



## POLICY FORUM

### DATA AND DEVELOPMENT

# Open data sharing and the Global South—Who benefits?

Limited capacity, deep mistrust pose challenges to sharing

By David Serwada,<sup>1,2</sup> Paul Ndebele,<sup>3</sup>  
M. Kate Grabowski,<sup>2,4</sup> Francis Bajunirwe,<sup>5</sup>  
Rhoda K. Wanyenze<sup>1</sup>

A growing number of government agencies, funding organizations, and publishers are endorsing the call for increased data sharing, especially in biomedical research, many with an ultimate goal of open data. Open data is among the least restrictive forms of data sharing, in contrast to managed access mechanisms, which typically have terms of use and in some cases oversight by the data generators themselves. But despite an ethically sound rationale and growing support for open data sharing in many parts of the world, concerns remain, particularly among researchers in low- and middle-income countries (LMICs) in Africa, Latin America, and parts of Asia and the Middle East that comprise the Global South. Drawing on our perspective as researchers and ethicists working in the Global South, we see opportunities

to improve community engagement, raise awareness, and build capacity, all toward improving research and data sharing involving researchers in LMICs.

The Scholarly Publishing and Academic Resource Coalition defines open data as being “freely available on the internet permitting any user to download, copy, analyze... without financial, legal or technical barriers other than those inseparable from gaining access to the internet itself.” Many of the broad challenges surrounding data sharing are not entirely new. Such challenges have been explored within the context of social, policy, and research norms of the time, resulting in approaches to equitable international collaborations involving researchers in LMICs, including data sharing standards through managed access that helped ensure appropriate recognition. But more recently, as a result of changing technology that has made data collection, storage, and sharing more feasible, along with changing social and research norms driving toward openness and sharing, this prior equilibrium has been disturbed, with equitable collaboration less easily assured.

Although it is arguable that most obstacles to open data sharing are not exclusive to the Global South, stark resource inequities and, in some cases, histories of colonial oppression, present a markedly different

A field researcher for INDEPTH collects data in the village of Kokmua in Ghana, September 2007.

landscape, often characterized by limited capacity and deep mistrust, for acceptance and implementation of open data policies.

### PARACHUTES AND FREE RIDERS

North-South research collaborations are almost universally funded by Northern donors with upstream flow of specimens and data from the South and rarely vice versa. Open data effectively makes the resources collected in one country freely available to others. African scientists have expressed concern that open data compromises national ownership and reopens the gates for “parachute-research” (i.e., Northern researchers absconding with data to their home countries). Other LMIC researchers have articulated fears over free-riding scientists using the data collected by others for their own career advancement (1, 2).

Studies on data sharing among LMIC researchers have found that although they are generally supportive of data sharing, there is considerably less enthusiasm for open data. There are concerns about data misuse, violations of patient privacy through participant reidentification, and possible humiliation and exploitation of the researchers themselves (3, 4). Lack of awareness and experience with data sharing outside of trusted collaborations were also identified as obstacles to public data release (1).

Capacity building efforts, such as the National Institutes of Health (NIH) Fogarty International Center, have done much to alleviate stereotypes of the parachute researcher symbolic of Northern colonial science. Indeed, the Council for International Organizations of Medical Science ethical guidelines for human research acknowledges an essential role for capacity development in ethical international collaborations (5). However, open data may jeopardize capacity building in the absence of risk management. Research data are used not only to advance science, but also to train new investigators, increasing national autonomy and equity. It is inevitable that open data sharing will give rise to scenarios where Southern investigators do not publish all intended analyses from data sets before investigators at more well-resourced institutions in the North do so first. This is because of disadvantages inherent to LMICs, including poor digital literacy, inadequate infrastructure, and minimal investment in data science training. Additionally, many researchers in resource-poor settings have limited time to publish scientific manuscripts because a disproportionate amount of energy is spent on implementing studies and programs for Northern collaborators.

<sup>1</sup>Makere University School of Public Health, Kampala, Uganda. <sup>2</sup>Rakai Health Sciences Program, Kalisiza, Uganda. <sup>3</sup>Medical Research Council of Zimbabwe, Harare, Zimbabwe. <sup>4</sup>Department of Pathology, Johns Hopkins University, Baltimore, MD, USA. <sup>5</sup>Department of Community Health, Mbarara University for Science and Technology, Mbarara, Uganda. Email: dserwada@imul.com

Although a major objective of open data sharing is to increase the numbers and diversity of scientists addressing societal problems, this also means that open data will result in researchers using and interpreting data from communities to which they have no connections. In an ideal world, these external users would engage local researchers and communities to ensure that their science is not harmful to local communities and to avoid misinterpretation of results. In reality, such interactions may not be practical, potentially imperiling rather than promoting public health. We recognize that not all projects will necessitate local expertise and some research findings may not have direct implications for community well-being, but Northern researchers using open data should strive to conduct ethical studies relevant to local communities and return results to them.

#### CREDIT, CAPACITY, AND ENGAGEMENT

So how can Northern researchers using open data from LMICs avoid the pitfalls of colonial science and move open data forward ethically and equitably? First, there is an urgent need to increase awareness of the benefits and risks of open data among scientists and other stakeholders in biomedical research in the Global South. Such efforts could be undertaken by funding bodies such as NIH or Wellcome Trust, which have spearheaded research ethics awareness and training on data sharing in many LMICs. Strategies to mitigate risk from open data while ensuring equity could be incorporated into these future trainings.

