TARGETED CONSULTATION DRAFT
National Health and Medical Research Council

PRINCIPLES FOR ACCESSING
AND USING PUBLICLY-FUNDED DATA
FOR HEALTH RESEARCH

Developed by NHMRC’s Prevention and Community Health Committee and the
NHMRC Data Reference Group

December 2014
Principles for Accessing and Using Publicly-Funded Data for Health Research

Purpose

Background:

Australia has many publicly-funded data holdings, including registries, administrative databases, clinical datasets and surveys, which can be used for scientific research into population health, health services and clinical care. The means of accessing these datasets or determining their fit-for-purpose to address a research question are often not well defined and this can cause confusion for researchers and data custodians.

These Principles for Accessing and Using Publicly Held Data for Health Research (the Principles) provide a framework for researchers and data custodians to consider when requests or applications are made for access to existing publicly-funded datasets for the purposes of health research. The Principles should be read in conjunction with the National Statement on Ethical Conduct in Human Research (2007)¹, the Australian Code for the Responsible Conduct of Research² and other policies and guidance such as those outlined in Appendix A.

The Principles are a guide for researchers who are requesting access to wholly or partially publicly funded datasets, and for organisations that hold publicly-funded datasets. It does not dictate a particular process that must be followed, nor restrict an organisation from using its own data access policies. Organisations holding relevant data that are not publicly-funded, such as those conducting clinical trials e.g. pharmaceutical companies, are also encouraged to consider and use the Principles.

The need for the Principles:

Researchers and data custodians often have different perspectives and priorities in data sharing. Priorities for custodians include adherence to legislation, privacy protection, ensuring that the suitability and limitations of the data for specific research applications are understood and the need to meet internal work demands. Researcher priorities overlap with custodian priorities, but also include expeditious study implementation, often determined by the time frames specified by research funding agencies. The Principles aim to facilitate communication between researchers and data custodians to foster a better understanding of respective roles and responsibilities, and to optimise the appropriate use of publicly-funded data for health-related research.

Scope and target audience

The Principles are for custodians and researchers using publicly funded health and health-related datasets, including linked datasets, whether held at national, state/territory or local level. This may include both aggregate and unit record data held in registries, administrative, survey-based and clinical datasets/databases, existing health research datasets/databases, and datasets held by universities.

Health-related data include datasets that are not health-specific but which may be used to address research questions about health or health services. Not all health or health-related data are publicly funded, however it should be noted that in certain situations, data collected by private individuals or organisations may be made available to government stakeholders under particular agreements. Wherever possible, these agreements should not preclude the further use of these datasets for research. Appendix B includes a map of health data key stakeholders and a list of examples of health-related data.

## Principles

### 1. Research use of publicly held health and health-related data should be maximised

- **a.** Research can be an important use of existing publicly-funded health and health-related data. Research contributes significantly to ensuring that the most important health issues for the Australian community are addressed.

- **b.** Use of existing datasets for research should be promoted, encouraged and maximised when the public benefit is expected to outweigh any risk to privacy or confidentiality.

- **c.** Researchers should expect that the conditions of data access and licensing, a description of an organisation’s data holdings, a description of the data quality and up to date and comprehensible metadata will be published and maintained.

- **d.** Publicly-funded health and health-related datasets, in non-identifiable form, should be available for research wherever legally and practically possible and should not be withheld from research. Lower risk re-identifiable data should be available either through public release or with minimal access requirements. Researchers should have the opportunity to apply to data custodians for access to high risk identifiable data.

- **e.** To maximise public benefit, key priority areas for research using publicly-funded data should be identified by researchers in collaboration with data custodians, policymakers, practitioners and the community. Major health issues and priorities determined by the NHMRC may be a useful reference point for prioritisation of access for health-related research.

- **f.** Health research using publicly-funded data should follow the principles of promoting public awareness and public engagement. This may include the publication of appropriate details of the research project, including its purpose and the datasets used at its outset and throughout the project.

### 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.** Have processes (e.g. data management plan) in place that not only cover use of the data for the purpose for which they were collected, but also any future research purposes for which they may be used. They should also include agency resource allocation for data provision.

