

13 September 2022

Gayle Milnes

The National Data Commissioner

[information@datacommissioner.gov.au](mailto:information@datacommissioner.gov.au)

# Draft Data Availability and Transparency Code

## Introduction

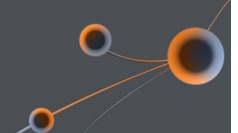
Research Australia is pleased to have the opportunity to make this submission in response to the Draft Data Availability and Transparency Code (Draft Code).

Research Australia is the national alliance representing the entire health and medical research pipeline, from the laboratory to the patient and the marketplace. We work to position Australian health and medical research as a significant driver of a healthy population and a healthy economy.

Research Australia has been an active contributor to the development of the Data Availability and Transparency Act, and we look forward to the future success of the Data Sharing Scheme. The Draft Code is an important step in this process and can play an important role in assisting data custodians and researchers to share data.

The provisions of the Act and the Draft Code that are referred to in this submission are provided as an appendix.

The following hypothetical case study is used to illustrate the issues raised in this submission.



## Estimating prescription opioid addiction- hypothetical case study

Research is proposed to estimate the prevalence of prescription opioid addiction in the Australian community.

The research is argued to be in the public interest because the research will provide information about the prevalence of prescription opioid addiction which could be used to inform decisions about what actions, if any, should be taken to address opioid addiction and what resources should be committed to such a program.

The data to be used is Pharmaceutical Benefits Scheme (PBS) prescription data on certain opioid prescriptions. The PBS data includes personal data with name and address of the patient, details about the prescriber and the dispensing pharmacist, type and number of prescriptions, the dates prescriptions are made and filled.

The dataset made available to the researcher is records at the patient level with information about the type and number of opioid prescriptions for the patient, and the dates prescriptions are filled, for a two-year period. The postcode of the patient's place of residence is provided but no other data about the identity of the patient, the prescriber and dispensing pharmacist are provided.

Combining this dataset with overseas research on the prevalence of opioid addiction among prescription opioid users based on the type of opioid prescribed, the number of prescriptions and the duration of use, an estimation is made that 0.3% of the Australian population is at risk of opioid addiction and half of this number (0.15%) are addicted to prescription opioids at any one time.

A paper is published in an academic journal which provides this statistic this finding and includes:

- a table showing aggregate numbers and types of prescriptions at a postcode level; and
- a table with estimations of addiction prevalence for capital cities and rural and regional areas, based on the prevalence of prescriptions in those areas.

## The Project Principle and public interest

The Project Principle requires consideration of whether a project for a research purpose can reasonably be expected to serve the public interest. Where the research has received Government funding, and where the research has received ethics approval, whether the research is in the public interest will already have been evaluated by the funding body and ethics committee, respectively.

Consultation Question: Is the approach to weighing arguments for and against the project serving the public interest appropriate? If not, how else could entities assess whether a project for the purpose of informing government policy and programs, or research and development, serves the public interest?

**Research Australia submits that the Code should provide guidance to decision makers to consider the provision of public funding and/or ethics approval as evidence for the project serving the public interest.**

## Outputs, final outputs and uses of outputs

Various parts of the Act and the Draft Code refer to outputs, final outputs, and to ‘uses of the outputs’. What is not clear is when the use of an output leads to a product which is itself an ‘output’ and when the use of an output leads to the creation of something which is, itself, no longer an output.

The distinction is important because the creation, use and sharing of outputs are subject to restrictions under the Act, affecting who can access them and how they can be used. Something which is the product of an output but is not itself an output is not subject to these restrictions.

### Research and data outputs

Where data is used for a research purpose, it can lead to a range of different research outputs (as distinct from data outputs), but most commonly as a paper published in an academic journal. This paper and the research findings it contains can then subsequently be used in a variety of ways which are not controlled by the researcher and may not even have been anticipated or envisaged by the researcher. This includes the paper being used by other researchers in their own research. It appears that all or part of a research paper reporting the findings of research which uses data shared under a data sharing agreement could be an ‘output’ under the Act.

Providing greater clarity in the Data Code about the boundary of the term ‘data output’ in specific circumstances and the extent to which a data sharing agreement can prescribe how the data and the data outputs are used, is going to be essential if researchers and the broader research community are to be able to use the research findings.

### The scope of ‘data’ and ‘output’

The definition of ‘data’ in section 9 of the Act is very broad. Under section 11A the output of the project is defined to include ‘any data that is the result or product of the user’s use of the shared data’. In the context of data used for a research purpose, this very broad definition leads to ambiguity, and potentially unintended consequences.

For example, in the hypothetical case study:

- Is a finding that ‘3% of the Australian population is at risk of addiction to prescription opioids’ an output of the project because it is the result or product of the user’s use of the shared data’?
- Is the table in the case study showing aggregated data at the postcode level an output of the project?
- Is the table with estimations of addiction prevalence for different areas an output of the project?

