that varies based on age). This means that each year of a person's life, no matter what their age, is treated equally. We adopt this approach.

Applying discount rate for present value[69]

The economic principle of time preference recognises that society generally places a higher preference on receiving a benefit at an earlier date compared with received it at a later date. The same logic applies to costs. Consequently, a social discount rate is applied to benefits and costs in each relevant year to discount future costs and benefits (and compound past costs and benefits), to account for the social rate of time preference.

To simplify somewhat, one can distinguish between two ways of determining the appropriate discount rate. The social rate of time preference (SRTP) reflects the rate at which society is willing to forgo current consumption in return for more consumption in the future. One can take the rate at which society can borrow money through its government as one measure of this.[70]

The other method takes into account the opportunity cost of the resources committed to a project, which can be measured by the social return on their next best use. Any money governments spend on a project could equally be placed in investments like equities with high expected rates of return.

Although some of the considerable margin between the interest rate at which governments can borrow and the expected rate of return on a balanced portfolio of equities represents a risk premium, the general, but by no means unanimous conclusion of domain experts is that a large part if not most of the debt-equity premium represents market failures of various kinds. As Simes and Gruen argue this discloses an opportunity for governments to arbitrage this margin by borrowing to invest in high return assets whilst managing for prudent risk taking, including valuing the additional risk taken on.[71]

---

[69] Thanks to John Quiggin, Ric Simes and George Argyrous for assistance in considering these arguments in this section, though none of them should be held responsible for our reasoning.

[70] Note this abstracts from risk – both the risk premium those lending money to the government apply in determining whether to lend and the risk premium that government should consider given the riskiness or otherwise of the project it intends to fund. We take this up below. It also abstracts from the many ways in which governments influence the wider financial and monetary environment in which they operate.

[71] Simes, R and Gruen, N 2003, "Risk Management and the Re-Invigoration of Reform: A paper delivered to the One-Day Symposium: Debt, Risk and the Role of Government: The bond market in a wider context, 27 February, ANU, available at https://lateraleconomics.com.au/wp-content/uploads/2014/02/Risk_and_reinvigorating_reform.pdf.

Advice from government on the appropriate discount rate to use is likewise varied. The Australian Government's Office of Best Practice Regulation (OBPR) focuses on the opportunity cost of funds and suggests a standard annual real discount rate of 7 per cent for assessing regulatory interventions, with sensitivity tests at 3 per cent and 10 per cent.[72]

The Victorian Department of Treasury and Finance[73] suggests that a 4 per cent real discount rate is appropriate for provision of goods and services in traditional core service delivery areas of government, such as public health, justice and education, based on a risk free rate plus a very small risk premium.

We follow the Victorian Government's guidance of 4 per cent for several reasons. Firstly we have experienced much lower rates of return on both debt and equity since the global financial crisis which suggests government discount rates should be adjusted downward as has occurred in some jurisdictions such as New Zealand.[74]

Second, abstracting from risk, the decision as to whether one uses the rate required to raise funds or the rate one could get for a portfolio of high return assets is to some extent arbitrary, with a 'first best' policy arguably involving borrowing up to some point at which the (risk adjusted) debt equity premium disappeared. In these circumstances we can distinguish between the use of a discount rate to set a hurdle rate of return, beyond which one decides to proceed with a project, and the use of a discount rate to calculate the total benefit-cost of a project if it goes ahead.

There is always a case for investing in a portfolio of high return assets ahead of a project that generates lower expected returns than this at comparable risk. However once it is clear that the project one is considering is clearly very advantageous at the more conservative higher discount rate – which, as shown in subsequent sections, investment in PHRN clearly is – estimating its value can legitimately be done at the lower 4 percent real rate of discount. Nevertheless, we also test 7 per cent and 3 per cent within low and high sensitivity test bundles (see later).

---

[72] Australian Government Department of the Prime Minister and Cabinet 2016, *Guidance Note – Cost–benefit analysis*, Office of Best Practice Regulation, February, pp. 7-8. Note that the social discount rate chosen has a major effect on the scale of the results.

[73] Victorian Department of Treasury and Finance 2013, *Economic Evaluation for Business Cases – Technical guidelines*, August

[74] Reflecting both the conservatism of the New Zealand Treasury and the higher interest rate structure of New Zealand's capital market – possibly reflecting lower rates of New Zealand saving – the discount rate recommend by New Zealand Treasury was at a very high 10 percent, then was lowered to eight percent, and is currently six per cent.