- **b.** Manage their datasets in accordance with good management principles and encourage, support and facilitate access to data to promote high quality and sustainable research outcomes.

- **c.** Protect the privacy and confidentiality of individuals, by ensuring that data releases for research purposes are legally permitted, and accord with relevant Commonwealth and state/territory legislation and policies.

- **d.** Provide researchers with information to assess the quality of the datasets and assist the researchers to determine whether the data is fit-for-purpose to address the proposed research question(s) prior to submitting an access request.

- **e.** Provide researchers with timely access to data for research on a best endeavours basis. This should include availability of aggregate data released in the public domain, lower risk re-identifiable and confidentialised data, and the ability to apply for high risk identifiable data.
f. Be transparent and proactive about:
   - requirements for ethics approval (if applicable)
   - how data custodians manage their data sets
   - how requests or applications for data access are made, assessed, and reviewed (e.g. risk assessment), and how long applications may take
   - timelines for provision of data
   - any charges associated with data provision
   - processes in place 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
   - requirements for researchers to sign confidentiality agreements

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

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

a. Consult with data custodians early and check that the data are appropriate for the intended research purpose. Researchers can use the Applying the ABS Data Quality Framework[^3] to help ensure the quality of the data is fit for their purpose, noting that research using this data is a secondary use from the original purpose of collection by the public organisation.

b. Comply with the National Statement on Ethical Conduct in Human Research (2007)^[4].

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, and destroyed after an agreed period. Researchers should meet obligations under legislation and in signed agreements with data custodians for the use of data, including restrictions imposed on the use and sharing of data.

d. Provide data custodians and their parent organisations with opportunities to review and comment on research reports, and respect concerns expressed by data custodian organisations in relation to the way their data have been used in the compilation of draft reports. Researchers should allow a reasonable amount of time for data custodians to review and comment on the draft reports.

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

f. Be transparent about:
   - declaring conflicts of interest
   - source of funding
   - avoiding misuse of data
   - the quality of datasets, including identifying any limitations
   - objective reporting of research findings
   - consultation with stakeholders
   - timeframes for research and publication of findings
   - compliance with the terms and conditions of data provision and usage
   - compliance with all relevant legislation, including the Privacy Act 1988

[^3]: http://www.abs.gov.au/AUSSTATS/abs@.nsf/Latestproducts/1520.0Main%20Features10May%202009?opendocument&tabname=Summary&prodno=1520.0&issue=May%202009&num=&view=
### Alignment with existing data policies and guidance

The Guide complements existing data principles, policies and guidance documents including but not limited to:

<table>
<thead>
<tr>
<th>Name</th>
<th>Link to website</th>
<th>Relevance to Principles</th>
</tr>
</thead>
<tbody>
<tr>
<td>Australian Government Information Management Office (AGIMO)</td>
<td><a href="http://www.finance.gov.au/sites/default/files/Big-Data-Strategy_0.pdf">http://www.finance.gov.au/sites/default/files/Big-Data-Strategy_0.pdf</a></td>
<td>Details the whole-of-government-approach to the use of big data. This approach can be used to provide better access to Australian government data for research.</td>
</tr>
<tr>
<td>Australian Public Service Better Practice Guide for Big Data</td>
<td><a href="http://www.finance.gov.au/sites/default/files/APS-Better-Practice-Guide-for-Big-Data.pdf">http://www.finance.gov.au/sites/default/files/APS-Better-Practice-Guide-for-Big-Data.pdf</a></td>
<td>Aims to address the key considerations for government agencies when growing their capability in big data and big data analytics. It is intended initially as an introductory and educative resource for agencies looking to introduce a capability and the specific challenges and opportunities that accompany such an implementation. It is also relevant to those in industry, research and academia who can work as partners with government. It states that further work is required to provide specific guidance and approaches for managing the responsible use of the data and data analytics to address vulnerabilities in relation to privacy, security, acquisition of data and the application of insights obtained from data.</td>
</tr>
<tr>
<td>A Good Practice Guide to Sharing Your Data with Others</td>
<td><a href="http://www.nss.gov.au/nss/home.nsf/NSS/E6C05AE57C80D737CA25761D002FD676?opendocument">http://www.nss.gov.au/nss/home.nsf/NSS/E6C05AE57C80D737CA25761D002FD676?opendocument</a></td>
<td>Provides a basic framework for data sharing agreements, including key factors to consider when developing data sharing agreements. It is designed to assist government data custodians, data providers and data users as well as assisting staff working on, or contributing to, data sharing projects.</td>
</tr>
<tr>
<td>National Aboriginal and Torres Strait Islander Health Data Principles</td>
<td><a href="www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=6442472790">www.aihw.gov.au/WorkArea/DownloadAsset.aspx?id=6442472790</a></td>
<td>Organisations with significant responsibilities in Aboriginal and Torres Strait Islander (ATSI) health data should apply these principles and establish meaningful partnerships with Aboriginal and Torres Strait Islander Australians. The principles set out a culturally respectful foundation for the collection, storage and use of ATSI Australian’s health and health-related information.</td>
</tr>
<tr>
<td>The NHMRC open access policy for dissemination of research findings</td>
<td><a href="http://www.nhmrc.gov.au/grants/policy/dissemination-research-findings">http://www.nhmrc.gov.au/grants/policy/dissemination-research-findings</a></td>
<td>To maximise the benefits from research, the open access policy requires that any publication arising from NHMRC supported research must be deposited into an open access institutional repository within a twelve month period from the data of publication.</td>
</tr>
</tbody>
</table>
The Australian Research Council (ARC) Open Access Policy

