



## DATA DEFINITIONS

### IDENTIFYING AND NON-IDENTIFYING INFORMATION

The National Coronial Information System (NCIS) is a database containing information about deaths reported to an Australian or New Zealand Coroner. It contains highly sensitive information about deceased persons. Approval must be granted by the Victorian Justice Human Research Ethics Committee (JHREC) to access the NCIS for research purposes. When making an application, researchers should be familiar with the following terms about identifying and non-identifying data and use these terms appropriately in their application. Below is a list of definitions and examples of identifying and non-identifying data elements.

### RECOMMENDED TERMINOLOGY

**Identifying data:** data where the identity of a specific individual can be reasonably ascertained from one, or a combination of data elements (including by reference to publicly available data sources). Identifiable data elements include: name, date of birth or residential address.

**Potentially identifying data:** data where a data element, while not identifying on its own, when combined with other available data elements (within one or multiple cases) could potentially identify a specific individual. Potentially identifying data elements include: age, gender, date or location of death. Contextual information such as time or location of death is considered potentially identifying.

**Re-identifiable data:** data from which identifiers such as name and date of birth have been removed and replaced by a code for the purpose of analysis or data linkage. The identifying data elements are retained and stored separately, making it possible to re-identify the data elements and ascertain the identity of a specific individual, if required.

**Non-identifiable data:** data elements which are non-identifying at the point of receipt and have never been labelled with individual identifiers; or data from which identifiers have been permanently removed, such that a specific individual cannot be identified from the data.

**De-identified data:** data which contains at least one identifiable element in the original extract but from which the identifying elements such as name and DOB have been removed for the purpose of analysis and interrogation. Please note: de-identified data is often an intermediary step for data that will be either re-identifiable or non-identifiable at the end of the process. The National Statement on Ethical Conduct in Human Research avoids the term

'de-identified data' as its meaning is unclear. Researchers using this term should pay extra attention to clarifying the meaning applied to it.

**Anonymous data:** the term 'Anonymous' is ambiguous and as such it should not be used when describing data. Please refrain from using this terminology and use 'non-identifiable data' as an alternative

Data definition	Example
<p><b>Identifying data</b></p>	<p>A case record including the person's name, date of birth and address, e.g. John Smith, 01/01/1900, 123 Smith Street</p> <p>In addition to the core data fields, the NCIS contains supporting documents such as the police reports, autopsy and toxicology reports and coroners findings, all of which contain identifying information about the deceased and potentially about Next-Of-Kin or other persons, and are therefore classified as identifying data.</p>
<p><b>Potentially identifying data</b></p>	<p>A cohort of cases with a low incidence in a specific region , e.g. Two deaths from jet-ski accidents in a small town</p> <p>Contextual data such as the location and mechanism of death are potentially identifying in areas where there is a low frequency of such deaths or a small population. Contextual data combined with demographic data such as gender or age range increase the possibility of the combination of data elements being potentially identifying</p>
<p><b>Re-identifiable data</b></p>	<p>A case record where the identifying data elements are extracted by the researcher but removed for the purpose of analysis. e.g. John Smith DOB 01/01/2000 is replaced with DE02331</p> <p>Scenario: A full record of the death of an individual is extracted and the identifying elements are replaced with a unique identifier so that only the mechanism and cause of death are used for research purposes. If necessary, the unique identifier can be used to re-identify the data.</p>
<p><b>Non-identifiable data</b></p>	<p>A case record that does not contain identifying data elements and there is no way to link the case to the individual deceased person. e.g. A record of the mechanism and cause of death only.</p>