UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Australian Centre for
Health Engagement,
Evidence & Values

November 2019

Exploring the Public Interest in and Social Licence for the Use of Linked Administrative Data in Therapeutic Development through a Scoping Review, Survey of Community Attitudes, and Hypothetical Case Studies

Prof Annette Braunack-Mayer
University of Wollongong

Dr Jackie Street
University of Wollongong

Ms Belinda Fabrianesi
University of Wollongong

Prof Stacy Carter
University of Wollongong

Dr Pauline O'Shaughnessy
University of Wollongong

Dr Lina Engelen
University of Wollongong

Ms Lucy Carolan
University of Wollongong

Ms Rebecca Bosward
University of Wollongong

Prof David Roder
University of South Australia

Ms Kylie Sproston
Bellberry Pty Ltd.

UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Australian Centre for
Health Engagement,
Evidence & Values

The Australian Centre for Health Engagement, Evidence and Values (ACHEEV) is the leading centre for high-quality deliberative and values-based research in health and health policy in the Asia-Pacific region.

We make health-related decisions more inclusive and democratic by constructively engaging diverse stakeholders, and more justified and legitimate by explicitly attending to ethics and values. Our methods include community juries, world cafés, discrete choice experiments, Delphi processes, dialogue groups and social research methods.

ACHEEV brings content experts, ethicists, methodologists, policy actors, health professionals, consumers and citizens together around contested or controversial issues in public health and healthcare.

uow.edu.au/social-sciences/research/centres-groups/acheev/
acheev-admin@uow.edu.au

For information on this report please contact:

Dr Jackie Street
streetj@uow.edu.au
+61 242528269

Suggested citation:

Annette Braunack-Mayer, Jackie Street, Belinda Fabrianesi, Stacy Carter, Pauline O'Shaughnessy, Lina Engelen, Lucy Carolan, Rebecca Bosward, David Roder, Kylie Sproston. Exploring the Public Interest in and Social Licence for the Use of Linked Administrative Data in Therapeutic Development through a Scoping Review, Survey of Community Attitudes, and Hypothetical Case Studies. A report for the Population Health Research Network, Australia. November 2019.

We acknowledge the Aboriginal custodians of the lands where we work, the Dharawal and Gadigal peoples, and their elders, past, present and future.

Contents

Executive Summary 4
Introduction 6
Aim 7
Structure of the report 7
What is the public interest? 8
What are 'publics'? Citizens, consumers and advocates 8
Conceptualising 'the public interest' 9
The public interest inheres in maximising goods and minimising bads 10
The public interest in sharing data with the private sector 10
The public interest and community attitudes 11
Summary 11
Scoping Review 12
Method 12
Search Strategy 12
Inclusion/Exclusion Criteria 13
Data Extraction 13
Collating, Summarizing and Reporting the Results 13
Results 14
Public support for sharing publicly held health data 14
Public concerns about sharing data with private industry 16
Circumstances under which government health data may be shared 17
Constructing a social licence or social contract to share government health data 21
Summary 23
Community Attitudes Survey 24
Method 24
Data Analyses 25
Findings 26
Demographics 26
Top Line Findings 27
Sociodemographic Patterning of Responses 30
Impact of consent preferences on conditions and views of sharing government data with private companies 36
Summary 39
Hypothetical Case Studies 40
Method 41
Recruitment 41
Data collection 41
Analysis 41
Results 42
Hypothetical 1: Submission to support inclusion of a new drug on the Pharmaceutical Benefits Scheme using Cancer Registry data 42
Hypothetical 2: Monitoring safety of therapeutic devices 43
Hypothetical 3: Sharing aggregate data through public websites 43
Hypothetical 4: Release of 'de-identified' health data 43
Conclusion 44
References 46

Appendix

Appendix 1 50
Logic grid – research question one

Appendix 2 51
Logic grid – research question two

Appendix 3 52
Scoping review inclusion/exclusion criteria

Appendix 4 53
Scoping review process and findings flowchart

Appendix 5 54
Summary of the final scoping review publications

Appendix 6 59
Summary of concerns about data sharing with sample quotes

Appendix 7 62
Summary of conditions which a proportion of participants believe needs to be met before data should be shared with private companies with sample quotes

Appendix 8 67
Community attitudes survey

Appendix 9 76
Table of counts of the sample characteristics used for calculating population weights

Appendix 10 77
Summary of demographics of online survey participants

Appendix 11 79
PHRN online survey question 6 analysis

Concerns about sharing government health data with private companies 79

Conditions under which government health data should be shared 80

Appendix 12 82
Stakeholder recruitment email

Appendix 13 83
Stakeholder interview schedule

Appendix 14 84
Hypothetical case studies

Hypothetical 1: Submission to support inclusion of a new drug on the Pharmaceutical Benefits Scheme using Cancer Registry data 84

Hypothetical 2: Monitoring safety of therapeutic devices 84

Hypothetical 3: Sharing aggregate data through public websites 86

Hypothetical 4: Release of 'de-identified' health data 87

List of Tables

Table 1 15
Proportion of participants willing to share health data with specific groups

Table 2 19
Conceptualisation of public benefit in health through data linkage and use

Table 3 26
Participant population adjusted demographics (n=2,537)

Table 4 31
Adjusted percentages of willingness to share government health data with the private companies by socio demographic patterning (n=2,537)

Table 5 32
Adjusted percentages of conditions on sharing government health data with the private companies by socio demographic patterning (n=2,537)

Table 6 34
Adjusted percentages of views on sharing government health data with the private companies by socio demographic patterning (n=2,537)

Table 7 37
Adjusted percentages of question 4 stratified by question 2 ("opt in" and "opt out")

Table 8 38
Adjusted percentages of question 4 stratified by question 2 response 'opt in' and response 'opt out'

List of Figures

Figure 1 28
Support for sharing government health data with the private sector (n=2,537)

Figure 2 28
Adjusted percentages of consent preferences (n=2,573)

Figure 3 29
Adjusted percentages of conditions on sharing government health data with the private companies (n=2,537)

Figure 4 30
Adjusted percentages of views on sharing government health data with private companies (n=2,537)

Executive Summary

The Population Health Research Network (PHRN) commissioned the Australian Centre for Health Engagement, Evidence and Values (ACHEEV) to develop a clearer understanding of the public interest in and social licence for the use of linked administrative government data by private companies in Australia. We carried out a systematic scoping review, community survey, interviews with private sector stakeholders and theoretical analysis to address this topic.

Whether or not a specific instance of sharing data with the private sector is 'in the public interest' will always need to be evaluated on a case by case basis. Weighing research outputs and privacy protections - conceptualised as both aggregative and corporate goods - will be at the centre of such public interest judgments.

Our scoping review found no previous studies that examined Australian public views. The small number of international studies reviewed put support for data sharing with the private sector at between 16% and 65%. The studies reported a complex suite of concerns, particularly about security, misuse and the profit motive and lack of public accountability in the private sector. There was broad agreement across studies that government health data could only be shared with the private sector if the research was of public benefit and in the public interest, access was tightly controlled and the data were anonymised. There was also support for informed consent for data use, safeguards such as independent oversight and a strong program of public engagement.

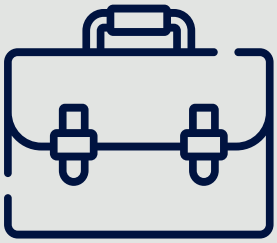
Our survey of 2,537 people recruited from across Australia found that between 52% and 58% of all respondents were willing to share their government health data with the private sector; lower proportions were in favour of sharing information to improve health services. A similar proportion of participants also wanted an opt-in method of consent. Overall, women, younger people, less well-educated people, people living in regional areas and, to some degree, people with poorer health status, were more concerned to impose conditions on release of health information. There was a very wide range of concerns about how private companies might use health information.

We need to take what publics say about sharing government health data seriously. The research outlined in this report, and recent studies, suggest that sharing government health data with private industry will require concerted and nuanced public engagement. Both government and the private sector will need to address the public's lack of understanding and lack of trust in the ways in which agencies collect, share, protect and use their personal data. We will need transparent, interactive and informed engagement that takes into account the capacity for and barriers to engagement.

Of the 2537 Australian people we surveyed...

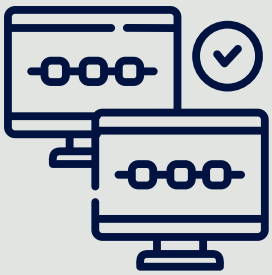
The gender split was roughly equal and most were in the 30-49 year age bracket, self-rating their health as good to excellent.

The majority lived in Metro areas such as Sydney, Melbourne and Brisbane.



Most had a University or Trade/ TAFE education, were full/part-time employed and did not work in the health industry.

Our survey found...



Between 52%-58% were willing to share their government health data with the private sector data.

Lower proportions were in favour of sharing information to improve health services.

A similar proportion of participants also wanted an opt-in method of consent.

Introduction

Every day, Australians produce large amounts of information about themselves through their interactions with government agencies. These large public sector datasets are a rich resource for service improvement, predictive analytics and new discoveries. Big data is already driving entire commercial market sectors; there is potential for similar benefits in using these big public sector datasets in similar ways.

Alongside the evident power of big data in the public sector, there is growing awareness of the real risks associated with its use. The potential risks of repurposing these data include:

- potential for privacy violations,
- loss of personal control over the ways and places in which personal data is presented,
- accuracy and misuse of data,
- systematic bias and the potential to reinforce existing prejudice in systems,
- harms to individuals, communities and agencies, including discrimination, reputational damage and embarrassment, and
- the potential to undermine public trust in service providers and systems.

These risks are especially acute in the public sector because administrative government data held is highly sensitive and of particular legal and ethical significance.

Effective and appropriate use of large public datasets to support therapeutic development is one area of focus for Australia's 2016 National Research Infrastructure Roadmap. This encompasses all stages of the development of new tests, devices and pharmaceuticals from discovery and proof of concept, pre-clinical and clinical phase, to registration and post-market phase.

The Population Health Research Network (PHRN) is currently conducting a project to better understand the needs of all stakeholders involved in therapeutic development (government, industry, researchers and the community) and to develop a strategy to successfully engage with this sector. The research described in this report has been conducted to develop a clearer understanding of the public interest in and social licence for the use of linked administrative government data by private companies.

Aim

The aim of this project was to examine community attitudes towards government sharing health data with private companies for research and development of treatments for disease and disability.

In consultation with the PHRN, we established the following objectives to address this aim:

1. Conduct a scoping review of relevant literature on public interest in, community attitudes towards and social licence for the use of linked government administrative data by private sector organisations for therapeutic development.
2. Explore current community attitudes to the use of linked administrative data by private sector organisations for therapeutic development in Australia.
3. Develop case studies (hypotheticals) of the kinds of therapeutic development research conducted by private sector organisations that would be in the public interest.
4. Provide a brief theoretical account of the meaning of 'public interest' as it applies in this context, linked to the hypothetical case studies developed for objective 3.

Structure of the report

The report begins with a brief account of how we might conceptualise the public interest in sharing government health data with the private sector. It then moves to two accounts of community attitudes toward sharing data reported through a scoping review of the international literature and a survey of the Australian population. The final section draws on the views of private sector stakeholders to craft four hypothetical examples of sharing government health data with the private sector.

What is the public interest?

What does it mean to say that something is in the public interest? 'The public interest' is called upon rhetorically in many contexts, including in political debate and in evaluating research and research ethics, and as a trump. What constitutes the public interest is taken as self-evident in many of these debates, but it can also be an empty signifier: a marker that does moral and political work despite lack of a deep, stable or even accessible meaning. So, while we attempt a conceptualisation and operationalisation of the public interest here, we note that its everyday meaning will remain fluid, and this instability will continue to do political and moral work.

We begin by considering publics, and then consider what it means for such publics to have interests and the nature of those interests. We then move to a discussion of how we might conceptualise the public interest in sharing data with the private sector in an Australian context.

What are 'publics'? Citizens, consumers and advocates

In the context of democratic engagement and deliberation, publics can be typologised into three groups: citizens, consumers and advocates. (1) To illustrate the difference, consider the following example: a proposal to share data from the public health system with a private pharmaceutical company.

An international pharmaceutical company wishes to bring its recently developed oncology medicine to the Australian market. The medicine extends life and has fewer side effects than existing medicines currently provided through the government-subsidised Pharmaceutical Benefits Scheme (PBS) for the same condition. The company approaches a Cancer Registry based in a State Department of Health with a request for aggregated, and therefore de-identified, data. Requests to the Cancer Registry for data are possible by application, and under a user-pay model. They would like to understand current treatment patterns in Australia, including age of onset of conditions, additional treatments, and information on variations in care provision across Australia.

Patient data held in the Cancer Registry are collected, without patient consent, from pathology laboratories, hospitals, radiotherapy and medical oncology departments, aged care facilities and the Registry of Births, Deaths and Marriages. The data collection is authorised under an Act of Parliament.

The Cancer Registry provides the company with aggregate data affording additional evidence to support a submission to the Pharmaceutical Benefits Advisory Committee. The submission is ultimately successful and the medicine is subsequently funded through Australia's Pharmaceutical Benefits Scheme.

Citizens in this example are the relatively disinterested general public, who may know very little about the issue (about the health system, cancer registries, data protection, medicines, or pharmaceutical companies). However, citizens have a stake in whether data is shared because they are members of the society in which sharing would take place. They, or people they know, may or may not benefit if the drug reaches the Australian market. They may or may not have contributed to the tax base that funds the Cancer Registry or the Pharmaceutical Benefits Scheme (PBS). Regardless, if given opportunity to learn about the proposal, their acceptance (or refusal) would be a basis for democratic accountability and legitimacy with respect to the proposal: they are the public at the heart of deliberative democratic engagement.

However, citizens are not the only public. The consumer public has more direct experience of the matter at hand: in this case, they might include, for example, people living with the cancer that could be treated with this drug. For most of this group, listing on the PBS will provide access to the drug that they would not otherwise have. Publics such as these can provide experiential knowledge to decision-making processes; it is, however, more difficult for them to be disinterested. They, understandably, will be seeking outcomes that might improve their personal experience of illness or the health system.

The final public is a public of advocates: direct stakeholders, who in this case might include pharmaceutical industry employees or committed data privacy advocates. (2) They have what approaches a traditional conflict of interest, where any primary interest they have in data sharing (such as providing benefit to patients through development of medicines) will be, or at least be perceived to be, compromised by a secondary interest (in, for example, maximising industry profit, or maintaining absolute data privacy at any cost, respectively).

These three publics are not independent. In respect of any data sharing decision, any particular person will be primarily a citizen, consumer or advocate. But they will often belong to two or even three of these publics. Most citizens are sometimes patients; some citizens have employment or community representation roles that locate them as advocates. Despite this lack of bright lines, it remains useful to ask 'which public?' when asking about 'the public interest'. A claim might be made, for example, that there is a 'public interest' in reimbursement of a high-cost drug which provides marginal survival benefit in metastatic cancer. However, on close examination, the interest may be primarily that of a consumer public rather than a citizen public, a distinction which arguably should have implications for reimbursement decisions.

Conceptualising 'the public interest'

There is general agreement in the literature that 'the public interest' cannot be easily or generally defined. The Australian Law Reform Commission argued in 2014 that there should not be a definition of 'public interest', citing the UK Joint Committee on Privacy and Injunctions. (3) The NSW Deputy Ombudsman has noted that, while the term is a central concept for democratic societies, it has never been "definitely defined" in either the courts or legislation. He commented that it is easier to be clear about what is not in the public interest: private interests, personal interests (of a decision-maker) and parochial interests. (4)

Laurie and Stevens, in their 2016 analysis of the legal and ethical implications of administrative data sharing in the UK, proposed that public administrative data holders should consider themselves bound by a public interest mandate. They observed that non-sharing is often the default, and that the public interest mandate proscribes such a default. However, they also argued that the public interest needs to be determined case by case in context, as it will differ between contexts. Public interest, they argue, "cannot be manufactured; it must be earned. This crucially involves the need to substantiate the public interests served by all uses of data" (5) but also entails a degree of uncertainty which also requires direct attention.

Laurie and Stevens' argument suggests that, while a broad commitment to the public interest might be something all public administrative data-holders should be expected to have and demonstrate, an interest in sharing (or not sharing) data will need to be examined and evidenced in each case. The public interest is not a single concept: there are multiple interests – for example, in privacy, health and wellbeing, and stable employment - and public interest judgments need to balance these on a case by case basis. We propose that, in most cases, there will be procedural goods that will serve the public interest, which might be very similar in all cases. These might include transparent communication, inclusive decision making reflecting stakeholders' perspectives, and contestability. (6) In contrast, data sharing proposals will also be substantively

different. Judging the public interest in a particular case of sharing will require consideration, for example, of what benefits or harms might result, the distribution of outcomes, and the possible effect on social bonds. It is these substantive dimensions that seem most likely to alter public interest judgements from case to case.

In Australia, the Commonwealth Privacy Act 1988 provides some guidance on how the public interest might be assessed in specific instances. Although the Act has no explicit definition of public interest, Sections 95 and 95A require weighing the public interest in privacy against the public interest in the conduct of research. Subsequent guidelines on Sections 95 and 95A have set out matters that Human Research Ethics Committees are to consider when doing this weighing, giving at least some substance to the meaning of the public interest in the conduct of research. These guidelines say that the public interest in research includes at least the following:

- scientific understanding and/or outcomes relating to public health or public safety (for Section 95A)
- identification, prevention or treatment of illness or disease
- improved delivery of health services
- enhanced scientific understanding or knowledge
- benefits to individuals, identifiable groups and the wider community
- the financial costs of not undertaking the research (to government, the public, the health care system etc.)
- the public importance of research (7, 8)

The public interest inheres in maximising goods and minimising bads

Claims that a case of data sharing is 'in the public interest' should be substantiated by some convincing evidence or argument regarding the balance of 'goods' and 'bads' that would result. We rely here on Widdows and Cordell's distinction between corporate and aggregative goods, which we have elsewhere argued extends also to bads. (9, 10) In their discussion of biobanks, Widdows and Cordell argue that communities' goods matter, and are of two kinds: aggregative (which result from adding up all of the goods experienced by all of the individuals in that community) and corporate. Corporate goods are

> *…social goods that attach to the community as a social whole such as justice, mutual advantage or law… [and]… require a community in order to be realized (such that it provides social or cultural institutions or values)… Importantly they also only exist as community and social goods. These goods are best understood as 'emergent social properties', in that they come about from the association and relations of individuals, yet are distinct from those of individuals (10)*

Data sharing, we would argue, potentially generates both aggregative and corporate goods and bads. Making a judgement regarding whether sharing data is in the public interest requires both adding up the effects for individuals (such as receiving new treatments, or having one's data hacked) and considering the collective and emergent corporate goods and bads (such as the advancement of knowledge, or making the health system untrustworthy). All of these goods and bads need to be weighed in making a final decision and on a case by case basis.

The public interest in sharing data with the private sector

Given the requirement in Australia that sharing data must comply with the Privacy Act, public interest arguments, at a minimum, must take into account the public interest in the research and the public interest in privacy. If we combine these two goods – research and privacy – with Widdows and Cordell's distinction between corporate and aggregative goods, we have at least a starting point for deciding whether a specific instance of data sharing might be in the public interest.

The public interest in the conduct of research will be met through generating goods such as better prevention or treatment of illness or disease and better health services. These are aggregative goods in the sense that individuals stand to benefit from better treatments and services; the good is a sum of the goods those individuals gain from longer or better life or employment. And, they are also corporate goods, because the benefits of better health services and treatments extend to future individuals, and contribute to a secure and flourishing society.

The second component relates to the public interest in privacy. The goods that a public interest in privacy serves are also both aggregative and corporate. As an aggregative good, the public interest in privacy is simply the accumulated privacy interests of all individuals who wish to control access by others to information about them. The corporate goods served by the public interest in privacy will include goods such as the maintenance of public confidence in hospitals or having a culture that supports open disclosure to clinicians of personal information. These goods, while certainly relevant for those individuals whose information is already held in public datasets, also offer corporate goods inasmuch as they have emergent properties that could not arise without collective efforts to maintain privacy as a social value, such as increased trust in public systems and a greater availability of useful information.

In the case study above, the cancer registry is a vehicle for the delivery of aggregate goods, as analysis of the data held in the registry may lead to access to a new drug for some patients. More generally, a registry is also a corporate good, as it contributes to new knowledge and, at least, has the potential to benefit future patients in general. There is also a clear public interest in the protection of privacy and, in the case of this registry, the Parliament has 'done the work' of balancing that public interest against other public interests to allow release of personal health information under certain strictly controlled circumstances.

Balancing the public interest in the conduct of research with the public interest in the protection of privacy as we suggest above is essentially the same task, regardless of whether data is shared with a public or private sector organisation. However, the capacity of private sector organisations to secure these goods may be constrained by factors that are distinctively commercial - for example, the need to limit access to the company's research outputs, protect intellectual property, or maximise market share. None of these factors are necessarily 'bads' in the sense we have identified above; at the very least, though, they are constraints on acting in the public interest.

In addition to the public interest in research and in privacy, there are other goods arising from sharing data with the private sector that may also serve the public interest. However, there is scope for advocates to argue almost any good into a calculation of whether data sharing is in the public interest. For example, a flourishing biomedical industry provides opportunities for satisfying and secure employment, something that is again both an aggregative good (benefits to employed individuals) and a corporate good (because there are emergent benefits for a society that is organised to ensure higher employment). As with all other claims that something is 'in the public interest', this claim will require evidence or argument to be substantiated. And, the public interest in satisfying employment (or, indeed, in any other good) should not be allowed to overwhelm or replace the more central goods of research outputs and protection of privacy.