http://www.arc.gov.au/applicants/open_access.htm

To maximise the benefits from research, publications resulting from research activities must be disseminated as broadly as possible to allow access by other researchers and the wider community and must be deposited into an open access institutional repository within a twelve month period from the date of publication.

State/territory policies such as:
- The DataVic Access Policy
- OpenGov NSW Data Portal
- The Commissioners for Law Enforcement Data Security (CLEDS):

http://data.nsw.gov.au

Provides lists of NSW and Victorian Government datasets that are accessible online.

CLEDS establishes standards and protocols of access to and release of law enforcement data.

NHMRC’s Ethical Considerations in Quality Assurance and Evaluation Activities (2014)


This advice is designed to assist organisations in deciding the appropriate level of oversight for quality assurance and evaluation. Organisations should consider this guidance when developing policies/advice on quality assurance activities, particularly as related to the triggers for ethical review.

Alignment with broader legislation, regulations and policies

The Guide also complements a range of broader legislation, regulations and policies, including but not limited to:

<table>
<thead>
<tr>
<th>Name</th>
<th>Link to website</th>
<th>Reason for relevance to Guide</th>
</tr>
</thead>
<tbody>
<tr>
<td>The Australian Institute of Health and Welfare Act 1987 (AIHW Act)</td>
<td><a href="http://www.aihw.gov.au/privacy-of-data/">http://www.aihw.gov.au/privacy-of-data/</a></td>
<td>The AIHW’s main functions relate to the collection and production of health-related and welfare-related information and statistics, and are specified in s.5 of the Act. The Act stipulates that the AIHW operates under a strict privacy regime which has its basis in s.29 of the Act. From the perspective of both subjects and providers of information held by AIHW, a major strength of s.29 is that it specifically imposes on recipients of data released by AIHW the same confidentiality obligations applying to AIHW staff. These include a specific restriction on being required to divulge or communicate any of that information to a court (s.29(1)(e)).</td>
</tr>
<tr>
<td><strong>The National Health Information Agreement</strong></td>
<td><a href="http://meteor.aihw.gov.au/content/index.phtml/itemld/182135">http://meteor.aihw.gov.au/content/index.phtml/itemld/182135</a></td>
<td>This is an agreement between the Australian Government and state/territory government health authorities to coordinate the development, collection and dissemination of health information in Australia, including the development, endorsement and maintenance of national data standards.</td>
</tr>
<tr>
<td><strong>The Freedom of Information Act 1982</strong></td>
<td><a href="http://www.comlaw.gov.au/Details/C2012C00231">http://www.comlaw.gov.au/Details/C2012C00231</a></td>
<td>Allows members of the public rights of access to official documents of the Government of the Commonwealth and of its agencies. Functions and powers given by this Act are to be performed and exercised, as far as possible, to facilitate and promote public access to information, promptly and at the lowest reasonable cost. The Parliament also intends to increase recognition that information held by the Government is to be managed for public purposes, and is a national resource.</td>
</tr>
<tr>
<td><strong>The National Health Act 1953</strong></td>
<td><a href="http://www.comlaw.gov.au/Details/C2013C00083">http://www.comlaw.gov.au/Details/C2013C00083</a></td>
<td>An Act relating to the provision of pharmaceutical, sickness and hospital benefits, and of medical and dental services. Section 135A Officers to Observe Secrecy prescribes the conditions under which information collected under this Act may be communicated to any person about the affairs of a third person. Subsection 3 of this section describes the limited situations when information can be communicated. Subsection 4 deems that an authority or person receiving information under subsection 3 is subject to the</td>
</tr>
<tr>
<td>Act/Section</td>
<td>Description</td>
<td>URL</td>
</tr>
<tr>
<td>-------------</td>
<td>-------------</td>
<td>-----</td>
</tr>
<tr>
<td><strong>The Health Insurance Act 1973</strong></td>
<td><a href="http://www.comlaw.gov.au/Details/C2014C00550">http://www.comlaw.gov.au/Details/C2014C00550</a></td>
<td>An Act providing for payments by way of medical benefits and payments for hospital services and for other purposes.</td>
</tr>
<tr>
<td><strong>State/Territory government legislation including privacy acts, health records acts, public health acts and other relevant legislation. For example, in NSW:</strong></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
PUBLICLY FUNDED HEALTH DATASETS STAKEHOLDER MAP

