



Defining ‘personal information’

The office of the Information Commissioner provides a broad and extensive answer to the question: “[What is personal information?](#)”. However, the explanation provided is not specific to any context. As such, a brief functional definition that applies to the setting of health care research may assist those seeking to conduct research.

All health information about an individual is personal information as it appears in medical records. Information about an individual’s characteristics, the treatment received and outcomes of any health care assessment or procedure when associated with that individual’s identity is personal information.

Does that mean that I can’t use my data in research?

Health information that is not reasonably identifiable is no longer considered personal information as defined by the Commonwealth Privacy Act and the Health Services Act. As such, non-personal data may be used in research without taking that legislation into consideration.

How do we define what is “reasonably identifiable”?

The answer to this question is multi-factorial and heavily context dependent.

Inherently identifiable information

“Certain information may be unique to a particular individual, and therefore may (in and of itself) establish a link to the particular person. (OAIC, 2017).”

Inherently identifiable information includes a participant’s name, Medicare number, telephone number, unique medical record number (UMRN) or detailed images of a facial features. These items are inherently identifiable and the presence of any item that directly identifies a person means the data contains personal information.

Cumulatively identifiable information

Specific data may not, in and of itself, render a person’s identity knowable. However, the volume of data may lead to a threshold being crossed where an individual becomes indirectly identifiable because of the specificity of the dataset becomes such that it becomes reasonable to assume an individual could be linked to their information.

This may disproportionately occur when undertaking research into very rare or uncommon conditions, procedures or activities because of the already limited pool of candidates.

The nature of the information collected may also contribute. For example, collecting information on a specific instance of health care provision, along with the specific location, time and date that care was provided may be sufficient for an individual to be identified.

Moderating the specificity of data collected may mitigate this to some degree. For example, collecting age or year of birth as opposed to full date of birth.

Information identifiable if linked to other sources

Individuals do not “[...] necessarily need to be identified from the specific information being handled. An individual can be ‘identifiable’ where the information is able to be linked with other information that could ultimately identify the individual” (OAIC, 2017).

Examples of these data may include an individual’s genetic information being compared with searchable genetic registries, recording a specific time a colleague completed an anonymous staff survey and comparing that to rostering information or timesheets.

What does this mean for me?

Thought needs to be given to the kinds of data collected, its format, potential uses and how it will be managed and protected. While the collection of directly identifiable data automatically means a data set contains personal information, processes may be implemented to ensure datasets that do not include inherently identifiable data, but are theoretically identifiable if scrutinised or linked, to be viewed as not containing personal information. This may include reducing the specificity of the data collected, limiting who can access data and providing assurances that no attempt will be made to find out the identity of individuals.

In the event of any ambiguity or if you would simply like further information, please contact the HREC office via scgh.hrec@health.wa.gov.au

A checklist has been provided below to assist applicants in assessing whether the data required by their research contains personal or non-personal information.

1	<p>Does the project involve tracking an individual across multiple data sets?</p> <p><i>Linking data across multiple datasets usually requires that identifying or unique information is used to track specific individuals.</i></p>	<p>Yes – move to Q1b</p> <p>No – move to Q2</p>
1a	<p>Will the linkage be undertaken by other parties not involved in the research and only be made available to the research team with personal information removed?</p> <p><i>These other parties may include WA Data Linkage branch, the Australian Institute of Health and Welfare or the Department of Human Services</i></p>	<p>Yes – Move to Q2</p> <p>No: I/my team is linking data- move to question 1b</p>
1b	<p>Will privacy protecting linkage methodologies be utilised?</p>	<p>Yes – Move to Q2</p> <p>No: This activity likely requires personal data</p>
2	<p>Does the research involve accessing, collecting or using information that is unique to a particular individual and could reasonably lead to that individual being identified?</p> <p><i>This may include: name, facial imaging, UMRN, telephone or Medicare number</i></p>	<p>Yes: This is personal information.</p> <p>No: move to Q3</p>



3	<p>Does the data required for this research include information which is not immediately identifying but could, when combined with other available data, lead to a person being reasonably identified?</p> <p><i>This may include: full date of birth, genome sequence etc.</i></p>	<p>Yes: this may be personal data; continue. No – move to Q4</p>
4	<p>Is the data being collected highly specific and/or concern a condition, process or activity sufficiently rare that the identity of individuals could be reasonably ascertained?</p>	<p>Yes: this may be personal data No – Move to Q5</p>
5	<p>Are there processes in place to avoid the possibility of identities of those involved becoming known?</p> <p><i>This may include:</i></p> <ul style="list-style-type: none"> • <i>limiting access to other parties who may be able to reidentify data</i> • <i>reducing the specificity of the data items collected,</i> • <i>providing an assurance that no such attempt will be made</i> 	<p>Yes No</p>



This document can be made available in alternative formats on request.

© North Metropolitan Health Service 2020