- Is the whole paper an output of the project?

There is currently a lack of clarity about what is an output of a research project and what is not. This lack of clarity has the potential to restrict the effectiveness of the Data Sharing Scheme.

Consultation Question: In practice, the output principle requires entities to agree how the accredited user will use shared data. Overall, how could the draft data code be improved to best assist entities apply the output principle?

**Research Australia submits the draft Code should provide clarification about what constitutes an output in relation to sharing of data for the purpose of research.**

Addressing the questions raised above with the case study would be useful.

Similarly, Clause 13(2) of the draft Code refers to pre-filled forms, aggregated data sets, mathematical models and publications as examples of ‘uses’ of the output. Pre-filled forms which contain personal information are a very different type of ‘use’ to an aggregated dataset or a research publication.

**Research Australia submits the Code should provide guidance on whether each of the examples provided on Clause 13(2) is:**

- a) both a use of the output and also an output in its own right, or;
- b) is considered to be a use of the output but not itself an output.

For research purposes this is an important distinction because of the restrictions that can be placed on outputs in the data sharing agreement. Research Australia would be pleased to assist in developing the guidance in this area.

## Requirements and procedures

Section 16(9) of the Act requires the custodian of the data and the accredited user to consider ‘requirements and procedures for use of the output of the project.’

If the output of the project is a broad concept, encompassing the findings and conclusions from the use of the data, then ‘requirements and procedures for use of the output of the project’ could equally be wide ranging.

Arguably the data custodian could use consideration of the ‘requirements and procedures’ to seek to restrict how and when the research findings can be published or communicated and this could extend to conclusions and findings drawn from the data shared by the custodian even if the data itself is not included in the publication or is included in a highly aggregated form.

**Research Australia submits the Data Code should provide guidance to data custodians about what requirements and procedures are reasonable, and what is not under the output principle. The guidance should include a consideration of the nature of the output and how closely the output is**

**related to the data; and how the output is related to, and serves, the public interest of the data project.**

For example, where the public interest is served by the wide publication and communication of the findings of the research, the Data Sharing Agreement should not seek to restrict how the findings are communicated. (Restrictions on how the data can be used and distributed may well be appropriate.)

## Designated Individuals

The Code has several clauses dealing with designated individuals of entities that are party to a data sharing agreement.

The use of the expression ‘permitted by the data sharing agreement to access the output’ in Clause 9 seems to be intended to have a narrow scope, associated with the data, rather than the broader concept of an output, which encompasses for example, published material.

[Consultation Question: Should the data sharing agreement include any additional details about the designated individual who is a foreign national?](#)

**Research Australia submits that Clause 9 should be amended to refer to individuals who are permitted by the data sharing agreement to access the *data*, rather than the *output*.**

This would be consistent with Clause 10, which refers to individuals ‘permitted by the data sharing agreement to access data’, and with Clause 20, which refers to ‘a designated individual for the Australian university who is permitted by the agreement to access data’.

Clause 19 of the Code seeks to restrict the individuals an entity allows to be involved in the entity’s collection or use of output. As highlighted above the output apparently/potentially includes an academic publication, which makes this provision untenable; the only way it could be complied with would be by not publishing the research.

It appears the objective of Clause 19 is to restrict the individuals who have access to the *data* rather than the full range of *outputs*. **Research Australia submits that Clause 19 should be amended so that it restricts who can collect or use the data rather than who can use the outputs.**

Again, this would be consistent with Clause 10, which refers to individuals ‘permitted by the data sharing agreement to access data.’, and with Clause 20, which refers to ‘a designated individual for the Australian university who is permitted by the agreement to access data’.

## Reporting to the Commissioner

Clause 21 provides guidance to entities about the information and documents required to be given under subsection 33(1) of the Act.

Sub-clauses 2 (k) and (l) require information to be provided, including a description of the final output and the circumstances in which any output of the project may exit the data sharing scheme.

Consultation Question: This part of the draft data code is informed by the list prescribed in section 130 of the Act. Is this an appropriate approach, and are there any additional details that should be provided to the Commissioner outside of that list?

**Research Australia submits that given the broad and somewhat ambiguous use of the term ‘output’ in the Act, the Commissioner should provide further guidance about the types of outputs referred to in sub clauses 21 (2) (k) and (l), and the circumstances in which an output of the project can exit the data sharing scheme.**

**Research Australia submits that to the greatest extent possible, annual reporting by users of data should be limited to information contained in the data sharing agreements and should be obtained directly from the register of data sharing agreements.**

## Conclusion

Research Australia is pleased to have had the opportunity to make this submission. We appreciate the Commissioner’s intention to provide a Data Code that will facilitate the sharing of data and believe our comments support the draft Data Code in this objective.