The base year for present value is 2016-17, the year in which the assessment is conducted.[75]

Figure 15 shows the effect of this, depicting the gross benefit of PHRN-related data linkage, in present value over the period of analysis. Table 16 summarises the cumulative results for each scenario. For example, Scenario A for PHRN produces a stream of gross benefits to 2040-41 having a present value to Australia of nearly $7.6 billion. Put another way, collectively Australians should be willing to pay nearly $7.6 billion for the improvement in health outcomes that are likely to be influenced by data-linkage based research projects.

Figure 15 – Gross benefit of PHRN-related data linkage, Scenario A, present value $2017

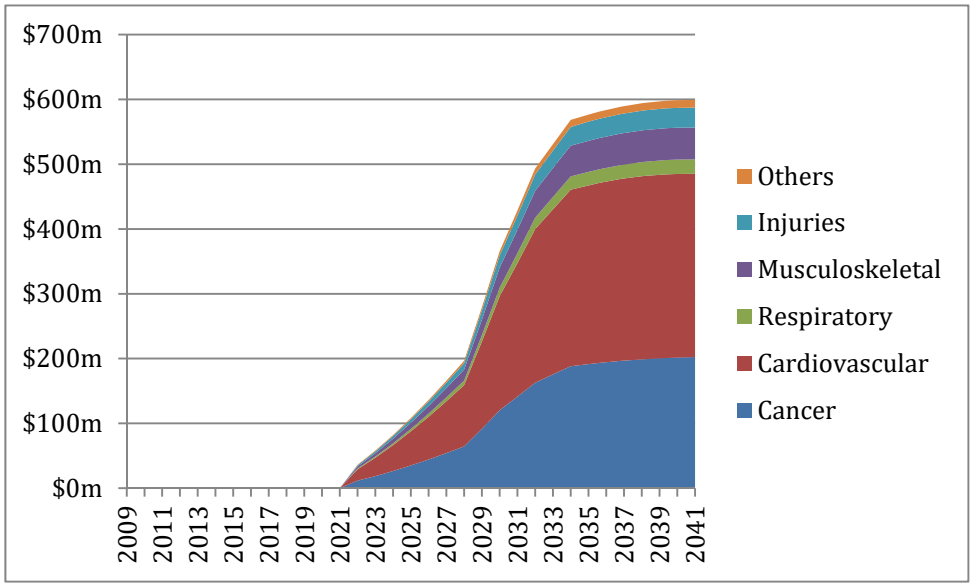


Table 16 – Summary of gross benefit of PHRN-related data linkage, various scenarios

Scenario	Gross benefit (present value, $2017)
Scenario A	$7,588 million
Scenario B	$6,718 million
Scenario C	$5,403 million
Scenario D	$2,637 million

[75] See Australian Government Department of Finance and Administration 2006, *Handbook of Cost-Benefit Analysis*, January, p.52

Costs

Financial costs

The costs in scope are those financed by NCRIS funding as well as cash and in-kind contributions by PHRN participants. In effect, for the purposes of this analysis, we treat the different sources of funding as a pooled fund. This matches the way that benefits have been approached, as there is no reasonable way to isolate different benefits to different parties when the funds are effectively pooled and put towards a single program of activity.

We assume that financial costs of PHRN from 2008-09 to 2015-16 are a reasonable proxy for the economic costs incurred by society (i.e. the opportunity cost of resources utilised for PHRN), given competitively determined market prices for labour and capital equipment.