Finally, private sector companies have interests that should not be considered when making a public interest judgment. In the case study above, the pharmaceutical company will have an interest in returning dividends to its shareholders; the parochial interests of a smallish group of individuals cannot count as being in the public interest. Neither can the strong reputation of the company as an industry leader in the development of new products count as being in the public interest, even if it benefits a large workforce.

The public interest and community attitudes

The public interest as we have conceptualised it above is not the same thing as 'what the public thinks'. Publics – citizens, consumers and advocates – can be mistaken, confused or misled about what is in their best interests. However, they can also be right and, even when they are wrong, what publics think is central to how we decide what is in the public interest in a given instance: public sentiment both reflects and shapes specific judgments about 'the public interest'. (11) We therefore need to take what publics say about sharing government health data seriously. The rest of this report does this by reporting community views collected through a scoping review, survey and interviews with stakeholders in the private sector.

Summary

Whether or not a specific instance of sharing data with the private sector is 'in the public interest' will always need to be evaluated on a case by case basis. We have suggested that weighing research outputs and privacy protections - conceptualised as both aggregative and corporate goods - will be at the centre of such public interest judgments.

Scoping review

To examine the overarching research question, 'What is the public interest in, community attitudes towards, and social licence for the use of government health data by private sector organisations for therapeutic development?' we conducted two separate search strategies to address the following sub-questions:

1. What are community attitudes towards the use of government health data by private sector organisations for therapeutic development?
2. What is the public interest and social licence for the use of government health data by private sector organisations for therapeutic development?

Papers describing models (e.g. five safes model), which may relate to the public interest, were extracted from the search results but we did not explicitly search for models. (12)

Method

Search strategy

In designing and conducting the scoping review, we drew on the work of Arksey and O'Malley (13) and Peters et al. (14) Since we were primarily concerned with the breadth of existing literature in the area, we did not assess or exclude papers based on quality but did note specific limitations of the studies.

Two logic grids (population, concept, context, outcomes) were developed for the study (see Appendix 1 & Appendix 2). The first search used terms describing citizens or patients, Big Data, the private health sector and views or perspectives, with these terms and relevant synonyms included in the searches. The second search used terms describing Big Data, social licence and public interest, with these terms and relevant synonyms included in the searches. Depending on the database, documents were sourced within the time period "last five years" or January 1st 2014 to April 1st 2019. We excluded papers published prior to 2014 because the field has developed rapidly in the last five years with significant increases in sharing and linking of data sets between government and private sectors (although not necessarily in Australia). (15) Earlier studies would not necessarily reflect current community views and judgements about the public interest. Studies published within the time period, but which reported data prior to 2014, were included.

We conducted a systematic literature search using four electronic databases PubMed, Scopus, Cinahl, and Web of Science. In addition, searches were conducted using Google Advanced and Google Scholar. These databases were selected for their coverage of quantitative and qualitative research with respect to the use of data analytics in health and, in particular, quantitative and qualitative research on community attitudes to data sharing in health. In Google Advanced and Google Scholar searches, where necessary, were restricted to the first 1000 hits. Additional 'pearled' relevant articles were extracted from the reference lists of included papers. Peer-reviewed and unpublished articles, reports, books and book chapters were included. Editorials or opinion pieces were excluded. There were no limitations on geographical location but only English language articles were included.

For the database searches we iteratively developed a search strategy based on the logic grid. Our final search strategy is shown in Appendix 3.

Inclusion/exclusion criteria

We screened title and abstract and, in the case of reports, the contents page using the following inclusion criteria:

1. Empirical studies using any methods
2. Discusses the sharing and linkage of data in research for therapeutic development (pharmaceuticals or medical devices)
3. Participant groups drawn from healthcare users, patients or the wider public
4. Examines the views, attitudes, opinions, perspectives, thoughts, awareness or acceptance with respect to the sharing of government health data with private industry for research and development. Government health data defined as information collected and held in the public sector including, but not restricted to, administrative data, electronic primary health care records, electronic hospital records, registries and national disease databases. Studies which were unclear about who was sharing the data (government or private industry), such as in the case of a cancer registry, were included. Similarly, studies which lacked clarity about the data recipient (government or private industry) were also included.
5. Described patient and public attitudes to sharing of health data for research in comparable areas such as biorepositories, genetic testing and genomic research but only where the research involved sharing data with private industry.

Studies were excluded if:

1. Duplicates, non-English articles and articles published prior to 2014
2. Exclusively described the use of digital methods, technologies or records in health care rather than public attitudes
3. Described the views of researchers, health professionals, industry experts, and government or key professional stakeholders rather than public or patient perspectives. Studies involving expert or stakeholder opinions were included if public and expert/stakeholder responses were reported separately.

Articles were screened based on title and abstract (by author JS, BF & RB). Full text screening was conducted by two authors (JS & BF). Where there was disagreement between the reviewers, the final decision for inclusion was made by the research team. Reference lists of included papers were reviewed and further articles identified. The flow chart in Appendix 4 summarises the review selection process and findings.

Data extraction

Two reviewers (BF & JS) extracted: title, author name, year of publication, location(s), aim(s), focus, public engaged, specific patient group, sample size, health technology, methodology, models (consent, data linkage, public interest), case studies, overarching results – access of private companies to public data, under what circumstances can public data be shared with private companies, consent, storage, definition of social /social contract, definition of public interest/public benefits, and bias/limitations which related to the research questions. One reviewer, JS, inductively coded, without a priori codes, the included articles using N-Vivo (www.qsrinternational.com/nvivo/home) to extract descriptive themes and develop analytical themes. A second reviewer, BF, coded two papers and used the extracted data and the research question to cross-check the coding framework. Differences were discussed and resolved.

Collating, summarizing and reporting the results

This scoping study sought to present an overview of all relevant material rather than synthesise evidence or to aggregate findings from different studies. We did not assess quality of evidence and consequently do not describe whether particular studies provide robust or generalizable finding, although limitations or potential sources of bias have been identified and reported as appropriate. A template for data extraction (described above) was used to provide a consistent approach to extraction and reporting of the findings.

Results

A systematic literature search of the four electronic databases, Google Scholar and Google Advanced generated the following number of articles: PubMed (797), Scopus (1,768), Cinahl (389), and Web of Science (1,844), Google Scholar (1,990) and Google Advanced (293).

A total of 6,788 articles were screened based on title and abstract (by author JS, BF & RB). Full text screening was conducted by two authors (JS & BF) on 221 articles (which included an additional four pearled articles). From there a total 33 publications were included. Appendix 4 reflects the review selection process and findings.

Of the 33 publications included in the review, 23 were peer-reviewed papers, seven reports, two conference proceedings and one a conference paper. Most papers reported on research conducted in the United Kingdom (n=17) and United States (n=7), with two studies set in Canada, two international and one each in Europe, New Zealand, South Korea, Switzerland and Thailand. Data collection within the included studies occurred during the period 2007-2018. A small number of studies did not report their data collection period.

Participants included broader public, affected patient groups, clinical stakeholders and private sector agencies, with ages ranging from 18 years to over 75 years. In total there were 25 papers focusing on the views of patients/ members of the public only. Seven studies reported expert opinions from clinical stakeholder or private sector agencies and the views of members of the wider public or patient groups.

Thirteen of the studies were qualitative, using focus groups, citizen juries, workshops, social assembly and one to one or group interviews. Sixteen of the studies were quantitative using online and in person surveys, and there were four mixed method studies. A summary of the final publications included can be found in Appendix 5.

Public support for sharing publicly held health data

We found no previous studies which examined Australian public views on the sharing of publicly-held health data with private industry for the purposes of therapeutic development.

Eight quantitative studies specifically asked participants if they would be willing to share their health data with commercial organisations (see Table 1 below). In general, willingness to share non-identified data was high with participants' own health provider and with academic researchers but fell if the data was to be shared with private companies. In addition, a large on-line discrete choice experiment in Scotland (16) excluded 461 of 1,465 respondents who began the questionnaire because they stated data linkage was "unacceptable under any circumstances" (ultimately 1,004 completed the survey).

In qualitative studies people expressed less willingness to share their government health data with for-profit private organisations than with other groups involved in health research such as health care professionals, university researchers and non-profit organisations. (17-22)

Willingness to share, across all studies, was, in part, dependent on the purpose of the sharing. A 2017 online survey of British adults (23) showed that 22% would be willing to share their medical records with an organisation that they knew: this fell to 4% if it involved an organisation that they did not know, yet, in the same survey, 47% were willing to share their medical data "if it helped develop new medicines or treatments". In a 2015 Northern Ireland survey, (24) participants (n=1202) were asked to consider whether a drug company should be given access to de-identified health data if they were working on a drug that might cure Alzheimer's disease. In this scenario, where there was a strong case for public benefit, three quarters of respondents agreed that the company should be given access.

Table 1: Proportion of participants willing to share health data with specific groups

	Family	Own health provider	Health services	Government departments	Academic researchers/ universities	Government researchers	For profit research companies	Commercial organisations like e.g. insurance	Charities/ non-profit organisations
Northern Ireland adults n=1202 (24)		91%	86%	73%	72%	-	-	41%	51%
Ipsos MORI for Wellcome Trust n= 2017 (25)						54%	53%	26% (insurance) 37% (marketing)	
USA Crowd sourced n=128 (26)		84%	-	28%	79%		24%	14%	-
USA genetics registry participants n=450 (27)	-	-	-	-	90-93% (same university) 76-87% (other universities)	64-76%	16-24%	-	84-91%
European leukodystrophy patients/carers n=195 (12)	37%	91-96% (specialist) 72-76% (family doctor)	-	-	64-90%	-	61-65%	-	-
Maori, New Zealand n=533 (19)	82%	-	-	77%	-	-	46%	-	-
South Korea, older adults n=170 (28)	80%*	-	66%*	-	39%*	24%*	25%*	19%*	-
London patients with diabetes n=404 (29)			51% 29%*					45% 15%*	49% 17%

*Identifiable data

The equivocal nature of attitudes towards data sharing was reflected in the qualitative studies. Across all the studies, sharing health data with the private sector was complicated by a complex interconnected network of conditions which participants placed on sharing, including:

- the purpose of the research,
- how the data will be shared,
- who will have access,
- the nature and security of the data,
- the potential for individual and societal harm,
- the nature of consent, and
- the safeguards in place to prevent misuse.

We describe these conditions in more detail below under the headings of concerns about data sharing and circumstances under which sharing would be acceptable.

Public scepticism about the acceptability of data sharing for health research is almost certainly exacerbated by a widespread lack of public understanding of data uses in the health sector and data research in general. Across several of the included studies it was clear that many people do not understand how the public and private health sectors work, the nature of data research, the extent of data collection, how data is owned and shared, the roles of different organisations and individuals within health research and the existing safeguards in place to regulate and control the flow of data. (17, 18, 25, 30-36) For example, the roles of academic institutions, non-government organisations and private companies in the development of new pharmaceuticals and devices were poorly understood amongst a sample from the UK public. (25) Similarly, a focus group study in Ontario, Canada indicated very low knowledge of research based on linked administrative health data. (18)

> *Is this actually happening today, where they're collecting a lot of data? General Public, Focus Group 2, Toronto 2017. (18)*

Many people also have little understanding of data analytics and statistics

> *It says so they can predict what will make you ill or better. How? Are they god? How can they work all that out? General public, Glasgow, 2016. (25)*

Public concerns about sharing data with private industry

Participants across the studies expressed numerous concerns about sharing their government health data generally and specific concerns about sharing their health data with private companies. Here we discuss the three most common concerns: security, misuse and profit making. A full list of the concerns is shown in see Appendix 6

Data security, data leaks and hacking

Participants across more than half the studies were concerned about confidentiality and data security. A large survey in West London indicated that 79% of the participants would worry about the security of their data if it were included in a large national system. The Wellcome Trust report on public attitudes to commercial access indicated that there was a belief amongst participants "that no amount of security could ever totally remove the risks involved in sharing data". (25, p.11) This included data leaks and hacking. Participants' concerns were related to the security of electronically stored data generally (17, 19, 25), the general disorganisation of hospital record keeping (20), concerns about 'selling on' data to other companies (25, 37) and their own general knowledge or prior experience through media reporting or individual targeting for marketing. (18, 25, 31, 38) Concerns about the possibility of third party access to data (21) and the "risk of unauthorised use or disclosure" (33) were also related to fear of surveillance, "being monitored and controlled" (21) and the potential for personal harm as discussed below. (25, 34)

Misuse of government health data

Discussion about the misuse of government health data were expressed both in terms of a general unease about the use of participants' data for purposes of which they were unaware or might oppose and concerns about specific harms that might result through misuse. Some participants saw sharing data as one more sign that "we are heading for a dystopian, surveillance-based society". (25) In one study involving older Swiss adults, participants expressed concern about the potential use of data in eugenics:

> *So, my only concern is, it has once been talked about, that it could be used to create the perfect human... or... that everyone would have blue eyes or a standard type or for military purposes. Of course, that is a big topic. I would be absolutely against that. No. 11, female, age 69, Switzerland. (21, p.8)*

Participants were concerned about becoming a "transparent citizen" with increased risk of selling or releasing government health data to "people or institutions that might gain unpredictable powers by it". (21, p.7) Across studies, these 'powers' related to the capacity to influence employment, provision of insurance cover, provision of financial services and health care provision. (17, 18, 20, 21, 25, 28, 31-34, 38) Participants were also concerned that even aggregate data might be used to stigmatise individuals based on their ethnicity or to segment, exploit or disadvantage vulnerable groups. (25, p.60) The impact of data depends on the context: for example, in a study conducted in Thailand participants explained that "migrant workers on the Thai-Myanmar border, may face increased stigmatization if they are identified as being a source of infectious diseases like malaria". (34, p.5)

Using government health data to generate profit

Although many participants acknowledged that there could be a role for commercial organisations in therapeutic development, they were concerned about the interests held by private industries, their lack of accountability and their need to make profits.

> *I'm fine with all of these organisations except businesses. Government usage is safer because there is responsible governance, but there is no corresponding obligation for private businesses who want to make a profit. General public, New Zealand, including First Nation peoples, 2017. (19, p.13)*

In several UK studies participants expressed concern that pharmaceutical companies would have access to publicly held health data to develop new drugs which they could then sell back to the National Health Service at considerable profit. (17, 30, 31)

> *Unfortunately, my belief is that when people start making a profit out of it that's when the ethics start getting a little bit less and a little bit less as the profit margin goes up the less ethical you are the more money you earn. Participant 1, person with diabetes, London, UK 2016. (29)*

Participants differentiated between private companies using government health data under regulated conditions for public benefit and unfettered access for generation of profit. One parent of a child with a rare disease commented:

> *Big pharma…Are they doing it with my consent, looking at a group to identify, make progress, come up with treatments, understand conditions more – I'd be comfortable with that. Or are they just given free rein on my daughter's medical records so they can stabilize business, play entrepreneurs, gamble on it – no that not OK. Parent of patient, Sheffield, UK, 2016. (25, p.57)*

Circumstances under which government health data may be shared

There was substantial agreement across studies about the circumstance under which government health data could be shared. The primary requirements were that:

1. the research should be of public benefit and in the public interest;
2. data should be securely stored;
3. access should be tightly controlled; and
4. the data should be anonymised.

There was also support for informed consent for the data use and safeguards such as independent oversight. A requirement for opt-in consent appeared to become less important if participants in the study had been able to discuss the associated issues with experts and deliberate at length. Faced with research of benefit not occurring because of governance issues, a majority of people (61%) canvassed in a UK survey would support commercial access to government health data. A quarter still did not want the research to occur if it were necessary for commercial organisations to have access to the data. A full list of circumstances which participants believed were necessary before data should be shared can be found in Appendix 7.

Public benefit and public interest

In many studies participants indicated that the purpose of the data use – for public benefit or in the public interest - was one of the most important considerations in the acceptability of data sharing. (12, 16, 17, 19, 20, 22, 25, 30, 31, 34, 39) Although, to some degree, these terms were used interchangeably, participants mainly focused on public benefit rather than public interest. One exception was a 2016 Scottish study in which participants suggested that in "research that would operate in the public interest…data would be used for appropriate and necessary purposes, and that research would (at least probably) ultimately lead to benefits for healthcare". (30, p.716)

Public benefit was seen as a broad and encompassing concept (see Table 2). In particular, cancer, dementia and mental health and research that improved health and quality of life through preventive measures were often highlighted as areas for particular focus. (22) Participants who were patients or relatives of patients with rare diseases saw data sharing as essential to support development of new treatments (12):

> *Patients are key to advancing research by providing data to researchers —the more information collected, the more it will promote advancement of research —in a rare disease like this, maximum participation is required for effective research. Patients or relatives of patients with leukodystrophies. (12, p.7)*

Monitoring the safety of drugs was widely seen as valuable but in one study of deliberative workshops British participants asked why the NHS could not conduct this work itself. (25, p.51)

Public benefit was also conceptualised as improving services for vulnerable groups. For example, in deliberative workshops, Scottish participants acknowledged that there was merit in health research targeting vulnerable groups to produce "benefits to particular smaller groups within the public" particularly those "in greatest need". (22, p.7) Ultimately this was also seen to be in the public and personal interest:

> *Just because you're not associated with it at the time it doesn't mean it won't impact you later on in your life. Female 3, Focus Group 3 Perth, UK 2018. (22, p.7)*

The participants in these workshops (22) recognised that there was a wide range of possible public benefits. As with participants in other studies, the workshop participants described the public benefit as "finding cures for diseases and making new drugs available". (22, p.7)

Concept	References
Disease diagnoses, treatments and cures with particular emphasis on cancer, dementia, mental health and rare diseases	(12, 17, 18, 21-26, 29, 31, 32, 35, 37, 40-42)
Improved population health and wellbeing including through prevention	(17, 21, 22, 29, 31, 33, 37, 41)
Monitoring the long-term safety and efficacy of drugs and treatments	(18, 25, 31, 35, 37, 41, 43)
Improved health services particularly improved health services, health and quality of life for vulnerable groups	(19, 22, 25, 29, 33, 35)
Improving research which will have impact	(20, 22, 29, 33, 43)
Creation and dissemination of new knowledge	(22, 33, 37, 41)
Improved allocation of resources, Cost-effective care	(22, 31, 35, 37)
Empowerment of individuals and communities including the perceived value of altruistic contribution to society	(22, 31, 40)
Improved health policy	(19, 35)
Giving “children the best start in life”	(21, 22)
Improving the lives of older people	(22)
Improvements to paediatric care	(25)
Improving the natural environment	(22)
Support for non-human life	(22)
Access to a wider skill set if private industry is involved	(18)
Ability to detect rare health events	(35)
Benefit to individuals	(17, 19, 25, 26, 31, 32, 43)

Table 2: Conceptualisation of public benefit in health through data linkage and use

Controlled access to government health data

There was consistency across the studies about who should be able to receive personal health data (see Table 1). Willingness to share health data was highest if the recipients were individual (e.g. own GP or specialist doctor) or organisational health care providers (e.g. UK National Health Service or US hospitals). (12, 22, 24, 26, 28, 29, 40) Participants were least willing to share their government health data with commercial companies, particularly insurance companies. (12, 19, 24-29) Academic researchers and non-profit organisations fell between these two extremes. (12, 24, 26, 28, 29) University researchers were trusted more than researchers working in private companies because they were regarded as more altruistic and less motivated by profit:

> *...you put your belief in the system that universities are there to try to sort of safeguard that this will be used for the correct reason. Mental health support group, Female 3, UK. (30, p.718)*

Willingness to share with research groups appeared to be related to age with several studies showing greater acceptance of data sharing with researchers amongst older age groups. (16, 24, 25, 28) However, the relationship was not always linear (25) and other studies demonstrated mixed findings suggesting the reasons were multifactorial. (23, 24, 27, 42) Willingness to share across the different entities also related to particular threats or benefits from doing so. For example, Kim et al. (28) suggested that older South Korean women, who were half their study group, may be unwilling to share their health data with government agencies since this might affect their access to health insurance or public welfare.

Participants reported a range of reasons for distrust of private companies. Private companies were regarded as motivated by profit. Participants in Scottish deliberative workshops believed that private companies had suppressed past 'cancer cures' and would suppress results in future research to increase their profits. (22) Participants could not see or understand that there could be any role for private companies in drug or device development. (25) They were also worried that the data would be sold on to others, particularly insurance or marketing companies, or they simply disliked the idea that private companies were making profit from their data:

> *Business involved changes things a lot for me – I'm unhappy with businesses getting personal data as they profit but don't have to give anything back. (19, p.13)*

Even with privacy and security safeguards, participants in a Canadian study believed that, as the number of people and organizations who accessed the data increased, risks to data security also increased. (18) In two studies, participants indicated that data users/organisations needed to be explicitly vetted before they should be allowed access. (24, 37)

Anonymisation of data

Anonymisation of government health data and the associated strategies of de-identification and aggregation were either a pre-requisite for acceptance of data sharing (16, 20, 25, 26, 34, 38) or they greatly increased willingness to share. (18, 19, 21, 30, 31, 33, 35) For example, in a large UK survey (n=2017) anonymity was seen as the second most important condition for sharing health data with commercial organisations. (25) In a smaller study, approximately half the participants (patients with diabetes n=404) indicated they would be more willing to share their government health data with NHS researchers provided it was anonymised, and this fell to 28% if identifiers were retained. (29) In the case of identifiable data and sharing with pharmaceutical companies, only 15% were willing to share their health data. (29) In UK focus groups, (8) (n=50) anonymisation of data was seen as particularly important if data was to be shared with private companies, expressed through concern that identifiable information could be misused. (30) In several studies, participants were concerned that even with identifiers removed it could still be possible to re-identify the data using 'jigsaw identification'. (24, 25, 41)

Consent

Consent was an important consideration in many studies. However, the need for and type of consent was highly contentious (20, 25) with participants in some studies comfortable with government health data sharing with no consent through to those who wanted explicit consent on every occasion. (25) Participants in two deliberative studies started from a position of explicit informed consent but after receiving information they shifted their views as it became apparent that it may be impossible or the cost of obtaining consent may be prohibitive. (25, 37) They moved to supporting data sharing without consent in cases with high public benefit. However, particularly where government health data would be shared with commercial entities, some participants still wished to have the opportunity to consent or refuse data use even if it meant that the research did not go ahead. (25) Some participants

suggested that their ongoing reservations related to a lack of clarity about the personal implications of sharing their data. (20, 25)

Safeguards

Participants across studies called for a range of rigorous governance structures to monitor and regulate access to government health data. For example, participants in a UK study (25) called for strict rules prohibiting passing data to third parties, anonymisation for data sharing, sanctions for misuse of data, secure data storage and oversight by an ethics committee. Many participants in the study wanted multiple safeguards to be instituted. Monitoring of individual access to the data by logging contact episodes was also suggested, the rationale being that this would serve as a deterrent for malpractice.

