

Data Availability and Transparency Code 2022 – Consultation on draft instrument

Submission

Population Health Research Network

14 September 2022



ABOUT THE PHRN

The Population Health Research Network (PHRN) is a national data linkage infrastructure network. The PHRN commenced in 2009 and is funded by the Australian Government's National Collaborative Research Infrastructure Strategy (NCRIS), with support from state and territory government agencies and academic partners. The University of Western Australia is lead agent for the PHRN. The PHRN's primary purpose is to build and support the operation of collaborative, nationwide data linkage infrastructure capable of securely and safely linking data collections from a wide range of sources including within and between jurisdictions and across sectors and providing access to linked data¹.

Our Roles

- We are a respected, independent and trusted broker, valued for bringing governments, organisations, individuals and data together securely.
- We collaborate to enhance and maintain significant, innovative research infrastructure to improve the nation's data linkage capability.
- We facilitate and grow the use of linked data in the areas of health and human services.
- We advocate for an improved authorising environment for better access, use and sharing of data.
- We support the whole of government focus on accessing, sharing and using data for the national good.

Our Vision

Linking life data to improve the wellbeing of all Australians

Our Mission

To lead and enable the linking of data for world class, action-oriented research

Dr Merran Smith

Chief Executive



www.phrn.org.au

¹ Flack, F. and Smith, M. (2019) "The Population Health Research Network - Population Data Centre Profile", *International Journal of Population Data Science*, 4(2). doi: 10.23889/ijpds.v4i2.1130.

PHRN RESPONSE TO THE CONSULTATION

Introduction

The PHRN is pleased to have the opportunity to comment on the Exposure Draft of the Data Availability and Transparency Code 2022 (the Data Code). The PHRN funds, supports and coordinates a national research infrastructure which links and shares person-level health and human services data for research in the public interest.

The Data Code will have a significant impact on how the *Data Availability and Transparency Act 2022* (the Act) is interpreted and implemented and it is important for the research community to have a clear understanding of the requirements of the Act.

We have provided specific feedback on the Output Principle and Privacy Protection questions.

Consultation questions

Output principle

15. 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?

Various parts of the Act and the Data 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 are in the public domain and are no longer controlled by the researcher. The paper and its findings can be used by anyone for purposes that cannot be anticipated or envisaged by the researcher. 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 research context outputs could include:

- Specific research findings e.g. 1% of the population have disease X
- Tables of aggregate analysis results
- Whole publications or reports

Clarity about what is an output is required for the smooth operation of the Data Sharing Scheme. In addition, the difference between outputs and uses of outputs (see Clause 13(2)) also requires clarification.

PHRN submits the draft Code should provide clarification about what constitutes an output and what is a use of an output in relation to sharing of data for the purpose of research.

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.

PHRN 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.

Designated Individuals

The Code has a number of 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.

PHRN submits the 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*.

PHRN submits that Clause 19 should be amended so that it restricts who can be collect or use the data rather than 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.

PHRN 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.

Privacy protections

16. One of the objects of the Act is to enable the sharing of data consistently with the Privacy Act and appropriate safeguards. Does this part of the draft data code strike the right balance between holding data custodians accountable to seek consent, and providing data custodians with an exception to collect consent in circumstances where it is genuinely unreasonable or impracticable to seek consent? How could the draft data code be improved to achieve the right balance? For example, could the National Health and Medical Research Council waiver of consent guidelines be used here?

International and national research ethics guidelines and privacy guidelines usually include some provision for waiving the requirement for consent. The most common criteria for consent to be waived are:

- It is impracticable to obtain consent
- That the collection, use and sharing of the data is low risk
- There is some public benefit or interest in the use of the data

The Australian Privacy Principle Guidelines, the section 95 Guidelines of the *Privacy Act* 1988 and the National Statement on Ethical Conduct in Human Research all include these three elements as well as additional criteria. The requirements in these guidelines and the draft DAT code are overlapping but there are differences and the wording varies. Currently human

research ethics committees (HREC) considering requests for waivers of consent for medical research must interpret and apply both sets of guidance. When considering research requests under the DATA scheme, data custodians will be required to apply the DAT Code which is worded differently to the two sets of guidelines the HREC will be using. The use of three different guidelines to assess the same research application will result in confusion and inconsistent decision making.

Given that one of the objects of the Act is to enable data sharing consistent with the Privacy Act it is unclear why the guidance in relation to waiving the requirement for consent in the DAT Code would be worded differently to the Privacy Act guidelines. It will be much easier for data custodians to comply with both the Act and the Privacy Act if the wording and the guidance is the same.

The specific wording in the draft Code clauses 16 (3-4) could be interpreted to be a more stringent requirement than the Australian Privacy Principle Guidelines and are definitely more stringent than the section 95 Guidelines and National Statement on Ethical Conduct in Human Research. If the wording 'It is not unreasonable or impracticable to seek consent merely because the consent of a very large number of individuals needs to be sought' remains it may preclude the use of the DAT Scheme for population level health research.

Conclusion

Achieving the appropriate level of detail and clarity in the Data Code is critical to the success and efficiency of the DATA Scheme. Therefore, the PHRN recommends that there is a further opportunity for stakeholders to comment on changes to the Draft Data Code that result from this consultation before it is finalised.