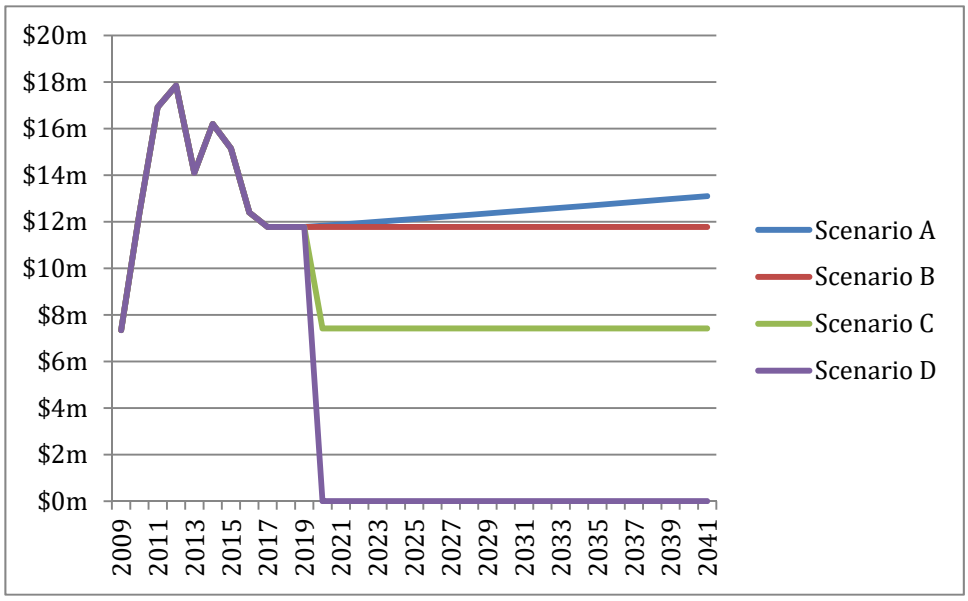
In-kind contributions are mostly personnel. It could be argued that personnel contributions should not be included if the participants would still incur these costs (i.e. employ the personnel) in the absence of PHRN. The opposite could also be argued – that participants (particularly those establishing a data linkage capacity for the first time) would not employ the labour time of such personnel in aggregate if it were not for PHRN, and so the opportunity cost of their wages and on-costs should be included. Taking a conservative approach, we include non-cash in-kind contributions as a cost of PHRN.

Full details of expenditure by year were not available. We assume total PHRN costs are phased across the years 2008-09 to 2015-16 proportionate to what is reported as expenditure in historical PHRN publications. For the purposes of this analysis, we assume total costs between 2016-17 and 2018-19 continue with an annual nominal increase of 2.5% (i.e. effectively stable in real terms).

The assumed (real) financial costs beyond 2019-20 vary by scenario, repeated below, and are depicted in Figure 16. This figure does not show the figures as present values, in order to more clearly depict the trend in real costs.

Scenario	Description
Scenario A	All annual resources slightly increase by 2.5% (in real terms) after 2019
Scenario B	All annual resources are maintained (in real terms) after 2019
Scenario C	Participant annual resources are maintained (in real terms) after 2019 but no further resources from the Australian Government through NCRIS or a similar initiative
Scenario D	No further resources for data linkage development after 2019.

Figure 16 – Financial cost profiles of different scenarios (constant $2017)



Marginal excess tax burden

A further economic cost included is the marginal excess burden of taxation. Sources of funding for PHRN are almost entirely government (whether through Australian Government initiatives or generally through government department budgets). Raising government revenue to invest in science is not costless to society. There is a cost to society from raising revenue through taxation.

Consistent with recent Productivity Commission analysis derived from Treasury analysis, we apply a rate of 24% on the financial costs to account for such 'deadweight loss'.[76]