> *That would make me feel a bit more comfortable because they would know, if for any reason the system had been abused, not that it would be but they would know…There'll be a shortlist of people who have accessed, it would be a deterrent of abuse. General Public, Belfast, UK. (25, p.63)*

Secure storage and independent oversight were widely recognised as essential conditions for data sharing. One participant in a European study likened the necessary controls to those found in the banking sector:

> *I'm just trying to say there is this framework, you know we say that there is a governance system in place which will protect the patient and we can look at them like we do the financial institutions and we're quite happy with how they exist, well they're quite well developed. There's a framework around this and we want some assurance. Patients/parents of patients with a rare disease, Europe. (38, p.1,405)*

One individual writing in response to attempts to share public administrative data sets in the UK said:

> *I want the data to be supervised by an independent forum of individuals whose remit is to follow strict published ethical guidelines relating to sharing, selling and profit making by the use of my data. UK, Comment posted to website Care.data. 22.01.14. (41, p.184)*

Clear explanations about how data would be recorded, anonymised and stored appeared to be helpful in building support for government health data sharing:

> *He explained to me that basically there's only one location where there's a cross-reference between the name of the participant and the identification process they're using on each individual patient's, or study participant's, file. So I don't have any issues with that. ClinSeq#120, NIH genomic research registry participant, USA. (20, p.967)*

Constructing a social licence or social contract to share government health data

Very few studies explicitly described the idea of a social licence or social contract to support data linkage and sharing, although trust was frequently described as foundational to public acceptance. However, many participants identified a number of ways in which trust in public and private organisations could be improved. These included reassurances to the public that data would be secure, better communication about the nature of public benefit from data sharing, meaningful ongoing public engagement and data sharing through trusted entities.

Reassurance that every effort is made to keep government health data safe

These measures have been outlined in the previous section ("Under what circumstances") and in Table 1 – summary of all conditions. In describing how trust in data sharing could be built, participants in the reviewed studies wanted effective governance structures in place to ensure data security and accountability. The key measures proposed were transparent data security, appropriate legislation to regulate data sharing, fines or penalties for individuals and companies who are negligent or misuse data and independent oversight.

These measures also needed to be well publicised and communicated to the public. (18, 23, 24, 37, 42, 44) For example, the writer in this letter to a UK newspaper reflected on the failure of government to persuade the public to share general practice data:

> *They blew it by being patronizing and disingenuous and by being unlucky enough to be preceded by Wikileaks exposures. They need to regain trust by apologizing for their previous abject failure and then by persuading*

> *us as individuals that a properly anonymised, secure version is safe and effective. Letter to The Guardian, 18.08.14. (41, p.183)*

Part of the challenge in releasing data is low public understanding of data use, data linkage and existing governance structures.(18, 25, 32, 33) For example, in a large UK study, many of the safeguards participants called for were already in place. (25) A US study with patients highlighted the need to inform and educate patients about current practices and protections. As one patient stated:

> *I think part of it comes down to, it's just patients getting enough education about the process, and the outcomes that we're looking for, to feel comfortable sharing that information. Patient from a patient advisory panel/ network, USA. (33, p.544)*

One report of multiple public engagement exercises in New Zealand suggested: "Organisations using data in this way, therefore, need to work harder to explain how data is being used, why it is needed, what the individual is gaining in exchange and what the business itself is gaining." (19, p.18) Another US study talked about providing "examples of trustworthiness" and "communicating details about research policies and procedures" through on-line platforms including social media as a way to facilitate data sharing. (40) Deliberative work with UK participants indicated that education on aggregation and anonymisation and "clear transparent online consent processes" without "confusing tick boxes or small print which is never read" would help to build public trust. (25, p.13) In particular, "participants felt that if they knew more about the processes and safeguards in place they might feel more empowered, and hence more open and trusting in the decision making process around data collection and sharing". (25, p.13) The NICE Citizens Council proposed that transparency could be ensured "through open days and information resources to explain what data is being used for, explaining precisely how it will be used and by giving reassurance that personal care data will not be passed on or sold to other organisations". (31, p.41)

Better communication with the public about public benefit

Beyond clarity about the mechanisms and safeguards for data sharing, participants in these studies also wanted clear information about the benefits which have been gained or which might accrue through sharing public data sets with private industry. Participants in several studies called for researchers and data custodians to actively publish and promote positive stories associated with use of data, to explain the reasons why it would benefit the public to use and share government health data, and to provide feedback on outcomes to participants included in research. (25, 30-32, 38)

> *It is important, I think the public should definitely be more informed and well informed and quite clearly explain to people why the data has been collected and what purpose and how it is used. I think they have a right to know. Black and Ethnic Minorities Group – Male 2, Scotland, Focus Groups. (30, p.719)*

Attention to building trust with the public through ongoing public engagement

Several studies, drawing on public, patient, carer and family views, called for increased public engagement. (12, 18, 19, 24, 30, 38) Aitken et al, 2016 (30) drew attention to the very different notions of public engagement expressed by stakeholders (researchers, social scientists, government analysts, data controllers and community representatives) compared with those emerging from focus groups with patients and the general public. The authors suggested that "much of the discussion at the stakeholder workshop could be viewed as exemplifying a deficit model of public engagement, whereby public trust can be 'improved' through the provision of appropriate (and selective) information" whereas "focus group participants indicated that they would appreciate a more open exchange of information and greater equity in the science–public relationship".(30, p.719) Stakeholders "discussed public engagement as a means of generating public trust in research/researchers whereas public participants saw it "as a potential indicator of the trustworthiness of the research and/or researchers".(30, p.719) Deliberative sessions conducted by NICE,(31) Ipsos Mori (25) and Tully et al. (37) suggested that, in response to information and discussion with others, the public shift in their views on governance for data linkage and sharing, albeit not always in the direction other stakeholders may want. Patients and carers from a rare disease group meeting saw patient input as important for "good governance" but also saw the need for capacity development for full patient involvement. (38)

Data sharing initiated and explained by individuals and organisations known and trusted by the public

Across studies it was clear that the public and patients were more likely to share personal data, including health data, with organisations or individuals they trusted. (18, 23, 30,

33, 41) The basis for the trust was not entirely clear but it appears to be related to: personal experience of organisations - for example, US participants in a small study trusted the academic organisation with whom they had already interacted but were less willing to their data with other academic groups (27) and 64% of UK survey respondents were willing to share some personal data with organisation they knew versus 36% where the organisation was unfamiliar. (23)

- community knowledge of how the data will be shared and used (19)
- trust in the regulatory mechanisms regulating data use and access in government organisations compared with private companies (19)
- lack of understanding or distrust of the motivation of private organisations (22, 25, 30, 31)

Channelling information through trusted entities could increase public acceptance of data sharing but some participants indicated that it could also erode trust in health care providers. (41)

> *Today I received the leaflet 'Better information means better care' together with a load of junk mail which I could have easily binned. I suspect many people will not give it a second look. There should have been some personal correspondence from one's GP practice informing patients about this rather than a mailshot. Comment posted to Care.data website, 21.01.14. (41, p.185)*

> *I do not trust the government with my data, and now I cannot trust my doctor o[r] the wider NHS. Comment posted to Care.data website, 05.05.14. (41, p.183)*

Government health data sharing with private industry where the private company only received aggregate results was generally more acceptable to public participants. (25, 38) Alternatively, participants suggested that private entities might gain trust if they were willing to "subject themselves to regulatory scrutiny". (25, p.56) In one study, participants suggested that private companies may be more responsive to customer feedback because of their commercial interests. This response reflected their experience with business transactions rather than the very specific instance of sharing health data. (19) In the same study participants expressed concern about businesses not acting for "public good". (19, p.13)

Summary

From 6,788 initial articles we identified a total of 23 peer-reviewed papers, seven reports, two conference proceedings and one conference paper which addressed community attitudes towards the use of government health data by private sector organisations for therapeutic development. Only a small number of papers internationally provided a quantitative estimate of public support for sharing data with the private sector with levels of support ranging 16-65%. (See Table 1)

This equivocal support was complicated by a complex suite of conditions that participants placed on sharing government health data with the private sector. Participants in the studies were concerned about data security, the potential for misuse and the fact that the private sector could make a profit from public data. They wanted to be confident that data sharing would only occur if the research were of public benefit (although views about what this meant varied), access to data were tightly controlled, the data were anonymised and securely stored, and there were rigorous governance structures to monitor and regulate access. Informed consent was also important, with studies that used deliberative methods to build understanding amongst participants more likely to find opt-out consent acceptable. Very few studies explicitly discussed a social licence or social contract to support data linkage and sharing, but many studies did emphasise the importance of trust for public acceptance.

In the studies in this review, public scepticism about the acceptability of data sharing for health research related in part to a lack of understanding of data uses in the health sector and data research in general. Participants in several studies called for better communication about data sharing, particularly through trusted entities, and a stronger program of public engagement.

We found no studies that examined Australian public views on the sharing of government health data with private industry for the purposes of therapeutic development.

Community Attitudes Survey

Building on existing research, including our own, on the use of linked administrative data in the public sector, we developed a community attitudes survey to explore community views on issues specifically related to the use of linked de-identified public administrative data by the private sector for therapeutic development. This section describes findings from the community attitudes survey.

Method

To develop the survey we carried out an extensive review of the literature and identified demographic and sociocultural factors that might influence how the public view sharing their personal health information with private industry (e.g. age, gender, health status, educational background and, experience working in the health industry). We searched the peer reviewed literature for existing tools to measure public attitudes to data sharing. We also summarised existing knowledge about patient attitudes, social licence and public interest in data sharing (see previous section). We developed a new instrument by combining existing questions from identified tools, with new questions drawing on insights from the literature.

Survey Monkey software was used to design an online version of the instrument. This was piloted with a convenience sample of the general population (n=10) aged 14 years and over. Pilot participants were selected to provide a diverse group with respect to age, gender, education, ethnicity, and presence or absence of long term illness. We asked participants to provide feedback on whether they understood each question, the design and layout as a whole, and were able to complete the survey under 6 minutes. We used the responses to refine the survey instrument, with the final survey taking approximately 9 minutes to complete. A copy of the final instrument is in Appendix 8.

The survey contained 29 items, including socio-demographic information (e.g. highest level of educational attainment) and health related information (e.g. long term health conditions) as well as possible experience with health data collection. We used the following five questions to assess views on sharing de-identified government health data with private companies:

- To what extent do you agree with the government sharing your health information with private companies, such as drug companies or medical device manufacturers? (Willingness include: to improve health services, for research, to develop new treatments and devices).
- What do you think about your health information being using by private companies for the development of new medicines or devices? (Options include: my information should not be used, able to opt in, able to opt out, I don't need to know).
- Would you like to be asked for your consent? (Options include: every time, just once, general consent)
- Imagine that the government has decided to share your health information with a private company. The company intends to use the information to help develop a new treatment for a disease. How important is it that each of the following conditions be met before the information is shared? (Conditions included: transparency, data storage, payment of data, purpose and strict rules and regulations)
- To what extent do you agree with the following statements about private companies using government health information to support development of new treatments? (Concerns included: trust in private companies, profits, purpose, secondary use of data without consent, possible re-identification and misuse of information.

A leading market research company McNair yellowSquares was employed to recruit a nationally representative sample of 2,500 participants by age, gender and location. McNair yellowSquares currently conducts community surveys for the NSW Department of Premier and Cabinet, ACT Health and the SA Population Health Survey. McNair yellowSquares only invited people who were registered on an Australian panel database held by the company.

After initial strata questions were presented (age, area of residence, gender), participants were provided with a half page summary of the topic explaining the concepts of data linkage, including potential benefits and risks. The 29 item instrument was then presented, with each survey question presented on a separate screen, followed by the demographic questions. To support participants understanding that each question referred to de-identified government health data, the following banner appeared at the top of each page: The questions below are about your government health information which has personal information removed, e.g. no name, no address, no date of birth, no Medicare number.

Data analyses

Data analysis was performed using Statistical Package for Social Sciences (SPSS). To produce proper population inference, we analysed the PHRN survey data using post stratification gender-by-age-by-state weights. We used the 2016 Australian Bureau of Statistics census data to obtain the Australian population characteristics of gender (2), age (4) and state (9) and calculated the survey weights based on the realized sample characteristics after we combined categories with small sample counts. Appendix 10 shows the counts of weighting characteristics from the survey data. A small coefficient of variation of the weights of 0.445 and a design effect of 1.198 suggest that the quality of the weights is reasonable for the subsequent analysis. The results in this report are gender-by-age-by-state weighted against the Australian population.

Findings

Demographics

Condensed population adjusted demographics are shown in Table 3. Please see Appendix 10 for full demographics.

Table 3: Participant population adjusted demographics (n=2,537)

Gender	N	%
Male	1,243	49.2
Female	1,285	50.5
Other	9	0.30
Region	**N**	**%**
Metro	1,641	63.2
Region	896	36.8
Employment	**N**	**%**
Full/part-time employed	1,481	59.6
Unemployed	120	5.5
Home duties	250	9.5
Student	112	6.9
Retired	456	14.4
Unable to work	107	3.5

Age	N	%
<29	552	38.6
30-49	873	26.7
50-64	652	20.2
65+	460	14.5
Education	**N**	**%**
No Educ/ Year 10	310	10.9
Year 12	422	18.3
Trade/Tafe	840	31.4
University	953	38.8
Self-rated Health	**N**	**%**
Poor/Fair	758	27.9
Good	991	38.3
Very Good/Excellent	788	33.9

Top Line Findings

The core finding across all questions is that, on average, respondents were equivocal about sharing health data with private companies. They tended toward support for strict controls, across a range of measures, if data were to be shared. This was a consistent trend across items that tested different dimensions of support for sharing.

General Willingness to Share

Figure 1 shows the degree of support for sharing health data for various purposes. Between 50 and 60% of all respondents were willing to share their data; fewer respondents were in favour of sharing information to improve health services.

Consent Preferences

Participants were asked for their consent preferences. There was a strong preference for 'opt in' consent (55%), which was more than three times more popular than any other option; 'opt out', 'my health information should not be used at all' and 'I don't need to know' all attracted 13% of preferences. For those who wanted to be asked for consent, 62.5% requested that they be asked 'every time', 23.6% requested 'get your general consent and be recontacted from time to time', while the remainder requested 'just once'. Figure 2 shows adjusted percentages of consent preferences.

Conditions on sharing

We gave participants a scenario in which the government had decided to share their health information with a private company and invited them to indicate how important various conditions would be for sharing their health information. The participants responded on a scale from 1-7 with the anchors 'Not important at all' and 'Very important', and 4 in the neutral position. For all statements but one about paying for data use, 80% of participants or more agreed that the condition was important. The one statement – private companies should pay for the use of the information – still had a majority of participants (61%) considering it to be important. Figure 3 shows participant responses to conditions of sharing government health data with the private companies.

Views about Private Companies

We provided a series of statements to assess participants' views about how private companies would use their health information. Respondents reported their level of agreement using a scale from 1 to 7 with the anchors 'strongly disagree' to 'strongly agree', and 4 in the neutral position, where 5-7 indicated broad agreement. There was wide variability in participants' responses to these statements. In general, the majority position was lack of trust in both companies and regulation. Just under half the respondents said that their data may be able to be re-identified and 23% did not think that re-identification was possible. Figure 4 shows participant views on sharing government health data with private companies

Figure 1: Support for sharing government health data with the private sector (n=2,537)

To what extent do you agree with the government sharing your health information with private companies, such as drug companies or medical device manufactures?

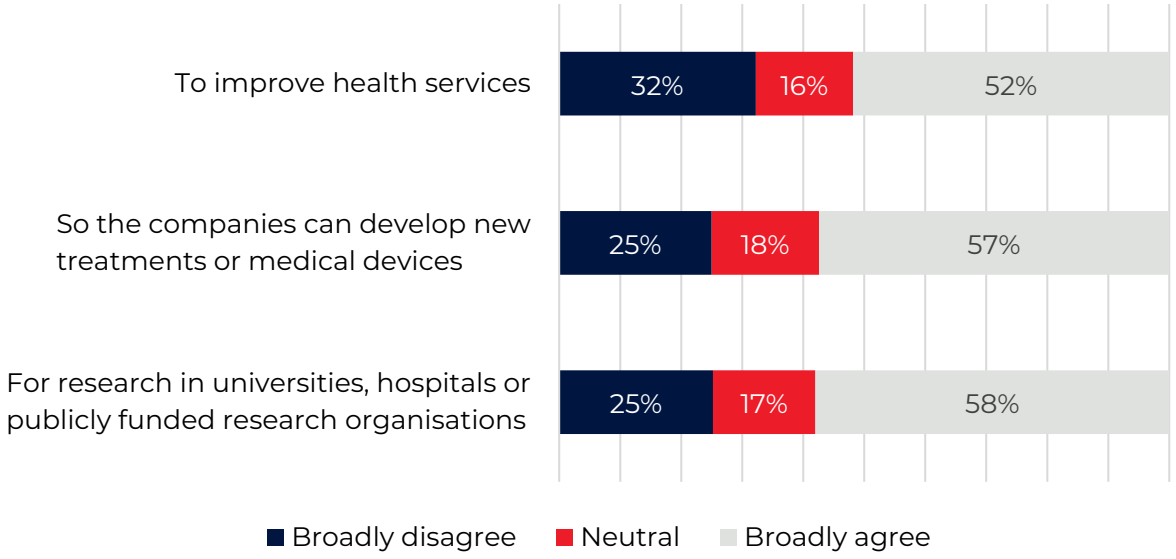


Figure 2: Adjusted percentages of consent preferences (n=2,573)

What do you think about your health information being used by private companies for the development of new medicines or devices?

