Australian Government
National Health and Medical Research Council

NHMRC

A Consumer Guide to the Principles for accessing and using publicly funded data for health research

January 2016

WORKING TO BUILD A HEALTHY AUSTRALIA

Publication Details

Publication title:	A Consumer Guide to the Principles for accessing and using publicly funded data for health research
Published:	January 2016
Publisher:	National Health and Medical Research Council
NHMRC Publication reference:	G11
Online version:	www.nhmrc.gov.au/guidelines/publications/g11
ISBN Online:	978-1-925129-53-3
Suggested citation:	National Health and Medical Research Council (2015) *A Consumer Guide to the Principles for accessing and using publicly funded data for health research* Canberra: National Health and Medical Research Council.

Copyright

Contact us

To obtain information regarding NHMRC publications or submit a copyright request, contact:

E: nhmrc.publications@nhmrc.gov.au

P: 13 000 NHMRC (13 000 64672)
or call (02) 6217 9000

KEY MESSAGES

- Australia has many publicly funded data holdings, including registries, administrative databases, clinical datasets and survey data. The use of these data for research into population health, health services and clinical care provides our greatest opportunity to unlock the value in these records for the benefit all Australians, helping us solve our major health challenges and lead to new and improved programs in prevention, screening and treatment.

- Publicly funded data holdings have usually been collected subject to privacy and confidentiality conditions, and use of these data for research will maintain these requirements and not significantly increase the risk to privacy and confidentiality.

- Researchers need to seek permission from the data custodian (the agency or organisation who holds the data) prior to gaining access.

- Researchers receiving access to publicly funded data are subject to strict ethics approval processes. These ensure that the research project is worthwhile and will add to important knowledge about the health of Australians that can not be obtained using other methods.

- Researchers who access this data will be under an obligation to share their findings with the public, with health and medical professionals, or with government and private agencies working in the health area.

- These Principles provide information and guidance for researchers and data custodians when researchers seek permission from the organisation to access and use data for their research.

- The Principles clarify the roles and responsibilities of those seeking the data and those providing access to this data.

- More details are in this Consumer Guide and in *NHMRC Principles for accessing and using publicly funded data.*

What is data?

Data are pieces of information or statistics. Health related data are information about a person or a group of people collected by a range of individuals or agencies as part of:

- clinical care of patients
- administration of healthcare service providers
- designing, monitoring, and evaluating health programs and services, and
- research.

Data can similarly be collected during other activities that can also impact on health (for example education or housing information). Those collecting the data vary widely, but some examples include clinicians, researchers, hospital administrators, community organisations, health insurance companies, and universities.

Data can be identifiable, non-identifiable or re-identifiable. These terms are defined in the Glossary at Appendix B.

Where and how is this data stored?

These data are often grouped together in datasets or data bases. These may be held at a national, state/territory or local level in universities, health service organisations and administrative areas of government agencies. They use the data in their varied roles related to providing care for, and monitoring and improving the health of all Australians. Private organisations and business can also hold datasets.

The rapid development in computing technology and the Internet have provided many opportunities to collect, store and share data. Databases are rapidly becoming an essential part of international health and science activities.

Common examples of these data sets include:

- Health, hospital and ambulance data
- Education, Social Security, immigration or sporting performance data
- Police, road safety, and crime data
- Environmental, geographic or demographics data, and
- Sales data (for example alcohol or processed food sales).

How is data used for health research?

Research is critical to improving the health and wellbeing of Australians by improving health services and clinical care, as well as prevention and screening activities. As part of their work, Australia's researchers often need to use data that have been funded, collected and held by various government agencies, health services and universities.

Researchers will use this data to examine the effect of specific factors on targeted health conditions. This will allow them to identify risk factors and emerging health trends early, creating an opportunity for preventing the condition developing, or reducing the frequency and severity of the condition.

How is my privacy guaranteed when a researcher asks to use this data?

Data are valuable because they have the potential to improve health outcomes for the Australian population or specific groups of people. However, they may include information about individual Australians who have accessed publicly funded services with the expectation that their privacy will be protected. The stories and the journeys of consumers have usually been provided with the obligation that their personal and health information will be kept confidential. At the time the data was collected most people will have given their consent for the information to be used for specific purposes e.g. in their treatment, for administrative purposes like billing, and for quality improvement of the service.

Over the past decade, Australia has developed many processes to ensure that data are used appropriately, with due consideration and respect for the individuals whose information is contained in the data.

Researchers need to ask permission to access and use these datasets or databases. They also need to abide by laws that protect the privacy of an individual and the confidentiality of their health information. Additionally, before a researcher can start their research, they need to get approval from an ethics committee, and part of this approval deals with ensuring the privacy and confidentiality of data.

What are the *Principles for Accessing &Using Publicly Funded Data for Health Research*?

These **Principles** provide information and guidance for researchers and organisations when researchers seek permission from organisations to access and use data for their research.