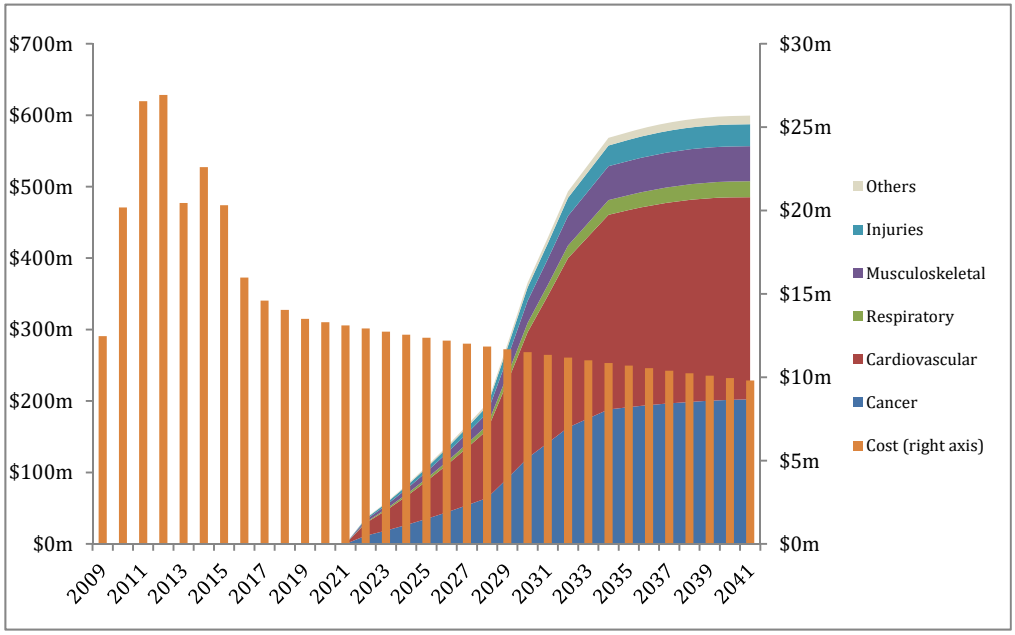
Net benefit results

The overall results for the indicative economic analysis show that like all infrastructure, the data infrastructure of PHRN involves incurring front-ended costs for an ongoing stream of benefits into the future.

Figure 17 for scenario A demonstrates this pattern, with the columns showing costs (left axis) and the filled area (right axis) showing benefit, as present values.

[76] Productivity Commission 2011, *Disability Care and Support*, inquiry report no.54, volume 2, 31 July, pp.955. This discusses studies undertaken for the Henry Tax review suggesting that the MEB of income tax is around 24 per cent. Using a MEB for income tax is reasonable given PHRN is principally funded from federal tax revenues.

Figure 17 – Summary of benefits (left axis) and costs (right axis) of PHRN over time, scenario A (present value $2017)



In summary, Scenario A indicatively shows a net economic benefit from PHRN-related data linkage to Australian society of $7.1 billion, expressed in 2016-17 dollars as a present value for the period to 2040-41. These benefits reflect PHRN-related research contributing to policy, practice and technology improvements that make Australians healthier, longer-living and safer. There is a benefit to cost ratio of 16.5, or over $16 in value for Australia for every $1 in cost. This is a substantial return on investment.

Other scenarios also show strong net benefits (see Table 17), with the extent of net benefit decreasing through scenarios B, C and D. In short, the fewer resources utilised for PHRN under the scenarios, the less benefit is generated. Scenario A allows Australians to be nearly $4.7 billion better off than Scenario D, in present value terms.

Table 17 – Summary of net benefit of PHRN by scenario (present value $2017)

Scenario	Benefits	Costs	Net benefit	Benefit to cost ratio
Scenario A	$7,588 m	$460 m	$7,128 m	16.5 to 1
Scenario B	$6,718 m	$403 m	$6,316 m	16.7 to 1
Scenario C	$5,403 m	$330 m	$5,072 m	16.3 to 1
Scenario D	$2,637 m	$208 m	$2,429 m	12.7 to 1

Notes on interpretation of net benefits and scenarios

The scenarios offer similar benefit/cost ratios until 2019, then diverge based on the activities feasible under different levels of resources amongst the higher cost scenarios. This is somewhat counter-intuitive as generally one would expect the application of investment to data linkage to experience diminishing marginal returns as is the case with investment in most activities. However, we think there are plausible reasons for the modelling outcomes reported.

Firstly, institutional factors around sources of funding can affect allocative efficiency. The NCRIS-type funding in Scenarios A and B can be allocated to what PHRN consider to be the best uses across jurisdictions nationally. By contrast in Scenario C without NCRIS-style funding beyond 2019, a higher proportion of resources are allocated by individual jurisdictions to uses within their jurisdictions which, overall, may have less marginal benefit.

Secondly, there may be network effects as the national set of linked data is developed to maturity.

Much of the PHRN costs to date involve establishing underlying infrastructure (e.g. technology, systems, processes, capability) for data linkage, and integrating the data collections that have been easier to link. Through the late 2010s and into the 2020s, particularly for Scenarios A and B, more higher-value linked data is expected to become available for research and other uses, which is the driver of usage. Implicit in Scenario D is that much of this expansion will not occur.

