Duncan Ivison
Deputy Vice-Chancellor (Research)

26 May 2017

Professor Anne Kelso AO
Chief Executive Officer
16 Marcus Clarke Street
Canberra City ACE 2600

Consultation on NHMRC Open Access Policy

Dear Professor Kelso,

Thank you for your invitation to provide feedback on the consultation draft of the NHMRC Open Access Policy.

In preparing our response we consulted stakeholders across the University engaged in gathering, curating, disseminating or reporting on research outputs (publications; data; IP), and in monitoring compliance with Medical Research Council and National Health (NHMRC) grant and fellowship scheme rules. These stakeholders include colleagues from the Library, Digital Research Support Team, Research Portfolio (Research office, commercialisation unit, clinical trials and ethics administration), and Human Research Ethics Committee (HREC) members committees.

While stakeholder feedback (see Attachment 1 for a summary) spanned the range of opinion from OA advocates to risk management and compliance-related roles, there was much agreement across the groups.

NHMRC Open Access Policy. There was strong support across stakeholders for NHMRC's commitment to benefit society by distributing the outcomes of the important research NHMRC fund quickly and widely to researchers and the wider community, while observing ethical and legal requirements.

The Publications policy is familiar to stakeholders, and the University will continue to invest in repository services and workflows to support open access archiving.

However, the cost of maintaining a repository infrastructure, research datasets, services and workflows is significant, and may not be within reach of smaller research institutions and commercial and government research partners.

The University considers the NHMRC's current approach alone will not optimise access to and return on investment in publicly funded research and innovation. To maximise the impact, readership and return on investment in funded Australian research and outputs, we recommend the publications policy directives – for individual researchers to archive their works, and that institutions maintain open access repositories - be supported by the NHMRC (and ARC, and other research funders) through established, scalable open access initiatives. We recommend the NHMRC and ARC (and others) consider partnership with PubMedCentral (PMC) international, [1] through archiving agreements between the major Australian research funding agencies, PMC and major publishers.

A model for this type of OA support in a research funding context similar to (but not the same as) Australia would be PMC Canada (part of PMC International).[2] The University

[1] PMC International: https://www.ncbi.nlm.nih.gov/pmc/about/pmci/
[2] See http://pubmedcentralcanada.ca/pmcc/static/aboutUs/ for an overview of PMC Canada, and https://www.ncbi.nlm.nih.gov/pmc/about/pmci/ for an overview of PMC International.

Room 646, Jane Foss Russell Building
G02
The University of Sydney
NSW 2006 Australia

T +61 2 8627 8150
F +61 2 8627 8151
E duncan.ivison@sydney.edu.au
sydney.edu.au

ABN 15 211 513 464
CRICOS 00026A

advises that CAUL (Council of University Librarians) would be well placed to investigate PMC International and other open access models and their impact, and provide a set of practical recommendations for consideration by the NHMRC and ARC.

Note: The University recommends CAUL's input into this issue as a peak industry body. While we have consulted with our University Library, The University of Sydney has not yet engaged CAUL on this issue.

Research Data. While we support NHMRC "encouraging" data sharing and use in health and medical research, we would prefer NHMRC adopt, or move towards adopting, the stronger stance being taken by international granting bodies and "mandate" data sharing and use subject to appropriate controls.

Striking a balance between openness and risk management and satisfying the interests across the spectrum of stakeholders will be difficult to achieve in practice. This was demonstrated by the recent debate in the UK sparked by the International Committee of Journal Editor's (ICMJE) proposal to share data from clinical trials. The UK stakeholders agreed, as do ours, on the importance of data sharing and that hurdles can and must be overcome to advance science for the public good – a solid foundation to build on.

The adoption of a research life-cycle approach – which incorporates strategies to improve the data literacy of researchers, university administrators, HREC members, biobank curators, clinical trials units, among others – is a large part of the solution. The NHMRC has a key role in driving this change in perspective through policy and funding rules, by engaging and educating health and medical research stakeholders, and developing guidelines and on-line training materials. The university shares this responsibility, and we would welcome the opportunity to work with the NHMRC in the research data space, towards "open by default" and the day it will be unimaginable that publicly-funded research outputs would not be shared and reused.