The Principles were developed by researchers, consumer representatives and organisations that hold data. They represent a common view about sharing the data, and the roles and responsibilities of all parties.

The Principles have been released by the National Health and Medical Research Council, and have the support of the Consumer Health Forum of Australia, the Australian Government Department of Health, the Australian Institute of Health and Welfare, the Australian Bureau of Statistics, the Australian Government Department of Human Services, the Australian Electoral Commission, the Australian Institute of Aboriginal and Torres Strait Islander Studies and Universities Australia.

A glossary defining terms in this document is at Appendix B.

Who are the Principles for?

The Principles are for organisations across Australia that hold this data (called data custodians) and researchers to understand the key aspects to be considered in making a decision about access to publically funded datasets.

Research can be complex and include a range of stakeholders. A map of all sectors and organisations involved in data access for research purposes is at Appendix A.

Why are the Principles needed?

There are many different policies, legislation, organisational processes and researcher practices guiding the access to data for research purposes. What is missing is a common set of principles that all parties agree is important.

Laws protecting individual privacy and the processes of ethics committees maximise the ethical use of data during research. Making sure individuals provide informed consent before participating in research is very important, and this includes how their data will be used during and after the research activity. In addition to these activities, it is important to make sure organisations and agencies have effective processes and systems in place as they handle requests from researchers for data (often called data governance).

Organisations usually have their own procedures and documents to manage this data access process. This process can vary considerably, because these organisations exist in different states and territories, and can be in different sectors (for example health, education, justice, welfare). The organisations can also have a different view or perspectives about sharing the data with researchers, and so Australians may be unfairly prevented from realising the research opportunities and benefits access to data can create.

Because of this, the Principles aim to:

- ensure that all parties have a common set of priorities about the use of data
- encourage communication between researchers and data custodians so they can better understand each other's roles and responsibilities, and
- make sure the use of data is optimized for the benefit of all Australians.

The Principles are provided in the following pages in the same format as in the core document NHMRC Principles for Accessing &Using Publicly Funded Data for Health Research (2016), available on the websites of NHMRC, the Australian Institute of Health and Welfare and the Australian Bureau of Statistics

Principles

1. Maximise the use of publicly funded health and health-related data for research

a. Research is an important use of publicly funded health and health-related data and is vital to unlocking their value to benefit Australians by driving improvements in disease prevention and clinical care.

b. The value of data lies in its use. Use of datasets for research should be promoted, encouraged and maximised for public benefit, while also protecting privacy and confidentiality.

c. To maximise public benefit, health and medical research should develop processes and systems to incorporate and support sustainable consumer and community involvement.

d. Public trust in the use of personal data for research should not be compromised.

2. Data custodians should recognise their responsibilities and accountabilities when providing access to data for research

Custodians of publicly funded health and health-related datasets should:

a. Encourage, support and facilitate access to data, including for data linkage, to promote high quality and sustainable research outcomes.

b. Ensure that data releases for research purposes protect the privacy of individuals and maintain confidentiality of their data in accordance with relevant legislation, policies and ethics guidelines.

c. Specify the conditions of data access and licensing, and provide a description of their data holdings to assist researchers to determine whether the data are fit-for-purpose.

d. Provide researchers, whose applications for access have met the required criteria and have been approved, with timely access to data for research on a 'best endeavours' basis.

e. Seek wherever possible to harmonise access, approval and data supply processes with other data custodians responsible for similar or related data sets.

f. Be transparent and proactive about:

- requirements for ethics approval (if applicable)
- how requests or applications for data access are made, assessed, and approved, and how long assessing applications may take
- timelines for provision of data
- any charges associated with data provision
- processes to follow up on outstanding data access requests
- processes to follow when seeking review of decisions to deny data access
- how data will be made available, and
- requirements for researchers to sign confidentiality, data use and/or other agreements.

Principles

3. Researchers should recognise their responsibilities and accountabilities when accessing and using publicly funded health and health related datasets

Researchers using publicly funded health and health-related datasets should:

a. Consult with data custodians early in the research project development: be clear about their research intent, and check that the data are appropriate for the intended research purpose.

b. Comply with the *National Statement on Ethical Conduct in Human Research* (2007)[1]; and be aware of relevant state, territory and Commonwealth legislation relating to privacy and use of data.

c. Ensure that the identity of individuals is not disclosed to unauthorised persons, that data are kept secure and only used for the purpose for which their use is approved.

d. Meet obligations under legislation and in signed agreements with data custodians for the use of data; including restrictions[2] imposed on the use, sharing and destruction of data; and any requirements for review of research findings and reports.

e. Publicly acknowledge in their publications, reports and presentations the data source, and support given by data custodians.

f. Be transparent and proactive about:

- declaring conflicts of interest
- source/s of funding
- measures taken to protect and prevent misuse of data
- the quality of datasets, including identifying any limitations
- objective reporting of research findings
- consultation with stakeholders including relevant consumer representatives
- timeframes for research and publication of findings
- considerations for research translation
- compliance with the terms and conditions of data provision and usage, and
- compliance with all relevant legislation, including the Privacy Act 1988.