Arguably, linked data becomes increasingly useful as more data collections are integrated (and as available data collections cover a longer time period), particularly high value data collections such as those of the Commonwealth. As this occurs, usage would be expected to grow at an accelerating rate. We might expect increasing returns to scale as high value data collections are integrated. However, beyond this period, when most high value data collections are integrated and the user community is mature, we would expect that the marginal benefit would decline with ongoing investment. For example, a given level of PHRN effort in 2025 may have a greater effect on outcomes than in 2035, all other things being equal. However, given the lag between research activity and real-world outcomes, the effects of this difference are mostly beyond the timeframe being modelled.

Notwithstanding the above, Lateral Economics' analysis (including the distinction between scenarios) is indicative and should not be taken to be too precise.

Sensitivity analysis

The indicative results above are based on best estimate assumptions, drawing from literature and stakeholder input. To test how sensitive these base case results are to the value of key inputs, we have developed two bundles of alternative assumptions as a 'pessimistic' case and 'optimistic' case, respectively.

The pessimistic case incorporates:

- a real social discount rate of 7% (base case 4%); and
- the proportion of health research impact in Australia being from Australian research at 3.8% of all research's impact on Australian health outcomes, which is the unadjusted Australian proportion of world health research by count (base case 10.9%).

The optimistic case incorporates:

- a real social discount rate of 3%;
- the proportion of health research impact in Australia being from Australian research at 18% of all research's impact on Australian health outcomes, which is the rate assumed for clinical practice; and
- a value of statistical life year 20% over the base case level given that the figure normally used seems a little low compared with other jurisdictions – such as the USA.

The net benefits for each of the four scenarios, under these alternative assumptions, are at Table 18. All figures are in present values for 2016-17.

Table 18 – Net benefits under pessimistic and optimistic cases

Scenario	Pessimistic	Base case	Optimistic
Scenario A	$1,209 m	$7,128 m	$17,357 m
Scenario B	$1,056 m	$6,316 m	$15,400 m
Scenario C	$850 m	$5,072 m	$12,310 m
Scenario D	$390 m	$2,429 m	$5,799 m

The key finding from this sensitivity analysis is that there is a large range of potential net benefits. For example, in Scenario A the net benefits could be as high as $17 billion or as low as $1.2 billion under alternative reasonable assumptions. They also suggest that PHRN will generate substantially more benefits than costs to Australia even under systematically pessimistic assumptions for Scenario D which involves the least development of data linkage.

Table 19, Table 20 and Table 21 provide further detail of gross benefits and costs under pessimistic and optimistic cases.[77]

Table 19 – Gross benefits under pessimistic and optimistic cases

Scenario	Pessimistic	Base case	Optimistic
Scenario A	$1,616 m	$7,588 m	$17,844 m
Scenario B	$1,427 m	$6,718 m	$15,820 m
Scenario C	$1,168 m	$5,403 m	$12,649 m
Scenario D	$619 m	$2,637 m	$6,000 m

Table 20 – Costs under pessimistic and optimistic cases

Scenario	Pessimistic	Base case	Optimistic
Scenario A	$407 m	$460 m	$487 m
Scenario B	$370 m	$403 m	$420 m
Scenario C	$318 m	$330 m	$339 m
Scenario D	$229 m	$208 m	$201 m

Table 21 – Benefit to cost ratios under pessimistic and optimistic cases

Scenario	Pessimistic	Base case	Optimistic
Scenario A	4.0 to 1	16.5 to 1	36.6 to 1
Scenario B	3.9 to 1	16.7 to 1	37.6 to 1
Scenario C	3.7 to 1	16.3 to 1	37.3 to 1
Scenario D	2.7 to 1	12.7 to 1	29.8 to 1

While we have attempted to be conservative in the base case, the indicative results are sensitive to the assumptions used and should be interpreted with care. Data linkage is an emerging area of public management and the realisation of much future benefit is highly dependent on how data custodians, user groups in government, academia and the professions, and end-users of research insights, learn and evolve their understanding and practices. This illustrates how important it is for the introduction of linked data to be broadly embraced by all those who can benefit from it in their work or in their lives.

[77] Note that costs increase in more optimistic cases, in this example, as the lower discount rate is the only sensitivity test assumption that affects costs. The lower discount rate increases the present value of future costs.