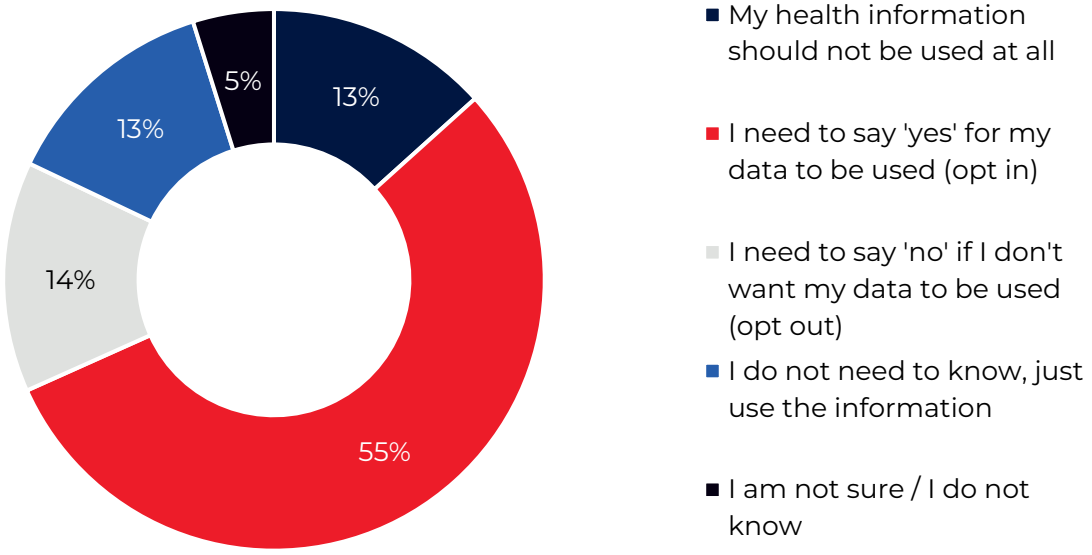


Figure 3: Adjusted percentages of conditions on sharing government health data with the private companies (n=2,537)

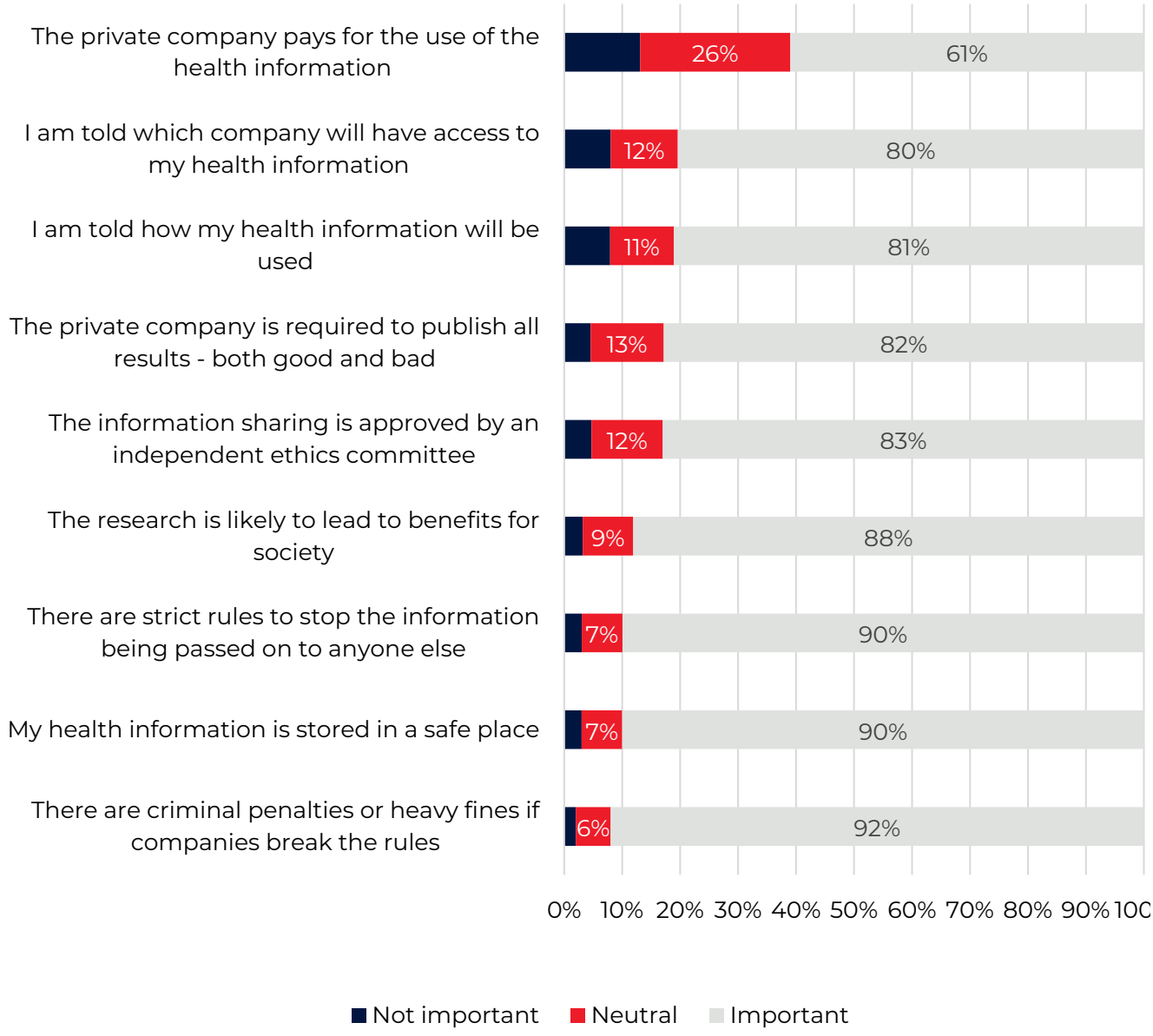


Figure 4: Adjusted percentages of views on sharing government health data with private companies (n=2,537)

To what extent do you agree with the following statements about private companies using government health information to support development of new treatments?

Sociodemographic patterning of responses

We investigated the impact of various sociodemographic variables on the participants' views on whether health information should be shared, and the conditions under which it might be acceptable (Table 4-6).

In some cohorts, there appeared to be no associations: for example, employment status, the presence of a chronic health condition, being a carer for someone with a chronic health condition, taking prescribed medication and working in the health industry were not associated with attitudes toward data sharing. However, there were sociodemographic patterns for several variables, including gender, age, region and education. Self-reported health status was also associated with the respondents' views. Detailed discussions of the sociodemographic patterning of responses are provided following Table 6.

Table 4 illustrates the sociodemographic patterning of responses to willingness to share government health data with private companies.

Table 5 illustrates the sociodemographic patterning of responses to conditions of sharing government health data with private companies.

Table 6 illustrates the sociodemographic patterning of responses to views on sharing government health data with private companies.

In Tables 4-6, the interpretation of the colours and shades are twofold. Dark blue indicates a large proportion in favour, while dark red indicates that the majority did not support a specific statement. A cell with a light shade of colour indicates that the proportion of supportive responses was around 50%.

Table 4: Adjusted percentages of willingness to share government health data with the private companies by socio demographic patterning (n=2,537)

To what extent do you agree with the government sharing your health information with private companies, such as drug companies or medical device manufacturers?

		To improve health services	For research in universities, hospitals or publicly funded research organisations	So the companies can develop new treatments or medical devices
Gender	Male	55.2%	61.6%	59.8%
	Female	48.7%	54.4%	55.3%
	Indeterminate/Intersex/Trans/ Gender diverse	34.4%	50.2%	33.0%
Age	18-29	49.2%	54.1%	56.4%
	30-44	49.6%	54.5%	54.9%
	45-59	52.2%	58.5%	55.4%
	60+	58.6%	67.8%	63.7%
Region	Metro	52.5%	59.4%	57.3%
	Regional	50.5%	54.7%	57.7%
Self-Reported Health Status	My health is poor	50.4%	52.3%	52.1%
	My health is fair	49.3%	56.4%	56.4%
	My health is good	50.3%	56.0%	56.3%
	My health is very good	55.0%	62.1%	60.0%
	My health is excellent	57.4%	61.6%	60.6%
Educational Level	No formal qualifications	60.2%	66.7%	64.5%
	Year 10 or school certificate	51.1%	52.7%	58.9%
	Year 12 or leaving certificate	50.4%	58.5%	57.4%
	Trade / apprenticeship	53.1%	60.6%	59.6%
	Other TAFE / Certificate	49.9%	57.2%	57.6%
	University degree / Higher degree	53.8%	59.3%	56.8%
Employment Status	Full time employed	54.0%	58.7%	58.7%
	Part time employed	50.5%	57.5%	56.0%
	Unemployed	45.6%	49.6%	53.1%
	Home duties	42.6%	48.9%	52.9%
	Student / Training	49.4%	56.6%	51.4%
	Retired	58.4%	69.3%	65.4%
	Unable to work (e.g. disability / Work Cover)	53.6%	50.5%	53.7%
Past/current employment in health industry and/or health services/ research?	Yes	52.3%	62.7%	62.8%
	No	52.0%	57.2%	56.8%
	I am not sure	42.6%	66.6%	53.4%
	I prefer not to answer	34.7%	40.5%	41.1%
Health status / chronic health condition	Yes	56.0%	64.0%	61.9%
	No	51.0%	56.6%	56.6%
	I am not sure	45.9%	51.5%	51.5%
Care for someone with chronic health condition	Yes	50.4%	61.9%	59.0%
	No	51.8%	57.1%	56.8%
	I am not sure	58.4%	66.6%	69.4%
Prescribed medications usage	Yes	53.8%	61.4%	60.4%
	No	50.4%	55.5%	55.0%
	I am not sure	42.0%	36.9%	54.3%
Has MyHealthRecord	Yes	59.3%	67.4%	68.1%
	No	42.1%	49.5%	47.7%
	I am not sure	54.7%	55.4%	55.0%

How important is it that each of the following conditions be met when information is shared with the private sector?

		I am told how my health information will be used	I am told which company will have access to my health information	My health information is stored in a safe place	The private company pays for the use of the health information	The information sharing is approved by an independent ethics committee	The private company is required to publish all results - both good and bad	The research is likely to lead to benefits for society	There are strict rules to stop the information being passed on to anyone else	There are criminal penalties or heavy fines if companies break the rules
Gender	Male	78.6%	78.7%	88.5%	64.3%	81.8%	83.0%	86.7%	88.1%	89.6%
	Female	83.3%	82.2%	92.3%	57.6%	84.3%	83.1%	89.7%	91.1%	93.2%
	Indeterminate /Intersex/Trans /Gender diverse	83.4%	75.6%	83.4%	75.3%	66.7%	59.2%	66.7%	66.7%	83.4%
Age	18-29	81.3%	81.3%	88.0%	60.0%	77.3%	78.4%	84.4%	85.9%	88.7%
	30-44	80.4%	80.7%	88.8%	62.7%	82.4%	81.6%	85.3%	88.3%	89.0%
	45-59	85.0%	83.9%	94.2%	64.4%	88.9%	86.1%	92.4%	93.6%	94.1%
	60+	77.5%	75.7%	93.0%	58.0%	88.6%	89.8%	93.8%	93.8%	96.1%
Region	Metro	81.9%	81.2%	90.2%	61.9%	83.5%	82.6%	88.2%	89.4%	90.8%
	Regional	79.2%	78.8%	91.0%	59.1%	82.0%	83.8%	87.8%	89.9%	92.7%
Self-Reported Health Status	My health is poor	76.6%	73.1%	83.2%	59.5%	83.7%	85.5%	86.9%	87.8%	91.7%
	My health is fair	81.9%	81.5%	89.8%	60.2%	82.0%	83.7%	87.5%	90.6%	88.9%
	My health is good	79.9%	80.0%	90.3%	61.0%	83.3%	82.1%	87.2%	90.0%	91.4%
	My health is very good	82.7%	82.3%	92.8%	61.2%	85.0%	82.8%	91.2%	90.3%	93.6%
	My health is excellent	81.0%	77.8%	89.4%	63.6%	77.7%	84.1%	85.5%	83.3%	90.9%
Educational Level	No formal qualifications	86.6%	86.7%	91.0%	71.1%	93.4%	86.6%	89.1%	88.9%	97.8%
	Year 10 or school certificate	81.3%	81.6%	88.4%	58.3%	83.0%	84.5%	90.2%	89.7%	93.2%
	Year 12 or leaving certificate	76.9%	76.7%	86.2%	57.6%	81.5%	80.6%	84.9%	86.8%	87.9%
	Trade / apprenticeship	77.1%	78.8%	94.5%	62.3%	84.6%	86.2%	88.1%	89.6%	91.9%
	Other TAFE / Certificate	81.5%	81.0%	93.1%	57.5%	81.7%	81.7%	88.7%	90.8%	92.7%
	University degree / Higher degree	83.2%	81.7%	90.4%	64.9%	84.1%	83.9%	88.9%	90.3%	91.2%

Table 5: Adjusted percentages of conditions on sharing government health data with the private companies by socio demographic patterning (n=2,537)

		I am told how my health information will be used	I am told which company will have access to my health information	My health information is stored in a safe place	The private company pays for the use of the health information	The information sharing is approved by an independent ethics committee	The private company is required to publish all results - both good and bad	The research is likely to lead to benefits for society	There are strict rules to stop the information being passed on to anyone else	There are criminal penalties or heavy fines if companies break the rules
Employment Status	Full time employed	81.4%	82.2%	91.0%	65.0%	84.8%	83.2%	87.8%	90.1%	91.0%
	Part time employed	80.7%	79.0%	88.8%	58.2%	82.1%	82.1%	86.1%	88.5%	89.4%
	Unemployed	79.0%	78.9%	86.5%	62.7%	64.1%	66.7%	79.5%	80.0%	84.1%
	Home duties	84.8%	84.6%	92.1%	55.5%	78.1%	83.3%	89.3%	89.7%	92.7%
	Student / Training	86.8%	84.5%	90.7%	59.7%	85.5%	82.0%	89.2%	87.6%	91.4%
	Retired	76.5%	75.8%	93.1%	57.4%	88.6%	90.3%	94.3%	93.5%	95.9%
	Unable to work (e.g. disability / Work Cover)	81.6%	76.2%	86.4%	66.4%	87.4%	84.9%	88.5%	94.2%	95.6%
Past/current employment in health industry and/or health services/ research?	Yes	80.4%	80.4%	91.3%	60.0%	87.4%	84.8%	91.1%	91.6%	94.6%
	No	81.5%	80.8%	90.7%	61.1%	82.7%	83.2%	88.0%	89.7%	91.2%
	I am not sure	46.6%	49.6%	64.6%	54.9%	53.8%	42.0%	69.8%	61.9%	65.3%
	I prefer not to answer	75.6%	82.0%	70.1%	75.6%	70.1%	64.2%	63.5%	70.1%	82.0%
Health status / chronic health condition	Yes	81.0%	80.2%	91.5%	57.0%	87.3%	87.2%	91.4%	91.9%	92.8%
	No	80.7%	80.4%	90.4%	61.9%	81.7%	82.2%	87.2%	89.0%	91.0%
	I am not sure	84.3%	82.4%	87.2%	65.3%	83.0%	76.8%	87.2%	88.1%	90.1%
Care for someone with chronic health condition	Yes	76.5%	78.4%	91.0%	59.9%	83.1%	84.5%	90.2%	90.1%	88.0%
	No	81.7%	80.6%	90.5%	60.9%	82.9%	82.9%	87.8%	89.9%	91.9%
	I am not sure	80.2%	85.5%	84.5%	67.7%	84.8%	78.5%	88.9%	77.1%	88.3%
Prescribed medications usage	Yes	81.5%	79.0%	90.5%	57.3%	85.7%	85.2%	89.7%	91.6%	92.8%
	No	80.9%	81.7%	90.8%	64.0%	80.9%	81.6%	87.1%	88.4%	90.2%
	I am not sure	70.2%	80.7%	76.9%	67.7%	74.8%	63.3%	74.2%	69.6%	87.0%
Has MyHealth Record	Yes	81.6%	78.7%	92.3%	58.2%	84.3%	84.8%	92.0%	93.3%	94.5%
	No	81.7%	83.3%	88.4%	65.2%	82.0%	81.7%	84.1%	85.7%	88.5%
	I am not sure	78.8%	79.0%	90.6%	59.0%	82.4%	81.8%	87.9%	89.5%	90.6%

Table 6: Adjusted percentages of views on sharing government health data with the private companies by socio demographic patterning (n=2,537)

To what extent do you agree with the following statements about private companies using government health information to support development of new treatments?

		Private companies can be trusted to store health information safely	Private companies should be allowed to make a profit from the use of this information	Private companies can be trusted to act for the good of society	If you give health information to a private company, you cannot control where it ends up	Someone may be able to work out who I am even though my personal information has been removed	The government won't be able to stop private companies from misusing this information, even if they try
Gender	Male	36.4%	32.2%	38.1%	63.1%	49.7%	59.7%
	Female	32.4%	18.0%	29.0%	58.6%	46.0%	57.5%
	Indeterminate /Intersex/Trans /Gender diverse	16.4%	8.3%	16.4%	67.3%	49.2%	66.9%
Age	18-29	37.2%	25.0%	35.8%	53.0%	47.8%	53.1%
	30-44	33.9%	24.4%	32.5%	60.1%	49.9%	57.9%
	45-59	31.3%	25.1%	31.1%	68.2%	51.1%	65.3%
	60+	32.2%	25.2%	32.1%	69.1%	42.8%	63.4%
Region	Metro	34.4%	26.1%	34.4%	60.4%	48.6%	59.1%
	Regional	34.1%	22.4%	31.3%	61.9%	46.0%	57.8%
Self-Reported Health Status	My health is poor	21.5%	16.1%	15.5%	75.5%	54.8%	69.1%
	My health is fair	32.2%	23.8%	32.6%	64.0%	50.8%	62.0%
	My health is good	35.0%	25.3%	34.6%	61.2%	46.2%	59.2%
	My health is very good	37.0%	26.7%	36.1%	56.4%	45.5%	54.2%
	My health is excellent	35.3%	25.7%	31.8%	56.3%	50.1%	54.8%
Educational Level	No formal qualifications	33.4%	26.7%	37.9%	64.3%	57.4%	64.3%
	Year 10 or school certificate	30.2%	18.8%	28.6%	63.2%	47.1%	55.7%
	Year 12 or leaving certificate	33.9%	23.2%	31.1%	55.0%	43.8%	53.6%
	Trade / apprenticeship	33.8%	26.9%	35.3%	64.6%	50.6%	62.8%
	Other TAFE / Certificate	33.8%	23.5%	34.7%	60.6%	46.3%	59.2%
	University degree / Higher degree	36.3%	28.2%	34.7%	63.0%	50.1%	60.9%

		Private companies can be trusted to store health information safely	Private companies should be allowed to make a profit from the use of this information	Private companies can be trusted to act for the good of society	If you give health information to a private company, you cannot control where it ends up	Someone may be able to work out who I am even though my personal information has been removed	The government won't be able to stop private companies from misusing this information, even if they try
Employment Status	Unemployed	32.7%	17.2%	30.5%	61.2%	51.3%	55.1%
	Home duties	28.3%	20.8%	29.7%	58.6%	51.3%	57.7%
	Student / Training	34.7%	27.9%	31.6%	51.1%	50.3%	55.9%
	Retired	33.2%	27.1%	33.1%	67.1%	43.4%	63.4%
	Unable to work (e.g. disability/ Work Cover)	33.5%	15.9%	25.8%	68.3%	49.1%	62.4%
Past/current employment in health industry and/or health services/ research?	Yes	35.2%	22.9%	33.5%	64.6%	51.8%	61.3%
	No	34.5%	25.1%	33.4%	60.5%	47.3%	58.2%
	I am not sure	19.2%	43.5%	38.4%	47.5%	37.7%	55.6%
	I prefer not to answer	11.3%	17.5%	23.3%	52.5%	40.8%	69.4%
Health status / chronic health condition	Yes	32.1%	22.4%	30.3%	68.8%	49.4%	61.5%
	No	35.7%	25.8%	34.9%	58.3%	47.3%	57.4%
	I am not sure	26.6%	24.1%	28.3%	60.8%	48.1%	63.0%
Care for someone with chronic health condition	Yes	34.3%	28.2%	35.3%	60.6%	50.8%	61.2%
	No	34.3%	24.5%	33.1%	60.9%	47.2%	58.0%
	I am not sure	33.7%	25.2%	35.5%	61.3%	54.7%	68.7%
Prescribed medications usage	Yes	32.5%	24.9%	33.2%	65.9%	46.9%	61.3%
	No	35.9%	24.4%	33.6%	56.5%	48.5%	56.0%
	I am not sure	33.4%	43.7%	35.4%	57.8%	49.8%	68.6%
Has MyHealthRecord	Yes	38.8%	28.1%	36.6%	61.3%	44.0%	57.7%
	No	31.2%	24.9%	32.2%	62.0%	55.5%	60.4%
	I am not sure	31.7%	19.5%	29.9%	58.2%	41.7%	57.4%

Gender

On average, women appeared less willing to share their health data and were more likely to place conditions on sharing. For example, women were less supportive of sharing data with private companies for all three purposes – to improve health services, for research, and so companies could develop new treatments and devices. When asked about how important a range of conditions would be for releasing health information to private companies, they rated the importance of all but one condition more important than men. Fewer women agreed that private companies should be able to make a profit from data, and that companies could be trusted to act for social good. Conversely, though, women were less likely to say one could not control where health information would end up if shared with a private company.

Age

Overall, older people (60+ years) were more willing to share their health information with private companies to improve health services, for research and to develop new treatments. They were less concerned than younger people about knowing which companies would have access to their data, but rated as more important that information be stored safely, that negative results be published, that there should be strict rules against passing information on to third parties, and that release of information should lead to public benefits. The three oldest age groups were more supportive of criminal penalties and the youngest age group were the ones least likely to support ethics committee oversight.

Region

People living in regional areas were less likely to support data sharing for research, and were more likely to consider it important to store information safely, to have strict rules to stop third party sharing, and to have more criminal sanctions. They were also less supportive of profit making by private companies.

Education

Level of education was associated with the importance of various conditions placed on the release of health data. Participants who had completed at least year 12 as their highest level of education were less likely than those who had not completed school to consider it important to know how their information would be used. They were more likely than less well educated participants to report that secure storage, oversight by an ethics committee, publication of all results (both good and bad), and controls on third party sharing were important when sharing data with private companies. Participants with higher levels of education were also more cautious about whether private companies could be trusted to act for the good of society and less likely to agree one could control where the information would end up.

Self-reported health status

Participants with poor self-reported health generally were more concerned with sharing health information. They were less supportive of sharing data to improve health services, academic research and developing new treatments or devices. A significantly smaller proportion of them showed confidence in private companies in terms of health data storage and access security, acting for the good of society and misuse of information, compared to other groups. Fewer of those in the 'poor health' group agreed to allow private companies to make a profit from using their health information.

Impact of consent preferences on conditions and views of sharing government data with private companies

We also assessed whether participants' conditions on and views about sharing government health data with private companies were related to their views about 'opt-in' or 'opt-out' consent.

Consent preferences on conditions of sharing

For most of the questions regarding the conditions of releasing health information to private companies to help develop new treatments, there were large differences in ratings of importance for the subgroup that responded 'opt in' and the subgroup that responded 'opt out'. Table 7 shows the responses of Question 4 stratified by the two response groups from Question 2, namely, 'opt in' (option b) and 'opt out' (option c). Large differences (more than 5%) in adjusted percentage of agreement between the 'opt in' and 'opt out' groups are highlighted in red. For the questions 4a, b, c, e and h, more than 5 percent people in the 'opt in' group rated the statements important (Likert scale 5, 6 or 7) comparing to the 'opt out' group. In questions regarding paying for use the health data (Q4d), benefits to society (Q4g) and criminal penalties for breaking rules (Q4i), both 'opt in' and 'opt out' participants share similar views.