As the volume of research data grows exponentially, this raises concerns about the significant and growing cost to individuals and institutions of compliance, including putting data into useful form for sharing, and providing infrastructure and developing standards and frameworks for data sharing.

Further consultation and communication of proposed policy change. We do ask the NHMRC to communicate and consult widely with their researcher community before implementing the proposed changes, preferably before the busy grant period at the beginning of the year 2018.

Please contact Grant Cruchley, Senior Manager, Research Policy and Analysis, DVC Research Portfolio (grant.cruchley@sydney.edu.au) if you would like to discuss any aspect of this submission.

Yours sincerely,

Signature removed for electronic distribution

Professor Duncan Ivison
Deputy Vice-Chancellor (Research)

Attachment 1: University of Sydney - Consultation on NHMRC Open Access Policy

Attachment 1

University of Sydney – Consultation on NHMRC Open Access Policy

Summary of Comments from stakeholders

1. OA Policy: Publications (And Appendix 1.)

NHMRC (and ARC) require that publication metadata "be made available via an institutional repository immediately upon publication". Research offices and university libraries can gather and publish research output metadata records on behalf of researchers within a reasonable timeframe, but "immediately upon publication" is almost impossible in practice. We would like clarification on what interpretation of 'immediately' would be acceptable to the NHMRC – e.g., within a month?

2. NHMRC Statement of on Data Sharing.

We would prefer NHMRC adopt, or move towards adopting, the stronger stance being taken by international granting bodies and "mandate" data sharing and use subject to appropriate controls.

While striking a balance between data openness and risk management may be difficult to achieve in practice, stakeholders agreed that any hurdles must be tackled when striving for open data as the ultimate goal.

The life-cycle approach. As we noted above, the adoption of a research life-cycle view – which incorporates strategies for improving the data literacy of stakeholders including researchers, HEPS, HREC members, biobank curators, clinical trials units, among others – is a large part of the solution.

University staff have observed a digital literacy skills gap is a critical barrier to establishing data sharing, data visualisation and data publishing practices.

Universities' Libraries and Research Support Services are well placed to address this gap, and are offering a suite of training and information services to support researchers to manage, curate and communicate their data outputs at various stages within the research lifecycle.

The University's responsibility for providing research data management support and training should be noted in the policy document (as the costs of data sharing are) or guidance to promote the research data management and data publication services already on offer.

The NHMRC has a key role in driving the change in perspectives towards the research life-cycle approach through policy and funding rules, by engaging and educating health and medical research stakeholders, and developing guidelines and on-line training materials, and so on.

Education of institutional HRECs and other data providers. HRECs and other data custodians must become enablers rather than blockers of data-sharing, develop significant data literacy, and fully embrace and promote a 're-use wherever possible' mind set. The following practices, for example, need to be overcome through training/information:

- Deleting Data. Stipulations are often made by HRECs and data providers for deleting data once the research project has been completed. For example, data sets that are linked by agencies such as Australian Institute of health and Welfare (AIHW) come

with the stipulation that they be deleted after use. Linking datasets is expensive, and public funds from grants are used for this purpose. It is reasonable to expect that linked datasets should be made available to other researchers.

- Broad consent. Broad consent enables unspecified data uses are covered. Some HRECs, however, are reluctant to grant such consent. For example -
    - A university administered tissue bank where all tissue donors sign broad consent, but the LHD HREC will not accept non-targeted consent and will not release tissues in future as-yet-unknown research projects. HREC should be aware that consent should not be needed for sharing anonymised/non-identifiable data.

Data management planning. One challenge when encouraging or mandating data sharing is overcoming the complexities and costs of preparing data for sharing, particularly where this hasn't been planned for before the point of collection.