With respect to capacity development, there must be a paradigm shift in how we recognize and credit research contributions. Journals requiring open data sharing should formally recognize the persons or groups who provide the primary data. This could be done through a taxonomy of contributorship delineating authorship roles (6). Any taxonomic contributions to data curation should include at least one author from the country where the data originate, which ideally would be mandated by the International Committee of Medical Journal Editors. Academic institutions, donors, and industry must formally recognize open data contributions and reward them in promotion and grant reviews. Without any direct benefits to data generators, there will be limited incentive for researchers to participate in open data initiatives. Increasing awareness and use of data repository software such as Dataverse, which facilitates scholarly citing of data, could aid in providing such credit.

Substantial increased donor investment in management of open data sets, training in data analysis, and infrastructure for data handling and storage in LMICs would im-

prove equity and reciprocity. Governments of LMICs, which have historically invested very little in biomedical sciences, should be core funders of these activities as well.

Ensuring community engagement and feedback in the era of open data will be challenging but is possible. Northern institutional review boards (IRBs) could play a critical role in determining whether local expertise is needed and whether a dissemination plan for results is necessary. However, it is more likely that deidentified secondary analyses using open data will not undergo IRB review. In these instances, publishers could play an important role in facilitating community engagement. One solution would be to make research findings open access to IP addresses originating in the source country (some journals already do this). Journals could require authors to provide

#### ***"There is no stopping data sharing...But it is...not without risk to...ethical international research collaborations..."***

data curators or ministries of health with the findings as a condition of publication, and ensure diversity and country or regional representation within the peer-review process. Evaluations and research are also needed to document the benefits and potential challenges arising from open access in communities providing data.

Although organizations such as the European Union Horizon 2020 program and the British Medical Research Council, among others, have developed data sharing policies and guidelines, there is a need to generate an international policy framework in which countries from LMICs fully participate to find solutions workable for all stakeholders (7). The World Health Organization (WHO) is one such organization that could help bring countries from the South and North together to formulate policies and guidelines. The WHO and research funding agencies could also facilitate country-level adaptations, capacity development, and operationalization of data sharing policies in LMICs, which are currently lacking. Once established, policies as well as adequate financial support for data sharing should be incorporated into research grants and material transfer agreements, ensuring that the necessary resources, procedures, and legal frameworks are in place prior to studies.

Some data sharing initiatives involving researchers in LMICs, such as the Malaria

Genomic Epidemiology Network (MalariaGEN), Human Heredity and Health in Africa (H3Africa), International Network for the Demographic Evaluation of Populations and Their Health (INDEPTH), and the Southern African Treatment and Resistance Network (SATuRN) collaborations, actively engage in data dissemination between countries in the Global South, not just from South to North. However, these are predominantly through managed data sharing mechanisms. Other efforts are promoting open data in LMICs, but these almost exclusively aggregate national statistics to increase government transparency (e.g., African Information Highway, Kenya Open Data Initiative). The African Academy of Sciences (AAS), in partnership with F1000-Research, announced plans to launch in 2018 a platform to enable AAS-funded and -affiliated researchers to publish quickly, without obstacles. The Bill & Melinda Gates Foundation and Wellcome Trust already host similar platforms, but they are not widely used yet. Additional efforts to encourage and sustain open data sharing are needed in the Global South. Future open data initiatives should be bottom-up, led by Southern researchers with investment from their own governments. International advocacy will be required to enhance LMIC investments in research capacity to make such sustainable data sharing objectives realities.

There is no stopping data sharing in biomedical research. Efforts to do so—which this is not—would be in vain. But it is worth noting that it is not without risk to important elements of ethical international research collaborations, including capacity development and community engagement. The global scientific community in the North and South must ensure that those societies providing and collecting the data, particularly in resource-limited settings, benefit from their contributions. Otherwise, we risk imperiling that which we seek to promote. ■

#### REFERENCES AND NOTES

- S. G. Denny, B. Silaigwana, D. Wassenaar, S. Bull, M. Parker, *J. Empir. Res. Hum. Res. Ethics* **10**, 290 (2015).
- K. Hate et al., *J. Empir. Res. Hum. Res. Ethics* **10**, 239 (2015).
- P. Y. Cheah et al., *J. Empir. Res. Hum. Res. Ethics* **10**, 278 (2015).
- S. Bull, N. Roberts, M. Parker, *J. Empir. Res. Hum. Res. Ethics* **10**, 225 (2015).
- Council for International Organization of Medical Sciences, *International Ethical Guidelines for Health-Related Research Involving Humans* (2016).
- L. Allen, J. Scott, A. Brand, M. Hlava, M. Altman, *Nature* **508**, 312 (2014).
- I. Jao et al., *J. Empir. Res. Hum. Res. Ethics* **10**, 264 (2015).

#### ACKNOWLEDGMENTS

All authors draw upon their experiences as participants in the 9th Annual Research Ethics Conference held by the Ugandan National Council for Science and Technology in Kampala, Uganda, in July 2017. The authors acknowledge J. Lessler and J. Hansen for their comments and feedback.

10.1126/science.aap8395

## Open data sharing and the Global South—Who benefits?

David Serwadda, Paul Ndebele, M. Kate Grabowski, Francis Bajunirwe, and Rhoda K. Wanyenze

Science 359 (6376), . DOI: 10.1126/science.aap8395

### View the article online

<https://www.science.org/doi/10.1126/science.aap8395>

### Permissions

<https://www.science.org/help/reprints-and-permissions>

Use of this article is subject to the [Terms of service](#)

---

Science (ISSN 1095-9203) is published by the American Association for the Advancement of Science. 1200 New York Avenue NW, Washington, DC 20005. The title *Science* is a registered trademark of AAAS.

Copyright © 2018 The Authors, some rights reserved; exclusive licensee American Association for the Advancement of Science. No claim to original U.S. Government Works