Table 7: Adjusted percentages of question 4 stratified by question 2 ("opt in" and "opt out")

How important is it that each of the following conditions be met when information is shared with the private sector?'	Opt In (n=1,356)	Opt out (n=352)
	Percentage (%)	Percentage (%)
I am told how my health information will be used (Q4a)	89.6	80.6
I am told which company will have access to my health information (Q4b)	87.7	80.6
My health information is stored in a safe place (Q4c)	94.8	88.7
The private company pays for the use of the health information (Q4d)	62.5	59.7
The information sharing is approved by an independent ethics committee (Q4e)	87.5	81.2
The private company is required to publish all results – both good and bad (Q4f)	87.0	81.9
The research is likely to lead to benefits for society (Q4g)	91.9	86.3
There are strict rules to stop the information being passed on to anyone else (Q4h)	94.0	88.0
There are criminal penalties or heavy fines if companies break the rules (Q4i)	95.1	89.9

Consent preferences on views of private companies

Table 8 shows the responses of Question 5 stratified by the two responses groups from Question 2, namely, 'opt in' (option b) and 'opt out' (option c). Differences of more than 5% in adjusted percentage of agreement between the 'opt in' and 'opt out' groups are highlighted in red. There were no clear differences between the two groups for most questions. However, there were two exceptions: a larger percentage of 'opt in' respondents agreed with the statement that they have no control of the data if the health information is given to a private company and with the possibility of identification even with removal of personal information.

Table 8: Adjusted percentages of question 4 stratified by question 2 response 'opt in' and response 'opt out'

To what extent do you agree with the following statements about private companies using government health information to support development of new treatments?	Opt In (n=1,356)	Opt out (n=352)
	Percentage (%)	Percentage (%)
Private companies can be trusted to store health information safely (Q5a)	36.4	33.7
Private companies should be allowed to make a profit from the use of this information (Q5b)	25.5	27.8
Private companies can be trusted to act for the good of society (Q5c)	34.9	29.9
If you give health information to a private company, you cannot control where it ends up (Q5d)	64.0	57.4
Someone may be able to work out who I am even though my personal information has been removed (Q5e)	50.1	44.6
The government won't be able to strop private companies from misusing this information, even if they try (Q5f)	60.3	56.6

Additional opinions from the open-ended question

The last survey question asked "Is there anything else you would like to tell us about your views on sharing government health information with private companies where the goal is to support the development of new treatments for diseases and disabilities?" Just under half (46% of all respondents) provided comments, primarily describing concerns about sharing government health information and conditions under which they would support sharing. The responses to Question 6 demonstrated the polarised nature of this debate with many participants expressing strong views in support of unconditional sharing of health information while many opposed it vehemently.

The most commonly raised concern was lack of trust in both private companies and the government. The respondents gave corporate interests, corruption and profit-making as the reasons for their general distrust of private companies. They expressed reluctance to share health information with private companies, if the end goal is profit generation and not societal benefit. In addition, the respondents referenced the poor track record of government in handling data and they questioned the ability of government to keep their data secure and prevent misuse. Support for regulated access to health information was linked to respondents' concerns about security.

The respondents explained that, if government health information is to be shared with private companies, certain conditions need to be met. The most common requirement was anonymisation of health information and a guarantee that all personal information be removed. In addition, a large subset of participants believed that data sharing needed to deliver public benefits or support the common good. They provided examples of public benefit, including developing new treatments, finding cures or improving the health of society. Giving consent was a prerequisite to sharing health information for many participants and the right to 'opt-in' rather than 'opt out' was highlighted by a subset. Please see a full analysis of question six in Appendix 11.

Summary

Overall, respondents were ambivalent about government sharing their de-identified government health data with private companies, with just over half of all respondents supporting private company use of health information. A similar proportion of participants also wanted an opt-in method of consent, which most researchers and data custodians would judge incompatible with large-scale data linkage activities. Overall, women, younger people, less well-educated people, people living in regional areas and, to some degree, people with poorer health status, were more concerned to impose conditions on release of health information. There was a very wide range of views about how private companies might use health information, suggesting that Australian society may be a long way from reaching consensus about allowing access by private companies to government health data.

This survey was conducted with members of the general public who had signed up to be part of a pre-existing online panel invited to participate in research. Therefore it carries a number of limitations. In particular, it may be likely that those who sign up to research panels are more supportive or at least more interested in research then the general public. The respondents also probably had a reasonable level of confidence in using information technology and felt comfortable using the internet, although what this means for their attitudes towards sharing their health data with the private sector is unclear.

Another limitation relates to the fact that most people have a limited understanding of data sharing, de-identification and current data linkage practices. (18) Despite our efforts to provide a clear description of data linkage in the introduction, participants may still have had difficulty in understanding how linked administrative health data may be used for research and development. This factor may have affected their capacity to understand and respond to the questions in the survey. The lack of understanding was evident in some of the qualitative comments in this report; for example, one participant wrote, "I don't think that that information is any use to anybody for developing new drugs or procedures".

Some caution should therefore be used in interpreting the findings in the survey.

Hypothetical Case Studies

The scoping review and, to some degree, the survey found that public understanding of how the private sector might use government data in the public interest is poor. The scoping review also found that greater and more detailed communication and engagement with the public is important to build public trust in both government and the private sector.

Clear examples of the kinds of therapeutic development research conducted by private sector organisations that might be in the public interest are an important part of building the case to share government health data with the private sector. Realistic examples that provide opportunities to debate what is 'in the public interest' in a specific case can also provide a means to step beyond the rhetoric around the 'public interest' that we noted in the first section of the report. In turn, this may support the development of governance structures, policies and guidelines that can meet the needs of the public and private sectors as we move toward greater sharing of data.

To support the PHRN in its work in this area, we interviewed a number of private sector stakeholders to develop hypothetical case studies (hypotheticals) to illustrate the ways in which government data might be shared with private companies for the purpose of developing new treatments for diseases and disabilities that are in the public interest.

The purpose of the hypothetical case studies was to:

- assist the PHRN to better understand the value of private company access to government health data, and what is deemed in the public interest;
- provide resources to use in future research; and
- illustrate to a range of stakeholders including the community, governments and researchers how linked data could be used by the private sector.

The hypotheticals are presented as written vignettes, with no person or organisation identified.

Method

The development of the hypotheticals was informed by the scoping review, engagement with the PHRN and interviews with relevant stakeholders.

Recruitment

Relevant stakeholders were identified at an advisory meeting in April 2018, whereby Biointelect facilitated a discussion between the PHRN principles, technical experts and other industry experts on data needs from each of the target segments: market access for medicines, market access for devices and data needs for clinical trials. The purpose of the meeting was to gain input from a small number of expert potential data users on the types of data required and where / how compromises might be made to enable a viable match between PHRN capabilities / governance and industry needs. Researcher ABM attended the meeting and was provided with a list of attendees who agreed to be contacted to be interviewed.

Twelve attendees were approached via an email invitation to participate (Appendix 12), and six expressed interest in being interviewed.

Data collection

The focus of the stakeholder interviews was to collect rich accounts of each stakeholder's experiences and/or augment information from other sources about the feasibility and relevance of international case studies in the Australian context. The research team focused on elucidating examples of therapeutic development in the public interest rather than the more complex study of the contextual dynamics within which they occur. A copy of the interview guide may be found in the Appendix 13.

Six interviews were conducted by ABM and JS by Zoom on the University of Wollongong Campus. All interviews took less than one and a half hours and were recorded but not transcribed.

Analysis

The interview data was analysed using a narrative approach drawing on close listening to the audio, interview reflections between research team members and relevant documents identified in the scoping review. Methods described by Riessman in Narrative Analysis, (1993) were used. (45) Riessman recognizes that narratives are subjective but suggests that it is "precisely because of their subjectivity – their rootedness in time, place and personal experience – that we value them". (45, p.5) The hypothetical case studies were constructed using the stories of the participants refracted against stories drawn from the literature. International cases were adapted to the Australian context based on the feedback from research participants.

There is no standard way to do narrative analysis; however, in general the research team used Labov's structural framework of orientation (time, place, and participants), complicating action (sequence of events), evaluation (significance and meaning), resolution (outcome) and coda (relevance to the issues). During the development of the hypotheticals the research team ensured each case study met the following criteria:

- Related to the use of administrative data for therapeutic development;
- Illustrated noteworthy social and ethical issues; and
- Relevant to the Australian context.

Participants of the stakeholder interviews received a copy of the final draft of the hypothetical case studies and asked to provide feedback. This included the opportunity to comment on and, if necessary, address any issues associated with inadvertent identification of themselves or any organization. Two participants provided feedback regarding content.

Results

The participants suggested four areas in which private sector access to government health data in Australia might be in the public interest:

- To support PBAC/MSAC submissions with enhanced evaluation of new drugs and devices
- To monitor treatment patterns to support equitable access to new cutting-edge technologies (for example, to facilitate more consistent treatment patterns across Australia)
- To monitor drugs and devices after release onto the Australian market to detect adverse events and safety concerns
- To provide data for comparative effectiveness research

Four abbreviated hypotheticals that address these areas are provided below. The full versions are in Appendix 14, together with supplementary questions that amend the scenario or raise further issues for discussion.

Hypothetical 1: Submission to support inclusion of a new drug on the Pharmaceutical Benefits Scheme using Cancer Registry data

An international pharmaceutical company wishes to bring its newly developed oncology drug to the Australian market. The drug extends life and has fewer side effects than existing drugs currently provided through the government-subsidised Pharmaceutical Benefits Scheme (PBS) for the same condition. The company approaches a Cancer Registry based in a State Department of Health with a request for aggregated, and therefore de-identified, data for which they will pay. They would like to have current treatment patterns in Australia, including by age of onset and additional treatments, and information on variations in care provision across Australia.

Patient data held in the Cancer Registry are collected, without patient consent, from pathology laboratories, hospitals, radiotherapy and medical oncology departments, aged care facilities and the Registry of Births, Deaths and Marriages. The data collection is authorised under an Act of Parliament.

The Cancer Registry provides the company with aggregate data which is sufficient for them to make a successful submission to have the drug funded through the PBS.

Hypothetical 2: Monitoring safety of therapeutic devices

A range of therapeutic implantable devices are made available in Australia for a debilitating condition. Initially the availability of the new treatment is widely welcomed by patient groups. The devices are funded by Medicare as part of a surgical procedure.

Within a year, it has become apparent that the devices can have serious side effects including chronic pain and infection but this knowledge is not widely disseminated. Some physicians and patient advocates call for controlled trials of the use of the devices. Health Technology Assessment bodies call for the devices to be 'archived' because of lack of evidence to support their safety. It is several years before the TGA acts on the adverse event reports submitted by patients and physicians to withdraw approval for the device. In light of these events, a clinician group, with support from government, establishes a registry to monitor similar devices.

A device company wishes to undertake ongoing monitoring of a recently approved device for treatment of a similar condition. Their clinical trials suggest this treatment is safe and effective but, given the recent history, the company is cautious. Monitoring the device will require linking of hospital emergency attendances, hospital surgery data, Medicare and PBS data and data from the registry. The company engages a university research group to undertake the data collection and analysis. Data linkage is undertaken by a government funded data linkage unit. This process takes 12 months because of delays with human research ethics committee and governance approvals. The new device proves to be associated with some adverse events, but only in a particular subset of patients. The indications for the use of the device are adjusted in the Medicare Benefits Schedule so that the device is no longer used for these patients.

Hypothetical 3: Sharing aggregate data through public websites

Health insurance companies are requesting access to aggregate estimates of particular patient outcomes, including more detailed information on disease survival times and associated complications. This data will help insurance companies to estimate future risk.

The government believes that a vibrant private health insurance sector is essential to reduce strain on public health services. It argues that aggregate health data should be available through government websites to assist the insurance industry.

Hypothetical 4: Release of 'de-identified' health data

In 2016 the Australian Department of Health released a de-identified data set containing 10% of Medicare Benefits Schedule and Pharmaceutical Benefits Scheme data. The data consisted of claims information made through the MBS since 1984 and through the PBS since 2003.

The release of the data was welcomed by researchers and consumer groups as an important tool for health systems research. The data was used in a range of ways by private companies, including to:

- Provide more detailed information in their submissions for public funding for new devices and drugs; and
- Identify subgroups of patients who were not receiving international standards of care for their condition.

One month after the release of the data, researchers at the University of Melbourne demonstrated that the encryption could be broken so that Medicare services provider numbers could potentially be identified. They also showed that some claimants could be identified by linking the dataset to other sources of information such as Facebook.

Conclusion

The aim of this project was to examine community attitudes towards government sharing health data with private companies for research and development of treatments for disease and disability. As we noted, what the public thinks about sharing data with the private sector is not same thing as acting 'in the public interest'. However, public sentiment both reflects and shapes specific judgments about the public interest and so we need to take what publics say about sharing government health data seriously.

We found no Australian studies that provided a quantitative estimate of public support for sharing data with the private sector. The small number of international studies that we found put support for data sharing at between 16 and 65%. Our survey of public support for sharing government health data with the private sector found a similar level of support, ranging from 52% to 58%.

Both the international literature and our survey found that the level of public support for sharing government health data with the private sector is bounded by a range of concerns. The respondents to our survey were concerned about data security and misuse, and the scoping review provided context for these views, indicating that people are concerned about the possibility of discrimination, surveillance and stigmatisation. We included an item in the survey about private sector payment for public data because this was an important concern in the international literature. We found that this was important, but not as important as other factors, for our survey respondents.

The flipside of these concerns was a complex suite of conditions placed on sharing data. These included (in both survey and scoping review): controls on access to data; maintenance of privacy and confidentiality; rigorous governance and regulatory structures; transparency about uses for shared data; and, most important of all, public benefit. However, views about how 'public benefit' can be defined varied across the studies we reported in the scoping review, paralleling the ill-defined nature of the term 'public interest' generally and the lack of studies explicitly discussing a social licence or social contract to support data sharing. For the private sector participants we interviewed, 'public interest' turned primarily on the possibility of benefit to patients through access to new treatments, broader access to existing treatments, and enhanced monitoring.

The scoping review also emphasised the importance of trust for public acceptance. Since the scoping review was completed, we have identified six new articles relevant to this research that also emphasise the importance of public benefit and trust. These papers support our findings that, provided the public benefit is clear, the public is generally comfortable with the use of anonymised government health data in research and service delivery. (46) However many people are still uncomfortable with the idea of private companies accessing their government health data. (47) There are particular concerns about passing information on for marketing or insurance purposes (46), with concerns about data privacy being key. (48) Trust in the Australian government, or lack thereof, was identified in two of the studies, highlighting concerns that government infrastructure and people lack the capacity to implement and manage data sharing and linkage adequately, both in general and with private companies. (47, 48) Tully et al suggest that, the more informed people feel, the more they are likely to support potential future uses of government health data by private companies. (49) Tully recommended that activities undertaken to share government health data with private companies must make the public benefit explicit.(49)

The research outlined in this report, and recent studies, suggest that sharing government health data with private industry will require concerted and nuanced public engagement. Both government and the private sector will need to address the public's lack of understanding and lack of trust in the ways in which agencies collect, share, protect and use their personal data. We will need transparent, interactive and informed engagement that takes into account the capacity for and barriers to engagement.

One of the outcomes of sustained engagement will be a better understanding of public views about sharing data in specific situations. A range of approaches will be needed to gain this understanding: population surveys; focus groups, particularly in vulnerable populations; public forums; publicly-focused websites for engagement and feedback; citizens' councils; and deliberative informed events such as citizens' juries. These strategies would provide information about public concerns and public values and would be crucial to the development of public understanding and a social licence for data sharing. This engagement would require public investment.

In addition to sustained engagement with the Australian public, government and the private sector will also need to do a better of job of making the public benefit in sharing data explicit. In part, this will be addressed by enhancing the public's understanding of how and why government health data are collected and used. It will also require deeper analysis of the meaning of 'public interest' through both conceptual and empirical work.

References

1. Degeling C, Carter SM, Rychetnik L. Which public and why deliberate? – A scoping review of public deliberation in public health and health policy research. Social Science & Medicine. 2015;131:114-21.
2. Lo B, Field, M. Conflict of interest in medical research, education, and practice. Washington: National Academies Press (US); 2009.
3. Australian Law Reform Commission. Serious invasions of privacy in the digital era. ALRC discussion paper 80. Discussion Paper. Sydney, NSW; 2014.
4. Wheeler C. What is the public interest? Government Solicitors Conference 6 September; Sydney: NSW Government Publication; 2016.
5. Laurie G, Stevens L. Developing a public interest mandate for the governance and use of administrative data in the United Kingdom. Journal of Law and Society. 2016;43(3):360-92.
6. Daniels N, Sabin JE. Accountability for reasonableness: an update. British Medical Journal. 2008;337:a1850.
7. National Health and Medical Research Council. Guidelines approved under section 95A of the privacy act 1988. 2015.
8. National Health and Medical Research Council. Guidelines approved under section 95 of the privacy act 1988. 1988.
9. Carter SM, Degeling C, Doust J, Barratt A. A definition and ethical evaluation of overdiagnosis. Journal of Medical Ethics. 2016;42(11):705-14.
10. Widdows H, Cordell S. Why communities and their goods matter: illustrated with the example of biobanks. Public Health Ethics. 2011;4:14-25.
11. Carter P, Laurie GT, Dixon-Woods M. The social licence for research: why care.data ran into trouble. Journal of Medical Ethics. 2015;41(5):404-9.
12. Darquy S, Moutel G, Lapointe AS, D'audiffret D, Champagnat J, Guerroui S, et al. Patient/family views on data sharing in rare diseases: study in the European LeukoTreat project. European Journal of Human Genetics. 2016;24(3):338.
13. Arksey H, O'Malley L. Scoping studies: towards a methodological framework. International Journal Social Research Methodology. 2005;8(1):19-32.
14. Peters MD, Godfrey CM, Khalil H, McInerney P, Parker D, Baldini C. Guidance for conducting systematic scoping reviews. International Journal of Evidence-Based Healthcare. 2015;13:141-6.
15. Aitken M, de St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC medical ethics. 2016;17(1):73.
16. Aitken M, McAteer G, Davidson S, Frostick C, Cunningham-Burley S. Public preferences regarding data Linkage for health research: discrete choice experiment. International Journal of Population Data Science. 2018;3(1).
17. Papoutsi C, Reed JE, Marston C, Lewis R, Majeed A, Bell D. Patient and public views about the security and privacy of Electronic Health Records (EHRs) in the UK: results from a mixed methods study. BMC Medical Informatics. 2015;15(1):86.

18. Paprica PA, de Melo MN, Schull MJ. Social licence and the general public's attitudes toward research based on linked administrative health data: a qualitative study. Canadian Medical Association Journal Open. 2019;7(1):E40.

19. Data futures partnership. A path to social licence: guidelines for trusted data use. New Zealand 2017.

20. Jamal L, Sapp JC, Lewis K, Yanes T, Facio FM, Biesecker LG, et al. Research participants' attitudes towards the confidentiality of genomic sequence information. European Journal of Human Genetics. 2014;22(8):964-8.

21. Mählmann L, Schee Gen Halfmann S, Von Wyl A, Brand A. Attitudes towards personal genomics and sharing of genetic data among older Swiss adults: a qualitative study. Public Health Genomics. 2018;20(5):293-306.

22. Aitken M, Porteous C, Creamer E, Cunningham-Burley S. Who benefits and how? Public expectations of public benefits from data-intensive health research. Big Data & Society. 2018;5(2):1-12.

23. YouGov for Open Data Institute. Knowledge & Opinion [Internet]. (ODI) ODI, editor. UK 2018. Available from: https://theodi.org/article/odi-survey-reveals-british-consumer-attitudes-to-sharing-personal-data/.

24. Robinson G, Dolk H. Public attitudes to data sharing in Northern Ireland. Ulster: Administrative Data Research Centre Northern Ireland (ARK); 2015.

25. Ipsos MORI, For the Wellcome Trust. The one-way mirror: public attitudes to commercial access to health data. London: Wellcome Trust; 2016.

26. Pickard KT, Swan M, editors. Big desire to share big health data: a shift in consumer attitudes toward personal health information. AAAI; 2014; Quebec City.

27. Goodman D, Johnson CO, Bowen D, Smith M, Wenzel L, Edwards K. De-identified genomic data sharing: the research participant perspective. Journal of community genetics. 2017;8(3):173-81.

28. Kim TK, Choi M. Older adults' willingness to share their personal and health information when adopting healthcare technology and services. International Journal of Medical Informatics. 2019;126:86-94.

29. Zalin A, Papoutsi C, Shotliff K, Majeed A, Marston C, Reed J. The use of information for diabetes research and care: patient views in West London. Practical Diabetes. 2016;33(3):81-6a.

30. Aitken M, Cunningham-Burley S, Pagliari C. Moving from trust to trustworthiness: experiences of public engagement in the Scottish health informatics programme. Science & Public Policy. 2016;43(5):713-23.

31. NICE Citizens Council. NICE citizens council reports. What ethical and practical issues need to be considered in the use of anonymised information derived from personal care records as part of the evaluation of treatments and delivery of care? London: National Institute for Health and Care Excellence (NICE); 2015.

32. Spencer K, Sanders C, Whitley EA, Lund D, Kaye J, Dixon WG. Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: a qualitative study. Journal of Medical Internet Research. 2016;18(4):e66.

33. Mazor KM, Richards A, Gallagher M, Arterburn DE, Raebel MA, Nowell WB, et al. Stakeholders' views on data sharing in multicenter studies. Future Virology. 2017;12(9):537-47.

34. Cheah PY, Jatupornpimol N, Hanboonkunupakarn B, Khirikoekkong N, Jittamala P, Pukrittayakamee S, et al. Challenges arising when seeking broad consent for health research data sharing: a qualitative study of perspectives in Thailand. BMC medical ethics. 2018;19(1):86.