To this end we would recommend that NHMRC follow the ARC's lead, and require a statement on intentions for data sharing for all grant and fellowship applications. This would prompt researchers to adopt the research life-cycle approach to data early in the process of developing the research methodology for the project. And, where applications are successful, this statement would form the basis of more thorough data management planning to identify opportunities for sharing and reuse data around research milestones, and put in place the processes and frameworks - ethics approval, consent from participants, data management and description decisions, record keeping - to ensure high-quality data in an appropriate form can be released ethically and legally. Arrangements for data sharing and reuse should be included in final reports to the NHMRC.

Sharing of Clinical Trails Data

A number of stakeholders referred to the debate generated in the UK generated by the proposal by the International Committee of Medical Journal Editors (ICMJE) to share data underlying published clinical trials. I have attached two papers for your information with which you may already be familiar.

- The report of a meeting held by the UK Academy of Medical Sciences and Wellcome in September 2016 to discuss the ICMJE proposal.
- Lisa Rosenbaum, M.D., Bridging the Data-Sharing Divide — Seeing the Devil in the Details, Not the Other Camp, April 26 2017 [N Engl J Med. 2017 Apr 26. doi: 0.1056/NEJMp1704482.]

This debate revealed the strongly held professional opinions and sensitivities of clinical trialists around the sharing and reuse of clinical trials data and demonstrates that questions around what data should be shared, and where, when and how that data is made accessible, are discipline specific, complex and contested. Rather than a top-down access model controlled by journals, the trialists had a strong preference for a bottom-up approach, citing the benefits of trialist mediated access to data for facilitating meaningful collaboration, and managing privacy issues and risks of misinterpretation of data; whereas geneticists have openly shared data for years.

Yet, the UK stakeholders agreed, as do ours, on the importance of data sharing and that hurdles can and must be overcome to advance science for the public good. So, while encouraging and or mandating data sharing, flexibility is required to enable those who understand the data best to determine the nature of the data provided and mode of access to serve the advance science.

De-identified/Non-identified Data

While de-identification allows sensitive data to be legally shared, the selection of a methodology to effectively de-identify data, while doing minimal damage to data quality,

requires particular expertise beyond most researchers' data literacy, and for some data sets the risk of re-identification poses a particular challenge.

Ideally researchers would have access to training and guidance (e.g., the ANDS guide for HREC members) and, where necessary, expert advice and assistance to prepare their research data for sharing safely. Such expertise does not come cheaply, and adds costs to research especially impacting large-scale clinical trials.

Re-identification of data. There are many examples of supposedly de-identified information being re-identified with apparent ease. So, we recommend a risk management approach that recognises there will be specific exceptions to the default presumption that data will be shared – e.g., small populations – isolated regions, rare diseases, etc.

Data sharing in public health emergencies

We endorse a recognition of data sharing in public health emergencies, and would like to see the provision expanded to include publications. Also, data and publications are being reused in new ways and at a larger scale, such as the mining of text and datasets for large scale analysis. Taking the Zika virus as an example, research literature was mined from open access resources as a faster way to gather information. These new ways of using data recognised in the context of data sharing, to ensure researchers are aware of possibilities for their data and the potential benefits.

3. NHMRC Principles for Accessing and Using Publicly Funded Data for Health Research.

Given that the volume of research data grows exponentially, there were concerns about the significant and growing cost to individuals and institutions of complying with NHMRC's principles including putting data into useful form for sharing, and providing infrastructure and developing standards and frameworks for data sharing.

To maximise the impact, readership and return on investment in Australian research funding, we recommend the policy directives – for individual researchers to archive their works, and that institutions maintain open access repositories (publications and datasets) - be supported by the NHMRC (and ARC, and CSIRO) through large scale open access initiatives such as PMC international, including archiving agreements between funders, PMC and major publishers.

A model for this type of OA support in a research funding context similar to (but not the same as) Australia would be PMC Canada (part of PMC International)[3]. The University advises that CAUL (Council of University Librarians) would be well placed to investigate PMC International and publisher support models and their impact, and provide a set of practical recommendations for consideration by the NHMRC and ARC.

[3] See http://pubmedcentralcanada.ca/pmcc/static/aboutUs/ for an overview of PMC Canada, and https://www.ncbi.nlm.nih.gov/pmc/about/pmci/ for an overview of PMC International.