---

[1] http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/e72_national_statement_130207.pdf

[2] NHMRC *National Statement on Ethical Conduct in Human Research (2007); Australian Code for the Responsible Conduct of Research; Section 95 and 95A of the Privacy Act*

APPENDIX A

Publicly Funded Health Datasets Stakeholder Map

DATA USERS AND/OR PROVIDERS

GOVERNMENT DEPARTMENTS

Includes but is not limited to federal, state/territory and local departments and policy makers in the areas of:

- Ageing
- Defence
- Communication
- Education
- Environment
- Health and human services
- Industry
- Prime Minister and Cabinet
- Social Services
- Treasury
- Veterans' Affairs

University faculties, schools of health and/or medicine (or equivalent), other researchers or research organisations

Public and private hospitals, private health service providers, clinical registries

Council of Australian Governments and relevant subcommittees, other state and territory governments and agencies, state government births, deaths and marriage registries

General public, NGOs

GOVERNMENT AGENCIES

Includes but is not limited to agencies and policy makers in the areas of:

- Aboriginal and Torres Strait Islander Health
- Cancer
- Defence Science and Technology
- Health and welfare
- Nutrition research funding
- Productivity
- Safety and quality in health care
- Sports
- Statistics
- Scientific and industrial Research
- Antarctic Research Expeditions

ADDITIONAL SOURCES OF DATA

- Ambulance data
- Australian Electoral Commission
- Crime data
- Demographic data
- Education data
- Environmental data
- Geographic data
- Immigration data
- Police data
- Road safety data
- Sales data
- Social Security data
- Sporting performance data

APPENDIX B

Glossary

Term	Definition
Confidentiality	The treatment of information that an individual has disclosed in a relationship of trust and with the expectation that it will not be used or divulged to others in ways that are inconsistent with the understanding of the original disclosure, without permission[3].
Data custodian[4]	Agencies responsible for managing the use, disclosure and protection of source data used in a statistical data integration project. Data custodians collect and hold information on behalf of a data provider (defined as an individual, household, business or other organisation which supplies data either for statistical or administrative purposes). The role of data custodians may also extend to producing source data, in addition to their role as a holder of datasets.
Data quality[5]	The seven dimensions of quality are: Institutional Environment, Relevance, Timeliness, Accuracy, Coherence, Interpretability and Accessibility. All seven dimensions should be included for the purpose of quality assessment and reporting. However, the seven dimensions are not necessarily equally weighted, as the importance of each dimension may vary depending on the data source and context.
Fit-for-purpose	The quality of the dataset is sufficient to be able to address the research question(s).
Health and Health related data	Includes any data required to inform health research/health status. This includes data sets with health specific and non-health specific data, as well as data sets linked to data where the primary issue is health (e.g. socio economic status, education, occupational health and safety). It includes data sets from bio specimens and tissue banks.
Identifiable data[6]	Data where the identity of a specific individual can reasonably be ascertained. Examples of identifiers include the individual's name, image, date of birth or address
Privacy	The individual's right or expectation that personal and other identifying information will not be disclosed.
Public benefit	Advantage gained for parts or the whole of society from accessing and using data for research purposes. Provision of data that will contribute to the generation of knowledge that will benefit the general community or lead to the formation of policies and practices to benefit the Australian population.
Publicly funded data	Data whose collection and maintenance rely on funding from the government at relevant jurisdictional levels. These include those collected for statistical, research, or administrative purposes.
Re-identifiable data	Data from which identifiers have been removed and replaced by a code, but it remains possible to re-identify a specific individual by, for example, using the code or linking different data sets
Research[7]	The original investigation undertaken to gain knowledge, understanding and insight. It is a broad concept and there is no simple, single way to define research for all disciplines.
Non-identifiable data	Data which have never been labelled with individual identifiers or from which identifiers have been permanently removed, and by means of which no specific individual can be identified. A subset of non-identifiable data are those that can be linked with other data so it can be known that they are about the same data subject, although the person's identity remains unknown.

---

[3] Population Health Research Network Glossary - http://www.phrn.org.au/media/27174/phrn%20glossary.pdf

[4] The National Statistics Service: http://www.nss.gov.au/nss/home.nsf/pages/Data+Integration+-+Roles+and+responsibilities+of+data+custodians?opendocument

[5] Australian Bureau of Statistics *ABS Data Quality Framework* at http://www.abs.gov.au/ausstats/abs@.nsf/mf/1520.0

[6] NHMRC's National Statement on Ethical Conduct in Human Research (2007) (Updated March 2014) https://www.nhmrc.gov.au/book/chapter-3-2-databanks

[7] *Australian Code for the responsible conduct of research* http://www.nhmrc.gov.au/_files_nhmrc/publications/attachments/r39_australian_code_responsible_conduct_research_150107.pdf