35. McCormick N, Hamilton C, Koehn CL, English K, Stordy A, Li L. Canadians' views about using big data in health research from a national online survey: a partnership of patient-consumers and researchers. Arthritis & Rheumatology. 2018;70.

36. Boeldt DL, Wineinger NE, Waalen J, Gollamudi S, Grossberg A, Steinhubl SR, et al. How consumers and physicians view new medical technology: comparative survey. Journal of Medical Internet Research. 2015;17(9):e215.

37. Tully MP, Bozentko K, Clement S, Hunn A, Hassan L, Norris R, et al. Investigating the extent to which patients should control access to patient records for research: a deliberative process using citizens' juries. Journal of Medical Internet Research. 2018;20(3).

38. McCormack P, Kole A, Gainotti S, Mascalzoni DA, Molster C, Lochmüller H, et al. 'You should at least ask'. The expectations, hopes and fears of rare disease patients on large-scale data and biomaterial sharing for genomics research. European Journal of Human Genetics. 2016;24(10):1403-8.

39. Grande D, Mitra N, Shah A, Wan F, Asch DA. The importance of purpose: moving beyond consent in the societal use of personal health information. Annals of Internal Medicine. 2014;161(12):855-62.

40. Platt J, Bollinger J, Dvoskin R, Kardia SL, Kaufman D. Public preferences regarding informed consent models for participation in population-based genomic research. Genetics in Medicine. 2014;16(1):11.

41. Sterckx S, Rakic V, Cockbain J, Borry P. "You hoped we would sleep walk into accepting the collection of our data": controversies surrounding the UK care. data scheme and their wider relevance for biomedical research. Medicine Health Care & Philosophy. 2016;19(2):177-90.

42. West B, Cumella A. Data sharing and technology: exploring the attitudes of people with asthma. Report London: Asthma UK; 2018.

43. Mursaleen LR, Stamford JA, Jones DA, Windle R, Isaacs T. Attitudes towards data collection, ownership and sharing among patients with parkinson's disease. Journal of Parkinsons Disease. 2017;7(3):523-31.

44. Satinsky EN, Driessens C, Crepaz-Keay D, Kousoulis A. Mental health service users' perceptions of data sharing and data protection: a qualitative report. Journal of Innovation in Health Informatics. 2018;25(4):239-42.

45. Riessman CK. Narrative analysis: qualitative research methods. US: SAGE; 1993 2nd September 1993. 88 p.

46. Understanding Patient Data. How do people feel about the use of data? UK; 2019.

47. Biddle N, Edwards, B, Gray, M & McEachern S. Public attitudes towards data governance in Australia. Canberra: Australian National University Centre for Social Research & Methods; 2019.

48. Lupton D. 'I'd like to think you could trust the government, but I don't really think we can': Australian women's attitudes to and experiences of My Health Record. Digital Health. 2019;5.

49. Tully MP, Hassan L, Oswald M, Ainsworth J. Commercial use of health data—A public "trial" by citizens' jury. Learning Health System. 2019;3(4):e10200.

50. Huskison T, Gilby N, Evans H, Stevens J, Tipping S, MORI I. Wellcome trust monitor wave 3: Chapter 6. Participation in medical research. London UK: Wellcome Trust; 2016.

51. Moss L, Shaw M, Piper I, Hawthorne C, Kinsella J. Sharing of big data in healthcare: public opinion, trust and privacy considerations for health informatics researchers. 10th International Conference in Health Informatics; Porto, Portugal, 2017.

52. Platt JE, Jacobson PD, Kardia SLR. Public trust in health information sharing: a measure of system trust. Health Serv Res. 2018;53(2):824-45.

Appendix

Appendix 1: Logic grid – Research Question One

The logic grid below was developed to explore the first research question:

What are community attitudes towards the use of government health data by private sector organisations for therapeutic development?

Logic grid for key terms describing population, concept, context and outcomes with respect to research question 1

Population	Concept or phenomena of interest	Context	Outcome/ themes
Community (communit* OR patient* OR public OR citizen* OR client* OR consumer*)	Big data (“health data” OR “health information” OR “Big data” OR “information sharing” OR “Data mining” OR “Data analytics” OR “Data linkage” OR “data sharing” OR “electronic health record” OR “electronic health data” OR “electronic medical record” OR “electronic medical data” OR “electronic patient record”)	Therapeutic development (medicine*OR “health technolog*” OR device* OR therapeutic*)	Attitudes, views or perspectives (attitude* OR perspective* OR view* OR opinion*)
Community (communit* OR patient* OR public OR citizen* OR client* OR consumer*)	Big data (“health data” OR “health information” OR “Big data” OR “information sharing” OR “Data mining” OR “Data analytics” OR “Data linkage” OR “data sharing” OR “electronic health record” OR “electronic health data” OR “electronic medical record” OR “electronic medical data” OR “electronic patient record”)	Private sector (“private sector” OR industry OR commercial)	Attitudes, views or perspectives (attitude* OR perspective* OR view* OR opinion*)

Appendix 2: Logic grid - Research Question Two

The logic grid below was developed to explore the second research question:

What is the public interest and social licence for the use of government health data by private sector organisations for therapeutic development?

Logic grid for terms describing exposure/context and outcomes/themes with respect to research question 2	
Concept/ phenomena of interest	Outcome/themes
Big data (“health data” OR “health information” OR “Big data” OR “Data mining” OR “Data analytics” OR “Data linkage” OR “data sharing” OR “information sharing” OR “electronic health record” OR “electronic health data” OR “electronic medical record” OR “electronic medical data” OR “electronic patient record”)	Social licence and public interest (“social licence” OR “public interest” OR “public good” OR “public benefit” OR “social trust” OR “social value”

Appendix 3: Scoping Review Inclusion/Exclusion Criteria

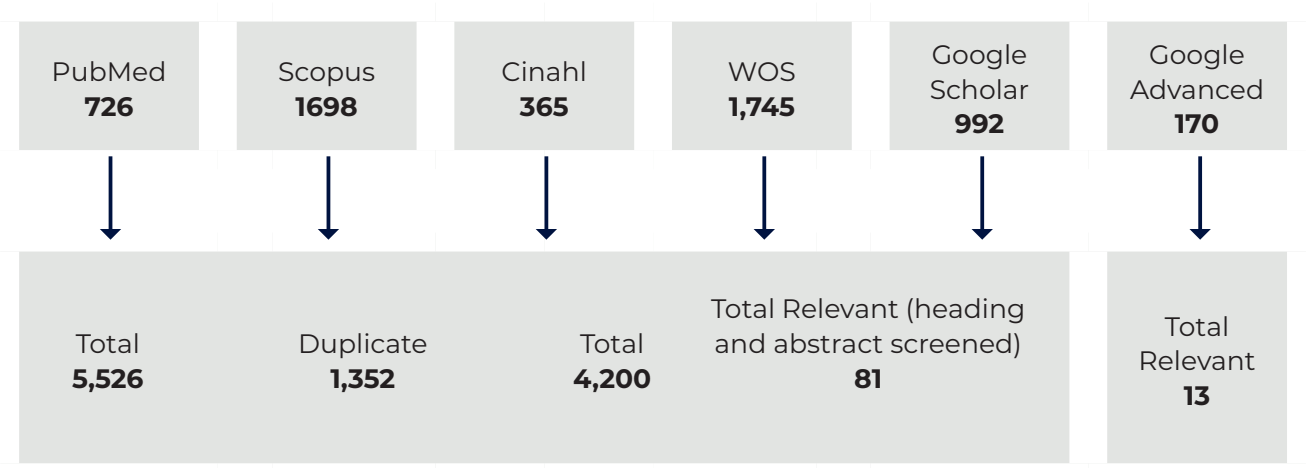
	Inclusion Criteria	Exclusion Criteria
Publication Date	March 2014 – March 2019	Before March 2014
Document Type	Journal articles, conference paper, review, book chapter, book, article in press and reports (govt. and non-govt.)	Thesis, blog, PowerPoint slide, and research agencies websites
Study Design	Empirical	Conceptual
Study Population	Community members, public, patient groups	Clinical stakeholders, private industry
Research Topic	Articles were selected because, with respect to the use of any administrative data (linked or not) by private sector organisations for therapeutic development, they describe empirical examples or understandings of: • community attitudes • public interest • social licence Private uses of public data for therapeutic development Private uses of public data outside therapeutic development in the biomedical/health sector	Outside of health: The paper is not related to data sharing in the health sector i.e. it is about data sharing in an alternate sector e.g. food sector, financial sector, environmental sector Technical ONLY: The paper ONLY describes technical methods for analysing, sharing and linking data Data sharing at individual level – e.g. sharing information between patients and doctors Marketing: The paper describes marketing or sales strategies for private industry

In addition, we set aside articles which describe:

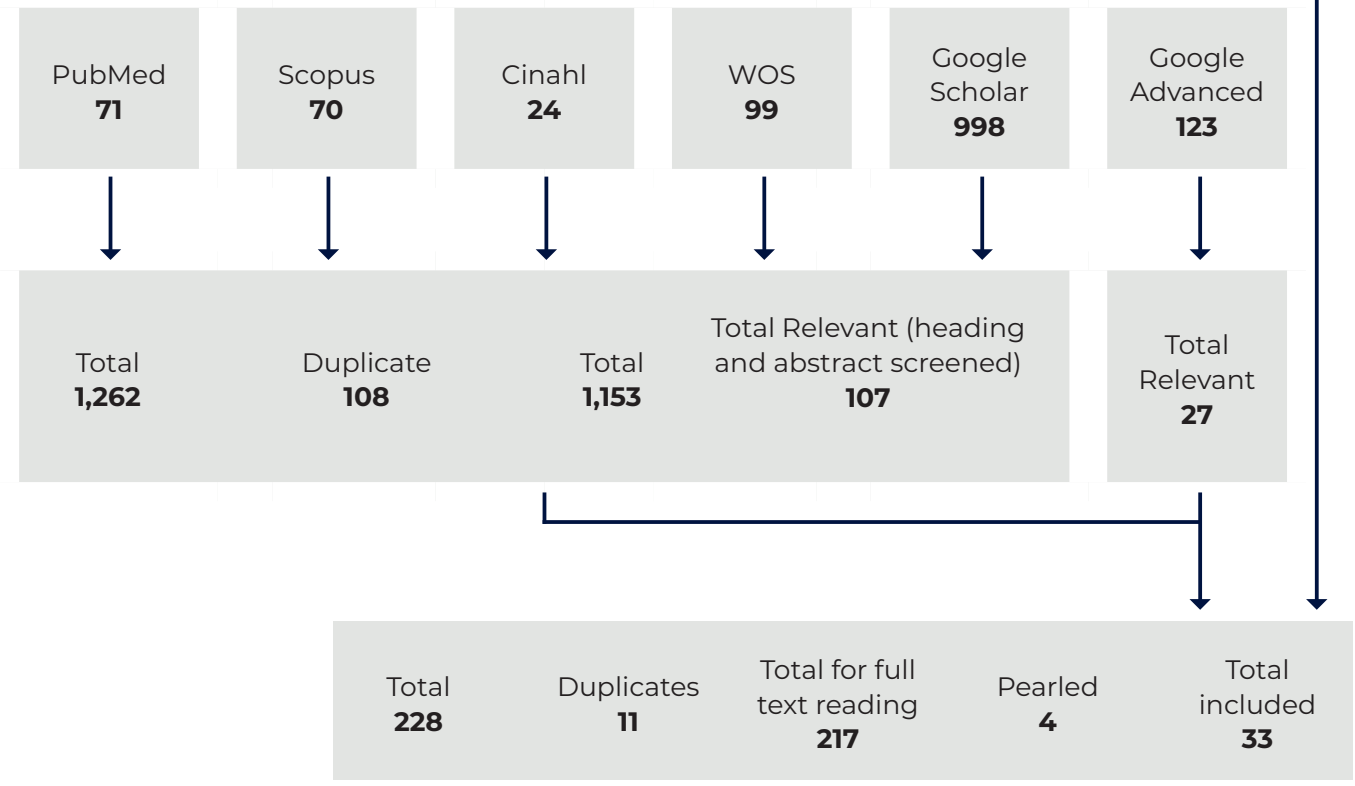
- Access models which permit acceptable ethical sharing which is in the public interest
- Reuse of data sets collected by private companies
- Social and ethical issues associated with sharing of administrative data including community attitudes to sharing of administrative data generally or in the public sector only
- Trust -discuss patient and public trust attitudes to data analytics and data linkage in health care

Appendix 4: Scoping Review Process and Findings Flowchart

Search 1: Community Attitudes



Search 2: Public Interest & Social Licence



Appendix 5: Summary of the Final Scoping Review Publications

Title	Author/s	Year (data collected)	Location/s	Publics engaged	Sample method (no.)	Stakeholder method (no.)
Moving from trust to trustworthiness: Experiences of public engagement in the Scottish health informatics programme	Aiken, Cunningham-Burley & Pagliari (30)	2016 (2010-2011)	UK	Broader public, clinical stakeholders	Focus groups (50)	Workshop (28)
Public preferences regarding data linkage for health research: a discrete choice experiment	Aitken, McAteer et al. (16)	2018 (2016)	UK	Broader public	Online questionnaire (1,004)	-
Who benefits and how? Public expectations of public benefits from data-intensive health research	Aitken, Porteous, et al. (22)	2018 (2017)	UK	Broader public	Deliberative workshops (69)	-
How consumers and physicians view new medical technology: Comparative study	Boeldt et al. (36)	2015 (2014)	US	Broader public, clinical stakeholders	Survey (1,102)	Survey (1406)
Challenges arising when seeking broad consent for health research data sharing: a qualitative study of perspectives in Thailand	Cheah et al. (34)	2018 (2017)	Thailand	Broader public, clinical stakeholders, affected patient group	Semi structured interviews and focus groups: 19	-
Patient views on data sharing in rare diseases: Study in European LeukoTreat project	Darquy et al. (12)	2016 (2012)	Europe	Affected patient group and their families	Online questionnaire: 195	-

A Path to Social Licence	Data Futures Partnership (19)	2017 (n/a)	NZ	Broader public, clinical stakeholders, private sector agencies	Workshops (379), Online survey (4,033), hui (94), online hui (60), online survey (533)	-
De-identified genomic data sharing: The research participant perspective	Goodman et al. (27)	2017 (2013)	US	Broader public, affected patient group and their families	Online survey (450)	-
The importance of purpose: moving beyond consent in the societal use of personal health data	Grande et al. (39)	2014 (2012)	US	Broader public	Online survey (3,064)	-
Wellcome Trust Monitor Report Wave 3: Tracking public views on science and biomedial research	Huskison et al. (50)	2016 (2016)	UK	Broader public	Online survey (1,524)	-
The One-Way Mirror - Public attitudes to commercial access to health data	Ipsos Mori (25)	2016 (2015)	UK	Broader public, clinical stakeholders, affected patient group	Deliberative workshops and workshops (246- include stakeholders), face to face survey (2,017)	Deliberative workshops and workshops (246 – include sample population)
Research participants' attitudes towards the confidentiality of genomic sequence information	Jamal et al. (20)	2014 (2011-2012)	US	Affected patient groups	Semi structured phone interview (30)	-

Oder adults willingness to share their personal and health information when adopting healthcare technology and services	Kim & Choi (28)	2019(2017)	South Korea	Broader public	Face to face survey (170)	-
Attitudes towards personal genomics and sharing of genetic data among older Swiss adults: A qualitative study	Mahlmann et al. (21)	2018 (2013-2014)	Switzerland	Broader public	Semi structured interviews (40)	-
Stakeholder views on data sharing in multicentre studies	Mazor et al. (33)	2017 (2015)	US	Broader public, clinical stakeholders, affected patient group	Semi-structured interviews (15)	Semi-structured interviews (19)
You should at least ask'. The expectations, hopes and fears of rare disease patients on large-scale data and biomaterial sharing for genomics research	McCormack et al.(38)	2016	International	Affected patient group	Focus groups (52)	-
Canadians' views about using big data in health research from a national online survey: A partnership of patient consumers and researchers	McCormick et al. (35)	2019 (2017)	Canada	Broader public, patient groups	Online survey (151)	
Sharing of Big Data in Healthcare: Public opinion, trust and privacy considerations for health informatics researchers	Moss et al. (51)	2017 (n/a)	UK	Broader public	Face to face survey (37)	-
Attitudes towards data collection, ownership and sharing among patients with Parkinson's disease	Mursaleen et al. (43)	2017 (2016)	UK	Affected patient group	Online survey (310)	-

What ethical and practical issues need to be considered in the use of anonymised information derived from personal care records as part of the evaluation of treatments an delivery of care	National Institute for Health and Care Excellence (31)	2015 (2015)	UK	Broader public (citizen council)	Deliberative meeting (30)	-
Patient and public views about the security and privacy of Electronic Health records in the UK: results from mixed methods study	Papoutsi et al. (17)	2015 (2011-2013)	UK	Affected patient group, clinical stakeholders	Survey (2,761), focus groups (114)	Focus group (6)
Social licence and the general public's attitudes toward research based on linked administrative health data: a qualitative study	Paprica et al. (18)	2019 (2015 & 2017)	Canada	Broader public	Focus group (65)	-
Big desire to share big health data: A shift in consumer attitudes towards personal health information	Pickard & Swan (26)	2014	International	Broader public	Online survey (128)	-
Public preferences regarding informed consent models for participation in population-based genomic research	Platt et al. (40)	2014 (2007-2008)	US	Broader public	Survey: 3,347	-
Public trust in health information sharing: A measure of system trust	Platt et al. (52)	2018 (2014)	US	Broader public	Online survey (1,011)	-
Public attitudes to data sharing in northern Ireland	Robinson & Dolk (24)	2015 (2015)	UK	Broader public	Survey (1,202)	-

Mental health service users' perceptions of data sharing and data protection: a short qualitative report	Satinsky et al. (44)	2018 (2017)	UK	Broader public	Focus groups (8)	-
Patient perspectives on sharing anonymized personal health data using a digital system for dynamic consent and research feedback: a qualitative study	Spencer, Sanders & Dixon (32)	2016	UK	Affected patient group	In-depth interviews and focus groups (40)	-
“You hoped we would sleep walk into accepting the collection of our data”: Controversies surrounding the UK care.data scheme and their wider relevance for biomedical research	Sterckx et al. (41)	2016 (2013-2015)	UK	Broader public	Qualitative analysis of blogs (256)	-
Investigating the extent to which patients should control access to patient records for research: A deliberative process using citizen juries	Tully et al. (37)	2018	UK	Broader public	Citizen juries (51)	-
Data Sharing and Technology: Exploring the attitudes of people with asthma	West & Cumella (42)	2018 (2018)	UK	Affected patient group	Online survey (3,054)	-
ODI Survey reveals British consumer attitudes to sharing personal data	YouGov (23)	2018 (2017)	UK	Broader public	Online survey (2,023)	-
The use of information for diabetes research and care: Patient views in West London	Zalin et al. (29)	2015 (2011-2012)	UK	Broader public, affected patient group	Survey (404), focus group (6)	-

Appendix 6: Summary of concerns about data sharing with sample quotes

Concern	Sample quote from a participant or from text in the article
Data security, data leaks and hacking	Best practice guidelines are fine – but people are just fallible. One person, without malice, discloses 1,000 people's sensitive data through an insecure email. Mistakes just happen, but once it's out there, you can't get it back. General Public, workshops, New Zealand. (19, p.13)
Loss of privacy lack of respect for confidentiality and privacy	'Drunk driving's still an offence even if you don't have an accident, because society thinks the potential of harm is sufficient. Are you seriously saying that medical confidentiality only matters in retrospect if the failure causes explicit harm, rather the potential for harm?' Letter to The Guardian, UK. 19.08.14. (31, p.182)
Data exposure causing stigmatisation	A small minority of participants described concern about risk to their privacy, speculating that patients with more sensitive health conditions may be "more guarded of what happens with their health information" (Participant #11) due to fear of stigmatization. Affected patient groups, UK (12, p.5)
Surveillance	Some participants felt that we are heading for a dystopian, surveillance-based society. For them, whether the data is anonymised or not was almost not relevant because they were worried about the future risks to privacy rights as a result of so much increased data sharing of any kind. General public, UK. (25, p.11)
Use of data to discriminate in insurance	'They do penalise people with illnesses. [Like] car insurance. My daughter's a prime example - she's got MS but she has to have her driving licence renewed every three years and she pays a higher premium. She's got MS but she is just as capable of driving as anyone'. – Participant, general public, Swansea, UK. (5, p.32)
Use of data to discriminate by an employer	'Yaa, the people who suffer from a disease, they want... they fear that it will become public, I think it is connected to their job, for example if they have an increased risk for diabetes, they might not get the job, since the employer might be worried that they will get sick, I guess that plays a role'. Participant, female, age 71. (21, p.8)

