

Ensuring the success of data sharing in Canada

David Moher^{ab*} and Kelly D. Cobey^{ab}

^aCentre for Journalology, Clinical Epidemiology Program, Ottawa Hospital Research Institute, Ottawa, ON K1H 8L6, Canada; ^bSchool of Epidemiology and Public Health, Faculty of Medicine, University of Ottawa, Ottawa, ON K1G 5Z3, Canada

*dmoher@ohri.ca

Abstract

The Canadian federal Tri-Agency Research Data Management Policy has recently been released. This will require Canadian universities and other research institutes to create and share strategic plans regarding data management and to equip their researchers with skills to complete data deposits. To help maximize the success of data sharing we outline five domains for research institutions to consider during implementation: training and education, paying for data sharing, audit and feedback, meta-science, and career advancement.

“Everybody’s knowledge, nobody’s property”¹

Open science (OS) refers to making the scientific process (e.g., protocols, materials) and its outputs (e.g., reports of completed research, data, code) freely and transparently available to everybody. There have been valiant efforts to ensure easy access to COVID-19 research reports and the sharing of its underlying data. A Wellcome initiative to mandate a set of OS practices for COVID-19 research (Wellcome Trust 2020) was started early on in the pandemic and subsequently endorsed by hundreds of organizations. Despite these efforts, we have not seen meaningful change; many of the materials and outputs of COVID-19 studies remain inaccessible. An analysis of 535 COVID-19 articles on pre-print servers found that “only 21% of authors included data availability statements, and only 11% of those made their data available in external repositories” (Sumner et al. 2020). None of the data underlying any of the COVID-19 vaccine trials is directly and easily available to the scientific community, patients, or the broader community (Baden et al. 2021). Only a handful of biomedical journals have strong data-sharing policies (Naudet et al. 2018) that prospective authors must agree to as part of the submission/acceptance process.

This is about to change in Canada. Canada’s Chief Science Advisor has established a “Roadmap to Open Science” that aims to create change and embed OS into all aspects of Canadian research culture (Canadian Federal Government 2021). A key component of Canada’s transition to OS will revolve around sharing research data. Canada’s federal Tri-Agencies have recently released their Tri-Agency Research Data Management Policy (RDM Policy 2021). This policy will require action on behalf of institutions, mandate data management plans for grant applications, and includes a strong preference for data sharing. This analysis focuses on five topics to help implement data sharing successfully in Canada (see Table 1).

¹ic.gc.ca/eic/site/063.nsf/eng/h_83F7624E.html.

OPEN ACCESS

Citation: Moher D and Cobey KD. 2021. Ensuring the success of data sharing in Canada. FACETS 6: 1534–1538. doi:[10.1139/facets-2021-0031](https://doi.org/10.1139/facets-2021-0031)

Handling Editor: Jules M. Blais

Received: March 17, 2021

Accepted: May 27, 2021

Published: September 9, 2021

Copyright: © 2021 Moher and Cobey. This work is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/) (CC BY 4.0), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author(s) and source are credited.

Published by: Canadian Science Publishing

Table 1. Five recommendations to implement and achieve successful data sharing in Canada.

No.	Recommendation to achieve data sharing	Implementation
1	Training and education	- Establish a core syllabus of general data-sharing training for all biomedical researchers that can be custom tailored at each institution if needed. - Establish discipline-specific training—allow researchers funding opportunities to lead this development. - Allow institutions the opportunity to tailor the standardized content to their specific needs.
2	Paying for data sharing	- Funders should allow researchers to budget anticipated costs of data sharing and acknowledge costs may vary considerably between research areas. - Institutions should invest in and create local core data sharing service resources. In doing so, institutions must consider that much research is not funded.
3	Audit and feedback	Stakeholders should implement automated dashboards to track rates of data sharing and establish benchmarks.
4	Meta-science	- Evaluate benchmarks and progress. - Funders should recognize the value of meta-science and make funds available for meta-science projects.
5	Career advancement	- Data sharing practices should be integrated in hiring, promotion, and tenure dossiers. - Data-sharing leadership (e.g., data champions) should be recognized in hiring, promotion, and tenure.

Training and education

Few biomedical researchers have had exposure to training to enable data sharing, nor is this part of typical graduate or clinical training. We have outlined the educational and training considerations (Taylor et al. 2020) for implementing data-sharing practices. While universities may elect to develop in-house educational resources, this runs a risk of creating heterogeneity of content and (or) missing key content. We recommend that a central committee be established to work collaboratively on a syllabus of core content that can be further tailored to meet individual institutional and discipline-specific needs. This approach will help reduce mixed messaging of content to researchers across the country and reduce duplication of effort and costs.