Data users and/or providers

Government departments
Includes but is not limited to departments and policy makers in areas of:
- Social Services
- Health and human services
- Ageing
- Industry
- Prime Minister and Cabinet
- Veterans’ Affairs
- Treasury
- Defence
- Communication
- Education
- State government departments of health and other relevant departments/agencies

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

Public and private hospitals, private 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 areas of:
- Health and welfare
- Statistics
- Research funding
- Safety and quality in health care
- Cancer
- Nutrition
- ATSI Health
- Productivity
- Sports
- Antarctic Research Expeditions
- Scientific and Industrial Research
- Defence Science and Technology

Examples of health-related data
- Ambulance data
- Crime data
- Demographic data
- Geographic data
- Police data
- Road safety data
- Sales data
- Social Security data
- Education data
- Sporting performance data
- Immigration data
## Glossary

<table>
<thead>
<tr>
<th>Term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Identifiable data⁶</td>
<td>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</td>
</tr>
<tr>
<td>Re-identifiable data⁶</td>
<td>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</td>
</tr>
<tr>
<td>Non-identifiable data⁶</td>
<td>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.</td>
</tr>
<tr>
<td>Health and Health related data</td>
<td>Includes any data required to inform health research/health status. This includes health data sets as well as data sets linked to data where the primary issue is health (e.g. socio economic status, education, occupational health and safety)</td>
</tr>
<tr>
<td>Privacy</td>
<td>The right or expectation that personal and other identifying information will not be disclosed.</td>
</tr>
<tr>
<td>Confidentiality</td>
<td>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 permission⁷.</td>
</tr>
<tr>
<td>Public benefit</td>
<td>Advantage gained for parts or the whole of society from accessing and using data for research purposes.</td>
</tr>
<tr>
<td></td>
<td>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</td>
</tr>
<tr>
<td>Fit-for-purpose</td>
<td>The quality of the dataset is sufficient to be able to address the research question(s).</td>
</tr>
<tr>
<td>Publicly-funded data</td>
<td>Data whose collection and maintenance rely on funding from the government at relevant jurisdictional levels. Datasets established through publicly funded modes, for example publicly funded health services.</td>
</tr>
<tr>
<td>Data Management Plan</td>
<td>The creation, collection, management, use and disposal of agency data is governed by a number of legislative and regulatory requirements. Government data needs to be managed in a way that ensures it is discoverable, accessible and useable.⁸ This should also include how the agency will address resource allocation in order to provide data.</td>
</tr>
</tbody>
</table>

---