Use of data for marketing	[Clinical trial participants] expressed concern about their identifiable information being used by third parties, such as telemarketers and insurance agencies, for purposes unrelated to health research. Affected patient groups, Focus groups, Thailand. (34, p.5)
Suppression of data for political ends	M2. It would depend where the university was getting the funding from. There's a chance the university could be getting their funding from a pharmaceutical company. F2: Again, that's trust, isn't it? You're trusting those researchers to be ethical with their findings...., Focus Group 3, Glasgow, Scotland (23, p.5)
Use of data to promote the interests of academic researchers	M2. 'It would depend where the university was getting the funding from. There's a chance the university could be getting their funding from a pharmaceutical company. F2: Again, that's trust, isn't it? You're trusting those researchers to be ethical with their findings....'. Participant, focusgGroup 3, Glasgow, Scotland. (22, p.5)
Use of data for purposes which are not in the public interest or for public benefit	'If there is such a partnership, I refuse to participate in the database. The pharma industry orients research in their own interests, not in the interests of patients'. Rare disease patient/ family member, survey, Europe. (12, p.8)
Use of data for eugenics	'So, my only concern is, it has once been talked about, that it could be used to create the perfect human... or... that everyone would have blue eyes or a standard type or for military purposes. Of course, that is a big topic. I would be absolutely against that'. - Participant, female, age 69, interviews, older adults, Switzerland. (21, p.9)
Use of data to deny treatment	'And then they combine all that together, and they say, okay, well, this person has got this and this and this. Wasting medication or treatment or whatever on this person, beyond this age is useless. Let's just let this person die'. - Participant, focus group 1, Thunder Bay, Canada. (18, p.E43)
Use of data to generate profit	Pharmaceutical companies hold a lot of power and the potential for life or death, and make huge profits out of life or death situations, one person suggested. Another said, We feel as a society we are at the mercy of pharmaceutical companies because they usually put profit before people. - General public, Citizens' Council, UK. (31, p.25)

Lack of transparency in how data is used	‘A deliberate decision was made to sell our data to a private company for purely commercial use, having nothing whatsoever to do with improving medical care or NHS services. This government has lied about its intentions for the NHS, and continues to lie and obstruct information about what it has done with it, and what is going on as we speak. This is all really too bad, because the existence of a large database like the NHS has such obvious benefits. However this is only where the NHS remains a public service, and where the information is used to benefit the public. NOT commercial businesses’. - Letter to The Guardian, UK, 18.08.14. (41, p.184)
Re-identification of anonymised data	‘I am concerned that even anonymised information could be combined with other information that’s easily available to de-anonymise and identify me. I’m also concerned that other moves that are planned for the future will further erode patient confidentiality beyond what has already been published’. - Comment posted to UK Care.data website. (41, p.182)
Concern that data may be sold on to others	‘In particular, it seemed to go against natural justice that a company could repurpose data originally generated from the public, and make money again and again from the same dataset’. – Participant -general public, UK. (25, p.59)
Erosion of trust in health care professionals	‘Patients will not confide in their doctor, certain personal information, that may be absolutely necessary for diagnoses and treatment, knowing that outside agencies may have access to it’. - Comment posted to the UK Care.data website, 26.02.14. (41, p.183)

Appendix 7: Summary of conditions which a proportion of participants believe needs to be met before data should be shared with private companies with sample quotes

Condition (references)	Sample quote from a participant or from text in the article
Purpose	
That there be clear public benefit (16, 17, 19, 21, 22, 24, 25, 30, 31, 33, 34, 39)	The involvement of a commercial organisation was seen as fairly easy to accept when participants could see clear potential for patients, society and future generations to benefit. This meant Linking data in the NHS and monitoring the safety of drugs were acceptable and seen as valuable. — General public, UK, 2015. (25, p.51)
That the use of the data be 'ethical' (20, 24)	'While I may be willing to share all my data for the purposes of improving health for the world at large, I find the language used vague -most probably- on purpose. There is no guarantee that my data will be used ethically, All it says is that there are "strict rules to protect" privacy. I want strict rules to protect my data from being used in research relating to the creation, marketing or deployment of weapons; I want my data to be protected from being sold to or shared with companies which engage in the patenting of genome products; I want my data to be protected from being sold to or shared with companies that engage in abusive hiring practices here and abroad; I want my data to be protected from being sold or shared with companies or individuals that treat the environment with contempt; I want my data to be protected from being sold to or shared with companies that have unacceptable top executive salaries; and this is just a sample. I want the data to be supervised by an independent forum of individuals whose remit is to follow strict published ethical guidelines relating to sharing, selling and profit making by the use of my data'. — Comment posted to UK Care.data website 22.01.14. (41, p.81)
That sharing of data not undermine equitable distribution of resources (22)	This related to wider discussions around the ways in which benefits of health research are realised and a widely held perception that currently the benefits are not realised equitably across society and that different groups or people in different locations across the UK experience health services differently as well as experiencing different health outcomes. Throughout these discussions a recurring theme was that the potential benefits of health research were not always or consistently realised. A range of factors were noted as limiting the realisation of public benefits from health research, these included commercial interests, political priorities and limited public funding. — General public, workshops, Scotland, 2018. (22, p.6)

Avoiding Harm	
That sharing the data causes no harm to individuals or society (25)	'Red lines: anything that risks personal harm, especially to vulnerable individuals'. - Participants describing circumstances under which data sharing should not occur, general public, survey, UK, 2016. (25, p.12)
Not for Profit	
That public companies not overly profit from the data sharing at the expense of the public (12, 16-19, 21, 22, 31, 33, 50)	The majority of respondents (62 per cent) chose: "Any profit made from research carried out using linked information should be invested into public services". Only 8 per cent chose "Any profit made from research carried out using linked information should be kept by those carrying out the research". — General public, Scotland. (16, p.11) 'I guess I just think maybe they [the private sector] could fund their own research. I'm not sure the taxpayers should pay for it. But I guess, as you said, if they're giving us an appropriate price or a better drug being released then I guess it's okay'. - Participant, focus groups, Toronto, Canada, 2017. (18, E43)
That private companies pay for access to data (24, 31)	'I would like them to get access to patient records but they should pay a fee to get the data as they will make a profit out of any new drug'. Participant, General public, Survey, Northern Ireland. (24, p.20)
Consent	
That consent be obtained before data is shared (12, 18, 19, 24, 25, 31, 32, 34, 38-42)	Council members suggested that systems should be established to ensure that researchers are transparent from the outset, with informed consent procedures requiring them to tell study participants how their data will be used and who might have access to it (including whether their data may be sold to other organisations), how long data will be kept and what will happen to data once the research study has finished. Explanations should be kept simple to ensure study participants can understand and they should be given written information. Informed consent procedures should include information about what a researcher plans to do with data if they discover incidental findings about a participant's health and wellbeing. — General public, Citizens Council, UK, 2015. (31, p.27)
That participants have the option to opt out from data sharing (19, 21, 31)	'... it has to be voluntary, because we have to respect different personal inhibition thresholds in terms of sharing data... you cannot force someone or share data behind someone's back. Out of respect'. - Participant, male, age 74, interviews Swiss older adults. (21, p.9) Overall, several Citizens Council members considered it was important to allow individuals privacy and the right to choose to share their data because they considered this was important to society as a whole. They felt society was stronger if people had freedom of choice. - General public, Citizens Council, UK, 2015. (9, p.39)

That participants be well informed in the consent process (19, 20, 31, 33, 34, 38)	'We have to take time to explain until they (the clinical trial participants) understand. The bulls will not pull the cart and take offright after you connect the cart to them. As for us, we have been committee members for many years so once you explain, we understand immediately. New people will not understand it...Carefully explain one by one step by step.' - CAB2, Community ethics advisory board member, male, focus groups, Mae Sot on the border of Thailand and Myanmar, Thailand, 2018. (34, p.6)
That consent procedures indicate what a researcher will do if they discover incidental findings relevant to a participant's health and wellbeing. (31)	'Informed consent procedures should include information about what a researcher plans to do with data if they discover incidental findings about a participant's health and wellbeing'. — Participant, general public, Citizens Council, UK, 2015. (31, p.27)
Reciprocity	
That the public be informed about how and why the data is being shared and the potential known risks of sharing the data (19-21, 25, 33)	'Guess what? Maybe we, as a generation, need to go out on a limb a little bit here, but it's that proper education of how this is going to be used, and proper thanks'. — Patients, interviews, USA. (33, p.544)
That participants in have access to their own data (29)	The majority of respondents said that they would prefer to have full access to their own medical history (91.52%) rather than limited access to their health information (4.09%). — Public and patients, survey, London, UK. (29, p.83)
That outcomes from any research involving their data be made public (12, 20, 24, 27, 32, 33)	'I just love this idea [Dynamic Consent], the updates they're great. If I was involved with something [research] and it got published, I could go on the Internet and click on that [dynamic interface] and it would give me all the published papers on it'. — Participant #4, patients, interview, UK, 2016. (32, p.7)
That participants in research have access to the developed technology if it proves useful (31)	The group considering a new treatment for a rare condition were concerned that a patient might contribute to research by sharing their data but might no longer have access to or be able to afford the drug once it was approved and not available as part of a research study. — General public, Citizens' Council, NICE, UK, 2015. (31, p.22)

That the benefit companies gain through data sharing be returned to the public in some way e.g. lower drug prices (17, 18)	'[If they] explain to me that the database is not only for medical purposes but would also get us access to more medical [services] in terms of the way the commissioning is taking place, then yes, you are making a good case to get me on the database, but if you are saying that, oh, I should just provide my [information] what's this all this research going on for?' — Participant, FG13, focus groups, patients, UK. (17, p.9)
Data Security and Confidentiality	
That the data be securely stored (18-20, 24, 25, 29, 37, 38, 41)	'Well I think because it's health data, it's really important to keep it safeguarded. It's not just some random information. It's personal information. Really personal information'. — Group 2, focus groups, general public, Toronto, Canada, 2015. (18, p.E43)
That personal privacy and confidentiality be respected and maintained (12, 18, 20, 21, 27, 29, 31)	Council members considered there was no difference in the need to protect data and keep it confidential depending on who was collecting and using data. Protection and confidentiality of personal data should be a top priority for all types of organisations, the group agreed. — General public, Citizens' Council, NICE, UK, 2015. (31, p.27)
That the organisations involved be seen to be highly trustworthy and competent (19, 37, 38)	The reasons the juries gave for why organizations should not have access to the data included several that were the opposite of the reasons for access, such as organizations that did not clearly indicate that the primary use of the data is for public benefit, who may use the data solely for private gain or commercial profit, or who did not have a trusted track record for protecting data. — General public, Citizens Jury, UK, 2015. (37, p.9)
That the data be anonymised or aggregate data (16, 18-21, 24-26, 30, 31, 33, 35, 38)	'With critical credentials removed I can see enormous benefit where data is used to forecast services for health and far beyond. I want my data to be part of that, as long as it's not recognisable as being mine'. — Participant, general public, workshops, New Zealand, 2017. (19, p.12)
That researchers be vetted (24, 25)	'Licensing and regulation --- if people are doing it for good then they should be happy for it to be regulated. Every organisation should be licensed and regulated and audited to ensure that there is a level of openness'. — Patients (rare conditions), Sheffield, UK. (25, p.62)
That researchers' activity on the system be tracked (25, 38)	Participants pointed out that the individual themselves may not be trustworthy. Many are reassured by the suggestion that every time somebody has access they should be required to enter a log. The public like that if anything goes wrong it is easy to trace the user and that in turn acts as a driver of good practice and means it is possible to monitor and enact consequences for excessive use. — Public and patients, workshops, UK, 2016. (25)

That the data analysis be undertaken in the public or university sector (24, 25)	'The pharma company should pay for it, the regulator or academics should do it'. — General Public, Sutton Coldfield, UK, 2016. (25, p.51)
Scientific Integrity	
That any research using data linkage and sharing be scientifically robust (19, 27, 31)	Good scientific practice to ensure the accuracy and validity of research design and data analysis. The Council felt that research that does not produce any useful findings because it is not scientifically robust is a waste of time and resources.— General public, Citizens' Council, NICE, UK, 2015 (9, p.43)
Accountability and Oversight	
That there be independent oversight or a strong regulatory body overseeing the research (16, 24, 25, 30, 33, 35, 37, 38)	'I found it encouraging that the information and privacy commissioner has an oversight over it and it renews every 3 years. I found that encouraging. Someone's keeping an eye on it'.— Participant, general public, group 1, Focus groups, Sudbury, Canada, 2017. (18, p.E43)
Criminal penalties (fines, prison sentences) in the event of wrongdoing (24, 25, 30)	Do you think that perhaps the reason we're not happy with many people having that level of power over our data, partly because we don't believe that the penalties for misusing data are severe enough? I mean, for me, that's a crucial point, I actually think there would be less mismanagement of data if the penalties of knowingly selling or giving away personalised information carry far greater criminal penalties [.. . Currently] They don't, I mean, you're not going to go to jail for it! Whereas, perhaps if you did people would be less likely to, you know, purposely sell personalised information. — Male1, social service researcher, Scotland, 2016. (30)
That collection of and access to data be appropriate – sufficient for purpose and no more (29, 31, 37)	Typically, these organizations clearly demonstrated that the primary goal for using the data was for public benefit (such as improved medical care and treatments, improved public health, or management of public funds) and made a clear and compelling case for why they need these patient records. They provided clear justification for how and why the data would be used, why it was relevant to their efforts, with whom it will be shared, and only access records they needed to perform their data analysis and could not get adequate data from other sources. — General public, Citizens' Juries, UK, 2018. (37, p.9)

Appendix 8: Community Attitudes Survey

Eligibility Criteria

What is your gender? One response only - drop down box

- ☐ Male
- ☐ Female
- ☐ Indeterminate/Intersex/Trans/Gender diverse
- ☐ I prefer not to respond

How old are you? One response only – drop down box

- ☐ Under 18
- ☐ 18-24
- ☐ 25-29
- ☐ 30-34
- ☐ 35-39
- ☐ 40-44
- ☐ 45-49
- ☐ 50-54
- ☐ 55-59
- ☐ 60-64
- ☐ 65-69
- ☐ 70-74
- ☐ 75+

Where do you currently live? One response only – drop down box

- ☐ Greater Sydney
- ☐ Rest of NSW
- ☐ Greater Melbourne
- ☐ Rest of VIC
- ☐ Greater Brisbane
- ☐ Rest of QLD
- ☐ Adelaide
- ☐ Rest of SA
- ☐ Perth
- ☐ Rest of WA
- ☐ TAS
- ☐ Northern Territory
- ☐ ACT

Linking data for development of new drugs and devices

Every day, Australians generate large amounts of information about themselves that is recorded in computers. This can include information about visits to doctors, medicines we take, hospital visits and blood tests.

Bringing together and linking these different pieces of information from lots of people provides statistics that can help us improve the quality of healthcare for all Australians. These statistics can help in the development of new treatments and make sure that the treatments we have are working and are safe. For example, by linking medical prescription and hospital emergency department statistics, researchers can discover unrecognised harms of new medicines.

At the moment health information is rarely shared and when it is, it is usually shared only between government organisations.

Although there are benefits to linking and using this information, some people are concerned about the possibility that their health information may be given to people who shouldn't have it, or that these people may be able to work out who we are. People are also worried about private companies misusing health information.

For this reason, it is standard practice before sharing and linking health information to exclude names, addresses, dates of birth and Medicare numbers. Despite this precaution, there have been a small number of cases where re-identification has occurred.

We would like to know what you think about sharing this information with private companies such as drug companies and medical device manufacturers where the goal is to support the development of new treatments for diseases and disabilities.

Please read the Participant Information Sheet below:

(Participant information sheet will appear on this page within the survey)

☐ I have read and understood the Participant Information Sheet

If you would like to view and/or download and/or print the Participant Information Sheet please go here. (This link will send participant to a PDF version of Participant Information Sheet)

The questions below are about your government health information which has personal information removed, e.g. no name, no address, no date of birth, no Medicare number.

1. To what extent do you agree with the government sharing your health information with private companies, such as drug companies or medical device manufacturers?

Matrix – one answer per row – evenly spaced shaded rows

The government can share my health information with private companies:	I strongly disagree with this			I neither agree nor disagree with this			I strongly agree with this
(a) To improve health services	1	2	3	4	5	6	7
(b) For research in universities, hospitals or publicly funded research organisations	1	2	3	4	5	6	7
(c) So the companies can develop new medical devices (e.g. pacemakers, cataract surgery)	1	2	3	4	5	6	7

2. What do you think about your health information being using by private companies for the development of new medicines or devices?

One answer only. Question logic: 2b go to question 3. 2a, 2c, 2d, 2e or 2f go to question 4

- ☐ (a) My health information should not be used at all
- ☐ (b) I need to say 'yes' for my data to be used (opt in)
- ☐ (c) I need to say 'no' if I don't want my data to be used (opt out)
- ☐ (d) I do not need to know, just use the information
- ☐ (e) I am not sure/I do not know
- ☐ (f) I do not understand this question

3. Would you like to be asked for your consent:

One answer only – drop down box.

- ☐ (a) Every time
- ☐ (b) Just once
- ☐ (c) Get your general consent and be re-contacted from time-to-time

4. Imagine that the Government has decided to share your health information with a private company. The company intends to use the information to help develop a new treatment for a disease. How important is it that each of the following conditions be met before the information is shared?

One answer per row

How important is each of the following:	Not important at all			Neither important nor unimportant			Very important
(a) I am told how my health information will be used	1	2	3	4	5	6	7
(b) I am told which company will have access to my health information	1	2	3	4	5	6	7
(c) My health information is stored in a safe place	1	2	3	4	5	6	7
(d) The private company pays for the use of the health information	1	2	3	4	5	6	7
(e) The information sharing is approved by an independent ethics committee	1	2	3	4	5	6	7
(f) The private company is required to publish all results – both good and bad	1	2	3	4	5	6	7
(g) The research is likely to lead to benefits for society	1	2	3	4	5	6	7
(h) There are strict rules to stop the information being passed on to anyone else	1	2	3	4	5	6	7
(i) There are criminal penalties or heavy fines if companies break the rules	1	2	3	4	5	6	7

5. To what extent do you agree with the following statements about private companies using government health information to support development of new treatments?

One answer per row

	Strongly disagree			Neither agree nor disagree			Strongly agree
(a) Private companies can be trusted to store health information safely	1	2	3	4	5	6	7
(b) Private companies should be allowed to make a profit from the use of this information	1	2	3	4	5	6	7
(c) Private companies can be trusted to act for the good of society	1	2	3	4	5	6	7
(d) If you give health information to a private company, you cannot control where it ends up	1	2	3	4	5	6	7
(e) Someone may be able to work out who I am even though my personal information has been removed	1	2	3	4	5	6	7
(f) The government won't be able to stop private companies from misusing this information, even if they try.	1	2	3	4	5	6	7

6. Is there anything else you would like to tell us about your views on sharing government health information with private companies where the goal is to support the development of new treatments for diseases and disabilities?

Demographics

7. In general, how would you rate your health?

Tick one box only

- ☐ My health is poor
- ☐ My health is fair
- ☐ My health is good
- ☐ My health is very good
- ☐ My health is excellent

8. About your health status:

Tick as many as apply. One answer per row.

	Yes	No	I am not sure/I do not know
(a) I have a chronic health condition			
(b) I care for someone with a chronic health condition			
(c) I take prescribed medication(s)			
(d) I have a My Health Record electronic health record			

9. Which best describes the highest educational qualification you have obtained?

One response only - drop down box

- ☐ No formal qualifications
- ☐ Year 10 or school certificate
- ☐ Year 12 or leaving certificate
- ☐ Trade/apprenticeship
- ☐ Other TAFE/Certificate
- ☐ University degree/Higher degree
- ☐ I prefer not to answer/I am not sure

10. What best describes your current employment status?

Tick one box only – drop down box

- [ ] Full time employed
- [ ] Part-time employed
- [ ] Unemployed
- [ ] Home duties
- [ ] Student/Training
- [ ] Retired
- [ ] Unable to work (e.g. disability/Work Cover)
- [ ] I prefer not to answer/I am not sure

11. Have you worked or do you currently work in the health industry and/or in health services or research?

One response only – drop down box

- [ ] Yes
- [ ] No
- [ ] I am not sure
- [ ] I prefer not to answer

If you would like to view and/or download and/or print the Participant Information Sheet please go here. (This link will send participant to a PDF version of Participant Information Sheet).