It will be critical to engage researchers as educational training is conceived of and developed. There are likely a few core data-sharing practices that will benefit from being taught across fields within medicine. For example, the FAIR principles (Findable, Accessible, Interoperable, Reusable) are likely key. It makes sense to roll out any training of the FAIR principles and other key data-sharing concepts using standardized content. This will help ensure that core messages seamlessly cross disciplines. However, within and between disciplines there will be a need for educational tailoring. For example, sharing sequencing data is different from sharing data from randomized trials. Even within disciplines there should a push to use core content across the field where possible. To do this effectively will require identifying ways to create dialogue with the research community to build grassroots informed training that identifies and addresses context-specific needs of different research communities.

Once a syllabus is developed, we recommend that a central group lead the development of the content and as needed contracting out particular modules for development. We have developed OS material that includes a content pretest and post-test with learners who reach a certain threshold (i.e., 80%) of knowledge being able to receive a certificate of competence. These practices may be relevant to consider. Whatever resources are developed they need to be easily scalable. These resources will also need to be available in English, French, and Indigenous languages. Similarly, making them available on an open educational resource platform will be critical. We also think it important to build a database of best-practice examples of how researchers have shared their data.

Paying for data sharing

While requiring data sharing itself is a few years off in Canada, there are important considerations to think about now to ensure its successful implementation when it is required. The resources required will take time to develop and implement. To facilitate data-sharing costs, we recommend that funders allow grant applicants to include a budget line item for data sharing. The costs might differ considerably between research areas and between clinical and preclinical research. Given that many universities will be starting from a low baseline of data sharing, it is likely that for the first few years the costs of data sharing may need to be shared jointly by funders and universities. We can imagine a scenario where for every successful grantee their home research institution will contribute to a central OS fund. The ratio of funding between an institution and funder could be 1 to 2; for every dollar provided by a funder 2 dollars would be given by the research organization. This would help facilitate data sharing of unfunded research at the institution. This fund could be built up based on all funding from successful grantees thus enhancing the sustainability for data sharing resources. Such a program could be in place for a limited period of time (e.g., 2 years) to develop the data-sharing core resource program.

Audit and feedback

There is no university or funder in Canada that can easily provide information on their proportion of data sharing. While mandates and policies are important, they are likely to have a greater chance of success if audit and feedback is provided. Developing and implementing an automated dashboard of OS practices will allow universities and funders to gauge and monitor their successes with implementing data sharing and other OS practices. Prototypes of such dashboards already exist ([Charité Metrics Dashboard 2021](#)).

Institutions may elect to have different immediate thresholds of achieving data sharing. For example, Institution A may set a low benchmark (e.g., 40%) after which they will institute a steep upward trajectory to reach 80% within two years. Alternatively, institution B may already have a series of educational and training modules shared across campus and have an expectation of a higher benchmark threshold (e.g., 65%). We believe it is critical for audit and feedback to be integrated into academic institutions and funders benchmarking data sharing and other OS practices. There are examples of how audits can enhance an entire system ([Goldacre et al. 2018](#)).

Meta-science

Meta-science has recently been highlighted as one of five science strategies for a better COVID-19 recovery by the Canadian federal government and the United Nations ([United Nations 2021](#)). Meta-scientists can facilitate a comparison of data sharing (e.g., FAIR principles) across and within disciplines, academic institutions, researchers, or funders to evaluate policy success. The recommendation to use standard content for education, the role of audit and feedback, and the integration of OS dashboards can be greatly enhanced through use of meta-science. There are few opportunities in the current Canadian funding structure to support researchers conducting meta-science. To facilitate this type of research we call upon Canadian federal and provisional funders to initiate specific calls for meta-science as has happened in Europe ([Wellcome Trust 2018](#); [Dutch Research Council 2020](#)).

Career advancement

In the current Canadian landscape of publish or perish ([Rice et al. 2021](#)), researchers may feel pressured to extract as many papers out of their collected data as possible. There is now a strong movement away from this mindset. The Declaration on Research Assessment (DORA) strongly advocates against using these metrics when assessing researchers for career advancement ([DORA 2021](#)).