Research Australia would be pleased to assist in the further development of the draft Code as it relates to research. Please feel free to contact Greg Mullins, Head of Policy, at Research Australia, [greg.mullins@researchaustralia.org](mailto:greg.mullins@researchaustralia.org) if you require further information or would like to discuss any aspect of this submission further.

Yours sincerely

A handwritten signature in black ink that reads "Nadia Levin". The signature is written in a cursive, flowing style.

Nadia Levin  
CEO  
Research Australia

## Appendix- relevant provisions of the Act and the draft Code

### The Data Availability and Transparency Act

#### 9 Definitions

In this Act:....

**data** means any information in a form capable of being communicated, analysed or processed (whether by an individual or by computer or other automated means).

#### 11A The data sharing project

*Project, and output and ADSP-enhanced data of project*

(1) A **project** involves at least both of the following elements:

- (a) an entity (the **sharer**) shares data with another entity (the **user**), either directly or through another entity (the **intermediary**);
- (b) the user collects the data and uses the **output** of the project, which is:
  - (i) the copy of the data collected by the user; and
  - (ii) any data that is the result or product of the user's use of the shared data.

#### Section 16(9)

*Output principle*

(9) The output principle is that the only output of the project is:

- (a) the final output; and
- (b) output the creation of which is reasonably necessary or incidental to creation of the final output.

(10) The output principle includes (but is not limited to) the following elements:

- (a) the data custodian of the data and the accredited user consider:
  - (i) the nature and intended use of the output of the project; and
  - (ii) requirements and procedures for use of the output of the project;
- (b) the final output contains only the data reasonably necessary to achieve the applicable data sharing purpose or data sharing purposes.

### The Draft Code

#### 9 People principle—actions in relation to conflicts of interest

(1) As part of satisfying itself that a project is consistent with the people principle set out in subsection 16(3) of the Act, an accredited entity must do the following:

- (a) identify any actual, potential or perceived conflicts of interest of those designated individuals for the entity who are permitted by the data sharing agreement to access the output, or ADSP-enhanced data, of the project;

(b) if any such conflicts are identified—notify them to the data custodian, and any other accredited entities who are parties to the data sharing agreement and take appropriate steps to manage the conflicts.

### **10 People principle—appropriate persons**

(1) As part of satisfying itself that a project is consistent with the people principle set out in subsection 16(3) of the Act, including the element set out in paragraph 16(4)(a), of the Act, an entity must take into account the matters set out in this section in relation to individuals who are:

- (a) designated individuals for an accredited entity that is party to the data sharing agreement; and
- (b) permitted by the data sharing agreement to access data.

### **13 Output principle**

(1) As part of satisfying itself that a project is consistent with the output principle set out in subsection 16(9) of the Act, including the elements set out in subsection 16(10) of the Act, an entity must have regard to the matters set out in this section.

(2) The entity must consider the nature and intended use of the output. These may include (but are not limited to) any of the following:

- (a) pre-filled forms;
- (b) aggregated data sets (tables or unit records) for further analysis;
- (c) mathematical models for monitoring government programs;
- (d) publications such as academic journals or government reports.

### **19 Access to data by designated individuals for accredited entities**

For the purposes of subsection 19(16) of the Act, a data sharing agreement must require accredited entities to ensure that involvement in the entity's collection or use of output, or ADSP-enhanced data, of the project is restricted to designated individuals for the entity:

- (a) who are Australian citizens or permanent residents; or
- (b) otherwise—who are specified in the agreement by their full names, nationalities and designations, along with a description of their involvement in the project.

### **20 Access to data by non-Australian designated individuals for Australian universities**

(1) For the purposes of subsection 19(16) of the Act, this section prescribes requirements to be met by a data sharing agreement if:

- (a) an accredited entity that is party to the agreement is an Australian university; and
- (b) a designated individual for the Australian university who is permitted by the agreement to access data is neither an Australian citizen nor a permanent resident.

### **21 Information and documents required at time of giving documents under subsection 33(1) of the Act**

(1) For the purposes of subsection 33(2) of the Act, the entity is required to give the Commissioner:

- (a) the information set out in subsection (2) of this section, in an approved form (if any); and
- (b) if the data sharing agreement, or variation, has an attachment—the attachment; and
- (c) any other information or documents the entity considers relevant in relation to registration of the data sharing agreement or variation.



(2) For the purposes of paragraph (1)(a), the information is the following:

(a) .....

(k) a description of the final output of the project;

(l) if output of the project may exit the data sharing scheme under section 20E of the Act—the circumstances in which the exit may occur;

(m) if the agreement has an expiry date—the expiry date.