Appendix 9: Table of counts of the sample characteristics used for calculating population weights

			Age				
Gender	**State**		**18-29**	**30-44**	**45-59**	**60+**	**Total**
Male	Greater Sydney		63	77	66	60	266
	Rest of NSW		26	31	35	45	137
	Greater Melbourne		60	70	58	56	244
	Rest of VIC		13	17	19	25	74
	Greater Brisbane		29	33	32	28	122
	Rest of QLD		25	30	32	37	124
	SA		20	22	25	25	92
	WA		29	39	33	31	132
	Other territories and states		13	15	15	14	57
		Total	**278**	**334**	**315**	**321**	**1248**
Female	Greater Sydney		61	77	64	66	268
	Rest of NSW		25	32	37	50	144
	Greater Melbourne		61	73	60	63	257
	Rest of VIC		13	17	20	28	78
	Greater Brisbane		29	36	31	30	126
	Rest of QLD		25	33	34	38	130
	SA		19	23	23	28	93
	WA		30	38	33	33	134
	Other territories and states		11	16	15	17	59
		Total	**274**	**345**	**317**	**353**	**1289**

Appendix 10: Summary of demographics of online survey participants

Gender	N	%	Adj. %
Male	1,243	48.9	49.2
Female	1,285	50.7	50.5
Indeterminate/ Intersex/Trans/ Gen	9	0.4	0.3

Age	N	%	Adj. %
18-24	234	9.2	16.5
25-29	318	12.5	22.1
30-34	237	9.3	7.2
35-39	253	10.0	7.7
40-44	189	7.5	5.8
45-49	194	7.7	6.0
50-54	196	7.7	6.1
55-59	242	9.5	7.4
60-64	214	8.4	6.7
65-69	195	7.7	6.2
70-74	146	5.8	4.6
75+	119	4.7	3.7

Self-rated health	N	%	Adj. %
Poor	129	5.1	4.5
Fair	629	24.8	23.4
Good	991	39.1	38.3
Very good	606	23.9	25.7
Excellent	182	7.2	8.2

Chronic Health Condition (CHC)	N	%	Adj. %
Yes	640	25.2	22.7
No	1,749	68.9	71.1
I am not sure	148	5.8	6.2

Care for someone with CHC	N	%	Adj. %
Yes	323	12.7	12.2
No	2,155	84.9	85.1
I am not sure	59	2.3	2.7

Taking prescribed medication	N	%	Adj. %
Yes	1,274	50.2	46.2
No	1,230	48.5	52.3
I am not sure	33	1.3	1.5

Place of living	N	%	Adj. %
Greater Sydney	534	21.1	20.1
Rest of NSW	281	11.1	11.3
Greater Melbourne	501	19.8	19.2
Rest of VIC	152	6.0	6.1
Greater Brisbane	248	9.8	9.7
Rest of QLD	254	10.0	10.4
Adelaide	145	5.7	5.6
Rest of SA	40	1.6	1.5
Perth	213	8.4	8.6
Rest of WA	53	2.1	2.0
TAS	54	2.1	2.2
NT	21	0.8	0.9
ACT	41	1.6	1.8

Employment status	N	%	Adj. %
Full Time	943	37.2	37.7
Part Time	538	21.2	21.9
Unemployed	120	4.7	5.5
Home duties	250	9.9	9.5
Student / Training	112	4.4	6.9
Retired	456	18.0	14.4
Unable to work	107	4.2	3.5
Prefer not to answer/Not sure	11	0.4	0.5

My Health record	N	%	Adj. %
Yes	1,039	41.0	39.7
No	913	36.0	37.4
I am not sure	585	23.0	22.9

Education	N	%	Adj. %
No formal qualifications	45	1.8	1.4
Year 10 or school certificate	265	10.5	9.5
Year 12 or leaving certificate	422	16.6	18.3
Trade / apprenticeship	166	6.5	6.0
Other TAFE / Certificate	674	26.6	25.4
University deg/ Higher deg	953	37.6	38.8
Prefer not to answer/Not sure	12	0.5	0.6

Worked in health industry	N	%	Adj. %
Yes	332	13.1	13.3
No	2,173	85.7	85.3
I am not sure	20	0.8	0.8
I prefer not to answer	12	0.5	0.5

Appendix 11: PHRN online survey question 6 analysis

Participants were asked: Is there anything else you would like to tell us about your views on sharing government health information with private companies where the goal is to support the development of new treatments for diseases and disabilities? Over half of survey participants (54%) did not have anything further to add. From the participants who did respond a number of themes emerged and are grouped into the following categories; concerns regarding sharing government health information, conditions that need to be met before sharing government health information, a willingness to share government health information unconditionally and a negative response to sharing.

Concerns about sharing government health data with private companies

When asked about the sharing of government health data with private companies, participants conveyed a range of concerns. The most significant were, lack of trust (in both private companies and the government), profit generation by private companies and the security of the health information once it has been shared. A summary of these concerns is discussed below.

Lack of trust in private companies

Participants expressed a general distrust of private companies and many participants gave self-interest, corruption and profiteering as the reason, for example:

> *Private companies are always in it for their bottom line. The entire concept is to make money. That doesn't lend to it always being beneficial for society*

Participants were specifically sceptical of the motives driving large pharmaceutical companies, poor corporate ethics and the lengths companies may go to in order to make a profit. Concern was also raised that private companies would be incapable of keeping their data secure from breaches or misuse.

Lack of trust in government

In reference to 'government', respondents did not differentiate between state and federal governments. Many participants accused the government of negligence, questioning the capability of the government to store their data safely, keep their data private, or prevent the misuse of data:

> *I don't trust that the Government has suitable security measures to protect my information, and that anything that they provide would therefore be suspect to misuse*

Participants referred to the government's poor track record and cited past mistakes:

> *The current government's record of online information processing has not been good. Look at what happened with the census*

> *Government is not very good at stopping anything in the past, E.g. bin full of census papers*

Participants questioned the ethical use of their health information by the government, accusing the government of corruption, especially if profit making is involved.

Profit generation

Participants expressed reluctance to share their government health data with private companies if the end goal is solely to generate a profit and not for the benefit of society. Participants commented that if the data is being accessed for free then the profits should be shared. For example one participant commented: "I would never trust a private drug company. They make billions and if they want my info than they should pay for it. It would be wrong to just give it to them on a whim of the Government especially one like we have now that is in the pockets of big business". Participants expressed concern that if a private company was motivated by profit generation or greed, then their actions in developing new treatments may be risky and unethical. Further, participants flagged that if profit

generation is the goal, then any new treatments that are developed would potentially not be available to the public at an affordable price. One participant commented:

> *These new treatments should not be priced out of the reach of most people once they are approved and on the market.*

Data security

Participants were concerned about the security and safety of their government health data once it has been shared, and any subsequent misuse of the data. These concerns included the hacking of stored health information, data leaks or breaches, and the competency of staff to handle the data (both in the public and private sector). For example, one participant commented:

> *Once any information like this becomes available in a database then it is just a matter of time until this information gets accessed illegally or misused*

Participants were also concerned about how the government health data will be used, in particular, if it will be shared, passed from one company to another or sold on to other companies.

There was support for the regulated access to government health data. This was linked to participants‘ concerns regarding security. Many participants highlighted the need for access to be tightly controlled and monitored, with strict regulations and security protocols in place, and penalties for data misuse. For example, one participant suggested:

> *There has to be regulation and rules to how the information is used and penalties if used in the wrong way*

Some participants suggested that the government oversee the process, for example:

> *This should be good for the entire society. However, government regulations should be strictly imposed on the private companies with regard to handling of this information to safeguard the privacy of those involved, how this information are used, and where this information ends after the research*

Other participants called for a third party authority to oversee the handling of their data: “There needs to be some sort of independent watchdog to ensure all rules are followed”.

Conditions under which government health data should be shared

Participants were in agreement that if government health data is shared with private companies then certain conditions need to be met. The most common requirements were: the anonymisation of health information, the goal of developing new treatments, and the need to obtain consent. There was also support for the requirement that the purpose of the research be for the common good (public benefit). A summary of these conditions is discussed below.

Anonymisation of government health data

The anonymisation (de-identification) of data was a prerequisite to sharing health information agreed upon by a substantial group of participants. Participants expressed a willingness to share data with a guarantee that all personal information was removed and they could no longer be identified. For example one participant commented:

> *Happy for data to be used providing there is no provision of identity*

> *I guess just have to maintain confidentiality in all individuals. Maybe can use the information but make it anonymous?*

Development of new treatments for diseases and disabilities

The purpose of sharing government health data was important to many participants. These participants indicated they would be willing to share data that would aid the development of new treatments for diseases and disabilities, to find cures, and to improve the health of society. One participant commented:

> *I would absolutely love to have more useful information passed on to anyone if it means fast tracking cures or medical relief for humanity*

A number of participants referred to their own personal experiences and supported sharing data in order to spare others suffering as they had, for example:

> *I would absolutely love to have more useful information passed on to anyone if it means fast tracking cures or medical relief for humanity*

Rather than refer specifically to 'the development of new treatments for diseases and disability', a large subset of participants supported data sharing for public benefit. Different terms were used by participants such as; "the common good", "the greater good", "to help others in the future", "to benefit society as a whole", "to help others", "benefit the community or public at large", "for the good of mankind", "for the good of all", "for the public good", "for the betterment of all", "benefit future generations", and "to benefit the population".

Consent

Giving consent (or permission) was a prerequisite to sharing health information for many participants. A common response given was:

> *My personal information should not be shared without my permission*

Some participants went further and requested an individual's consent be sought every time the data is used, not just the first time, and others wanted to be told exactly who their data would be shared with and what their data would be used for:

> *I don't believe it should be shared with anyone unless I have previously given my express permission for it to be used and then it must only be used in the way or for the purpose I have given permission. If this is to be changed in any way then I should be again asked to give my permission without which the data should not be used*

The right to 'opt-in' rather than 'opt out' of data sharing was highlighted by a subset of participants

Views regarding sharing government health information

When answering survey question six, many participants expressed strong views in support of and a willingness to share health information unconditionally while many vehemently opposed sharing. These participants did not cite concerns or conditions, rather they were clear in their views, for or against data sharing. For example a participant in support of sharing wrote:

> *I am quite supportive of moving forward with this*

> *There is no way I would EVER allow my personal information or my family's for that matter, to be shared*

Appendix 12: Stakeholder recruitment email

Dear Ms/Mr

I am writing on behalf of Prof Annette Braunack-Mayer in regards to your interest in participating in the research Exploring the Public Interest in and Social Licence for the Use of Linked Administrative Data in Therapeutic Development.

If you are still interested we are asking that you participate in an interview by telephone or face-to-face of 30-45 minutes at a time and place convenient to you. In the interview we would ask you about your understanding of the value of private sector access to linked administrative (government) data sets. We anticipate that you may draw on your own experience, the experience of others or particular case studies from the Australian or overseas experience.

Please find attached a Participant Information Sheet and Participant Consent Form. Please read both forms and sign, scan and email back the Participant Consent Form. I will then be in contact to organise a suitable interview time.

If you have any questions please do not hesitate to email or call me (02 4298 1312).

Prof Annette Braunack-Mayer is looking forward to meeting you again.

Kind regards

Belinda

Appendix 13: Stakeholder interview schedule

Information will be provided at the beginning of the interview about consent process, recording of interview, publication process and measures in place to ensure confidentiality. Participants will be cautioned not to use identifying details in talking about real case studies. The intent of the interview will be described: i.e. the development of hypothetical case studies

So to start, can you tell me what your role is at [name of company or institution]?

How long have you been working in this role?

Have you worked in similar roles elsewhere?

What sort of therapies does your company develop?

Prompts

- Drugs – e.g. chronic disease, rare disease, cancer, diabetes, vaccines etc
- Devices – e.g. stents, pacemakers, artificial hips, robotic surgery devices
- Services – e.g. robotic versus laparoscopic surgery

In your work have you ever thought it might be useful to have access to linked administrative data sets held in the government sector?

Prompts

- Why?
- On more than one occasion?
- Have circumstances changed which have made this need more or less likely?

Have you attempted to gain access to linked administrative data sets held in a government sector? What was the response?

Do you know of cases outside of your company/institution, perhaps overseas, where access to linked administrative data sets held in the government sector would have been helpful in development of therapies?

Here is a case study describing a particular therapeutic development which used linked administrative data sets that I would like you to consider: [Explain case developed from the scoping review/research team where access by private companies to linked administrative data promoted or supported the development of new therapies by private companies.}

Given our purpose here today, does this seem like a valuable example for Australian data custodians to consider when deliberating on the provision of access to private companies to linked de-identified administrative data sets?

Prompts

- Why is it helpful, unhelpful?
- Is it relevant to the Australian context?
- How would you change it?

What barriers do you see to private companies accessing linked administrative data sets held in the government sector?

Prompts

- Lack of support in government
- Pressure groups/media pressure
- Ethical dilemmas

Assuming access is granted, what barriers, if any, do you see to private companies being able to successfully use linked administrative data sets held in the government sector to support therapeutic development?

Prompts

- Lack of data
- Lack of expertise
- Cost of access
- Cost of analysis

What do you think are the really big issues in private sector access to linked administrative (government) data sets which will need to be resolved in the future?

Appendix 14: Hypothetical case studies

Hypothetical 1: Submission to support inclusion of a new drug on the Pharmaceutical Benefits Scheme using Cancer Registry data

An international pharmaceutical company wishes to bring its recently developed oncology medicine to the Australian market. The medicine extends life and has fewer side effects than existing medicines currently provided through the government-subsidised Pharmaceutical Benefits Scheme (PBS) for the same condition. The company approaches a Cancer Registry based in a State Department of Health with a request for aggregated, and therefore de-identified, data. Requests to the Cancer Registry for data are possible by application, and under a user-pay model. They would like to understand current treatment patterns in Australia, including age of onset of conditions, additional treatments, and information on variations in care provision across Australia.

Patient data held in the Cancer Registry are collected, without patient consent, from pathology laboratories, hospitals, radiotherapy and medical oncology departments, aged care facilities and the Registry of Births, Deaths and Marriages. The data collection is authorised under an Act of Parliament.

The Cancer Registry provides the company with aggregate data affording additional evidence to support a submission to the Pharmaceutical Benefits Advisory Committee. The submission is ultimately successful and the medicine is subsequently funded through Australia's Pharmaceutical Benefits Scheme.

An international pharmaceutical company wishes to bring its recently developed oncology medicine to the Australian market. The medicine extends life and has fewer side effects than existing medicines currently provided through the government-subsidised Pharmaceutical Benefits Scheme (PBS) for the same condition. Patient support groups are calling for the medicine to be available in Australia and for it to be publicly funded. Following registration, the company is intending to make a submission to the Pharmaceutical Benefits Advisory Committee (PBAC) to have the drug listed on the Pharmaceutical Benefits Scheme. To support their application the company wishes to know what the current treatment patterns are for people with this cancer in Australia, including how long patients survive on current treatment options, with a breakdown by age of onset and additional treatments. They are particularly interested in variation in care provision across Australia. They intend to use the data to develop a case which demonstrates patient need for better treatments in Australia and that this would be a cost-effective use of public funds.

The company approaches a Cancer Registry based in a State Department of Health with a request for aggregated, and therefore de-identified, data, which is possible upon application and under a user-pay model. Patient data held in the Cancer Registry are collected, without patient consent, from pathology laboratories, hospitals, radiotherapy and medical oncology departments, aged care facilities and the Registry of Births, Deaths and Marriages. The data collection is authorised under an Act of Parliament. Patients are not aware that their data are being used in this way for any research conducted with data from the registry.

The Cancer Registry provides the company with aggregate data the company's submission. The medicine subsequently receives a positive recommendation from the PBAC, and is listed on the PBS. The company receives no data which could identify patients. The payment contributes to the Cancer Registry's funding base.

Questions

1. Would it be acceptable for the Company to receive de-identified data (i.e. without names, ages and addresses) which had not been aggregated, so that they could conduct their own analyses?

2. Would it be acceptable for the Company to work with a research partner in an academic institution? The

academic researchers would receive de-identified data (i.e. without names, ages and addresses) which had not been aggregated, so that they could conduct the analyses.

3. The company and its academic research collaborator decide that the information is useful and relevant, and decide to request linkage of data to other Australian health care datasets. The data would be protected within a secure research environment and the company would not have access to any patient-level data, as the academic collaborator would conduct the analyses with appropriate IRB approval. Would it be acceptable to seek potential for data linkage and IRB review with these conditions to further understand, and report on, medicines utilisation and health outcomes in Australian practice?

Hypothetical 2: Monitoring safety of therapeutic devices

A range of therapeutic implantable devices are made available in Australia for a debilitating condition. Initially the availability of the new treatment is widely welcomed by patient groups. The devices are funded by Medicare as part of a surgical procedure.

Within a year, it has become apparent that the devices can have serious side effects including chronic pain and infection but this knowledge is not widely disseminated. Some physicians and patient advocates call for controlled trials of the use of the devices. Health Technology Assessment bodies call for the devices to be 'archived' because of lack of evidence to support their safety. It is several years before the TGA acts on the adverse event reports submitted by patients and physicians to withdraw approval for the device. In light of these events, a clinician group, with support from government, establishes a registry to monitor similar devices.

A device company wishes to undertake ongoing monitoring of a recently approved device for treatment of a similar condition. Their clinical trials suggest this treatment is safe and effective but, given the recent history, the company is cautious. Monitoring the device will require linking of hospital emergency attendances, hospital surgery data, Medicare and PBS data and data from the registry. The company engages a university research group to undertake the data collection and analysis. Data linkage is undertaken by a government funded data linkage unit. This process takes 12 months because of delays with human research ethics committee and governance approvals. The new device proves to be associated with some adverse events, but only in a particular subset of patients. The indications for the use of the device are adjusted in the Medicare Benefits Schedule so that the device is no longer used for these patients.

A range of therapeutic implantable devices are made available in Australia for a debilitating condition. Initially the availability of the new treatment is widely welcomed by patient groups. The devices are funded by Medicare as part of a surgical procedure.

Within a year, it has become apparent that the devices can have serious side effects including chronic pain and infection but this knowledge is not widely disseminated. Some physicians and patient advocates call for controlled trials of the use of the devices. Health Technology Assessment bodies call for the devices to be 'archived' because of lack of evidence to support their safety. It is several years before the TGA acts on the adverse event reports submitted by patients and physicians to withdraw approval for the device. There is considerable public outcry about this delay and a senate enquiry calls for increased monitoring of implantable devices. In light of these events, a clinician group, with support from government, establishes a registry to monitor similar devices.

A device company wishes to undertake ongoing monitoring of a recently approved device for treatment of a similar condition. Their clinical trials suggest this treatment is safe and effective but, given the recent history, the company is cautious. Monitoring the device will require linking of hospital emergency attendances, hospital surgery data, Medicare and PBS data and data from the registry. The

company engages a university research group to undertake the data collection and analysis. Data linkage is undertaken by a government funded data linkage unit. This process takes 12 months because of delays with human research ethics committee and governance approvals. The new device proves to be associated with some adverse events, but only in a particular subset of patients. The indications for the use of the device are adjusted in the Medicare Benefits Schedule so that the device is no longer used for these patients.

Questions

1. Cost of linkage and analysis of data: Industry argues that the public will benefit from ongoing monitoring of all implantable devices and calls for the government to share the cost of data linkage and analysis for monitoring these new devices. Should the government share this cost with private industry?
2. Funding the registry: Setting up and maintaining the registry is expensive and requires ongoing funding. The registry charges for the data which they provide to companies but at least initially finds that this is insufficient to ensure the sustainability of the registry. Who should fund the gap?

Hypothetical 3: Sharing aggregate data through public websites

Health insurance companies are requesting access to aggregate estimates of particular patient outcomes, including more detailed information on disease survival times and associated complications. This data will help insurance companies to estimate future risk.

The government believes that a vibrant private health insurance sector is essential to reduce strain on public health services. It argues that aggregate health data should be available through government websites to assist the insurance industry.

Private health insurance companies are requesting access to aggregate estimates of particular patient outcomes, including more detailed information on disease survival times and associated complications. They argue that they need improved access to health data to enable better estimates of occurrence of various levels of possible losses and exposures.

A recent government review has raised concerns about the viability of the private health insurance sector in Australia due to a range of factors including global uncertainties, an ageing population, increases in the number of natural disasters, and falling public confidence in the value of insurance. The review has also expressed concern that expensive treatments are being excluded from coverage, further fuelling lack of confidence in the health insurance sector.

The government believes that a vibrant private health insurance sector is essential to prevent strain on public health services, particularly in public hospitals where availability of beds, waiting lists for treatments and ambulance ‘ramping’ causes community concern. It argues that aggregate health data should be available through government websites in a timely manner. Although such information is already publicly available, it can be difficult to access, it is rarely up-to-date or at a level of detail sufficient to assist the insurance industry.

Questions

1. In addition to aggregate data, would it be acceptable for private health insurance companies to access unit-level data?

Hypothetical 4: Release of 'de-identified' health data

In 2016 the Australian Department of Health released a de-identified data set containing 10% of Medicare Benefits Schedule and Pharmaceutical Benefits Scheme data. The data consisted of claims information made through the MBS since 1984 and through the PBS since 2003.

The release of the data was welcomed by researchers and consumer groups as an important tool for health systems research. The data was used in a range of ways by private companies, including to:

- Provide more detailed information in their submissions for public funding for new devices and drugs; and
- Identify subgroups of patients who were not receiving international standards of care for their condition.

One month after the release of the data, researchers at the University of Melbourne demonstrated that the encryption could be broken so that Medicare services provider numbers could potentially be identified. They also showed that some claimants could be identified by linking the dataset to other sources of information such as Facebook.

In 2016 the Australian Department of Health released a de-identified data set containing 10% of Medicare Benefits Schedule and Pharmaceutical Benefits Scheme data. The data consisted of claims information for claims made through the MBS since 1984 and through the PBS since 2003.

The release of the data was welcomed by researchers and consumer groups as an important tool for health systems research. The data was used in a range of ways by private companies, including to:

a. Provide more detailed information in their submissions for public funding for new devices and drugs. For example, the release of the MBS dataset provided device manufacturers with insight into the use of alternatives to the particular device they were submitting in their application to the Medical Services Advisory Committee. The dataset provided an essential piece of information that was needed to assess the value of the device. The device was approved for inclusion in the Medicare Benefits Schedule.

b. Identify subgroups of patients who were not receiving international and national standards of care for their condition. For example, a company highlighted information about a patient subgroup who were not being treated with a drug because of misplaced concerns about safety. They presented their findings at conferences, to medical and government stakeholders and advocacy organisations. The company believed this improved care for these patients as more were treated with the drug.

One month after the release of the data, researchers at the University of Melbourne demonstrated that the encryption could be broken providing the potential for Medicare services provider numbers to be identified. They also showed that some claimants could be identified by linking the dataset to other sources of information such as Facebook. The data set was rapidly removed.

A subsequent investigation by the Office of the Australian Information Commissioner proposed that two lessons could be taken from the incident. First, de-identification of a unit-level dataset was difficult, if not impossible, and they recommended release of unit-level data should be limited to trusted recipients. Second, they called for improved governance to address approval processes, risk management processes and cross-government coordination.

Questions

1. With more stringent governance processes in place, would release of de-identified Medicare data be acceptable?

UNIVERSITY
OF WOLLONGONG
AUSTRALIA

Australian Centre for
Health Engagement,
Evidence & Values

For information on this report please contact:

Dr Jackie Street
streetj@uow.edu.au
+61 242528269