About 2000 organizations, globally, have signed DORA, including 27 in Canada (e.g., University of Calgary; Lunenfeld-Tanenbaum Research Institute, Toronto; Tri-Agencies). Importantly, patients participating in clinical trials appear keen on having their data shared (Mello et al. 2018). This is an important group of stakeholders without whom clinical research is not possible. Key to the success will be incentivizing and rewarding data sharing as part of a researcher's career advancement. Universities will need to integrate data sharing as part of promotion and tenure criteria. Cambridge University first developed the concept of data champions. These are typically researchers who volunteer to provide advice and assistance to the research community about ways to share their data. Champions themselves receive support and are managed by Cambridge's Office for Scholarly Communication. The Technical University of Delft, the Netherlands, have taken this one step further. Data champions at their institution can now include their role providing data-sharing help as part of their promotion and tenure portfolio. We think building networks of data champions should be seriously considered within all faculties of medicines in Canada.

Declarations

KDC is an academic representative on the Policy Committee for Research Data Canada and sits of the Advisory Board of DORA (Declaration on Research Assessment). DM leads the Hong Kong Principles initiative.

Author contributions

DM conceived and designed the study. DM and KDC analyzed and interpreted the data. DM and KDC contributed resources. DM and KDC drafted or revised the manuscript.

Competing interests

David Moher is a subject editor.

Data availability statement

All relevant data are within the paper.

References

- Baden LR, El Sahly HM, Essink B, Kotloff K, Frey S, Novak R, et al. 2021. Efficacy and safety of the mRNA-1273 SARS-CoV-2 vaccine. *The New England Journal of Medicine*, 384(5): 403–416. PMID: 33378609 DOI: 10.1056/NEJMoa2035389
- Canadian Federal Government. 2021. Tri-Agency Statement of Principles on Digital Data Management [online]: Available from ic.gc.ca/eic/site/063.nsf/eng/h_83F7624E.html.
- Charité. 2021. Charité Metrics Dashboard[online]: Available from s-quest.bihealth.org/charite-dashboard/.
- DORA. 2021. Declaration on Research Assessment [online]: Available from sfdora.org/.
- Dutch Research Council. 2020. New funding instrument to stimulate Open Science [online]: Available from nwo.nl/en/news/new-funding-instrument-stimulate-open-science.
- Goldacre B, DeVito NJ, Heneghan C, Irving F, Bacon S, Fleminger J, et al. 2018. Compliance with requirement to report results on the EU Clinical Trials Register: cohort study and web resource. *BMJ*, 362: k3218. PMID: 30209058 DOI: 10.1136/bmj.k3218

Mello MM, Lieou V, and Goodman SN. 2018. Clinical trial participants' views of the risks and benefits of data sharing. *The New England Journal of Medicine*, 378: 2202–2211. PMID: [29874542](#) DOI: [10.1056/NEJMsa1713258](#)

Naudet F, Sakarovitch C, Janiaud P, Cristea I, Fanelli D, Moher D, et al. 2018. Data sharing and reanalysis of randomized controlled trials in leading biomedical journals with a full data sharing policy: survey of studies published in *The BMJ* and *PLOS Medicine*. *BMJ*, 360: k400. PMID: [29440066](#) DOI: [10.1136/bmj.k400](#)

RDM Policy. 2021. Launch of the Tri-Agency Research Data Management Policy and a new requirement for postsecondary institutions and research hospitals [online]: Available from [ic.gc.ca/eic/site/063.nsf/eng/h_98222.html](#).

Rice DB, Raffoul H, Ioannidis JPA, and Moher D. 2021. Academic criteria for promotion and tenure in faculties of medicine: a cross-sectional study of the Canadian U15 universities. *FACETS*, 6: 58–70. DOI: [10.1139/facets-2020-0044](#)

Sumner JQ, Haynes L, Nathan S, Hudson-Vitale C, and McIntosh LD. 2020. Reproducibility and reporting practices in COVID-19 preprint manuscripts. *medRxiv*, 2020.03.24.20042796. DOI: [10.1101/2020.03.24.20042796](#)

Taylor F, Moher D, and Cobey K. 2020. Creating an audit and feedback loop to monitor and enhance data sharing in Canada. NDRIO: call for white papers on Canada's future DRI ecosystem. DOI: [10.20381/s9px-zq30](#)

United Nations. 2021. Open Dialogue with the UN Deputy Secretary-General on Science for Development in the Context of COVID-19 [online]: Available from [youtube.com/watch?v=7irnt-nkTp0&t=1s](#).

Wellcome Trust. 2018. New fund to support groundbreaking open research [online]: Available from [wellcome.org/news/new-fund-support-groundbreaking-open-research](#).

Wellcome Trust. 2020. Sharing research data and findings relevant to the novel coronavirus (COVID-19) outbreak [online]: Available from [wellcome.org/coronavirus-covid-19/open-data](#).