

Ethical principles for promoting health research data sharing with sub-Saharan Africa

Evelyn Anane-Sarpong  | Tenzin Wangmo | Marcel Tanner

Correspondence

Evelyn Anane-Sarpong, Department of Community Medicine, School of Medical Sciences, University of Cape Coast, Cape Coast, Ghana.

Email: evelynanane@yahoo.com

Abstract

A powerful feature of global health research is data-sharing with regions which bear the heaviest burden of disease. It offers novel opportunities for aggregating data to address critical global health challenges in ways higher than relying on individual studies. Yet there exist important stratifiers of the capacity to share data, particularly across the Global North-South divide. Systemic challenges that characterize sub-Saharan Africa and disadvantage the region's scientific productivity threaten the burgeoning data-sharing culture too. Like all endeavors requiring equal commitments under unequal circumstances, a strong ethical impetus is needed to help reduce inequities and imbalances to encourage adherence. This article discusses mandatory data-sharing in relation to peculiar challenges faced by sub-Saharan African scientists to suggest ethical principles for rethinking and reframing solutions. We propose six principles which mirror guidelines from the Institute of Medicine and encapsulate principles from the Emanuel Framework, Nairobi Data Sharing Principles, and the COHRED guidelines.

KEYWORDS

data sharing, fairness, global South, health research, inequities, sub-Saharan Africa, systemic challenges

1 | INTRODUCTION

The dominant paradigm of health research that necessarily involves primary data collection is decisively shifting to incorporate secondary data. Advances in technological and analytical tools help scientists to source different health research datasets to help in the creation of new knowledge, insights, and innovation for improving and saving lives.¹ Combined datasets, used alone or in addition to primary data, yield research solutions often superior to

results from individual studies.² Other advantages make data-sharing desirable:

- Reduced duplication of research data collection;
- Access to data that cannot be readily replicated;
- Rapid reuse of shared data to generate new insights;
- Maximized scientific knowledge returns and value on research investments as new analysts bring novel ideas and perspectives to the data;

¹Pisani E, Aaby P, Breugelmanns JG. et al.(2016).Beyond Open Data: Realising the Health Benefits of Sharing Data. *BMJ*. 355:1-5; Merson L, Gaye O, Guerin PJ. (2016). Avoiding Data Dumpsters--Toward Equitable and Useful Data Sharing. *New England Journal of Medicine*.374(25): 2414-5; Wellcome Trust. (2016). Sharing research data to improve public health: full joint statement by funders of health research. Available from: <https://wellcome.ac.uk/what-we-do/our-work/sharing-research-data-improve-publichealth-full-joint-statement-funders-health>

²Pisani E, Merson L, Ghataure A, Castillo G, Castillo A-M, Moride Y. (2018). Sharing health research data in low-resource settings: Supporting necessary infrastructure and building on good practices. Available from: https://wellcome.figshare.com/articles/Sharing_health_research_data_in_low-resource_settings_Supporting_necessary_infrastructure_and_building_on_good_practices/6042047; Gottesman M. (2015). Data Sharing: Greater Than the Sum of All Parts. *The NIH Catalyst*.525(5).

- Validation of research to promote reliability of results and improved methodology for strengthening findings;
- Re-purposed analyses to address issues left unexplored in original studies;
- Enhanced statistical significance (e.g., rare diseases through merging of datasets and combined analysis/methods otherwise difficult to realize within small samples);
- Respect for and recognition of research participants' altruism; and
- Acceleration of knowledge translation into health products and procedures.³

Data-sharing is in itself thought to promote trust, integrity, and completeness in science.⁴ It is increasingly becoming a new marker of scientists' responsibility and openness.⁵ To these ends, there are increasing global efforts to make all possible data findable, rapidly available, ethical, equitable, eternal, accessible, interoperable, and reliable.⁶

Particularly, for the Global South (or South) which comprises developing countries primarily in the southern hemisphere, including sub-Saharan Africa,⁷ sharing health research data provides an effective avenue for increasing research. The costly processes of contact with research participants, data collection, and management are reduced in studies that can rely on shared research data.⁸ Consequently, there is a growing body of literature and global actors pushing for mandatory early and complete data-sharing.⁹ In 2008, a draft international code on public health data-sharing was discussed in Mali.¹⁰ Earlier, and for over a decade starting 2005, the WHO began encouraging transparency in research through data-sharing. Other global actors have joined in codifying rules and guidelines to promote health

data-sharing. Notable funding organizations including the National Institutes of Health and the Bill and Melinda Gates Foundation have made data-sharing a condition for sponsorship.¹¹ Several influential Journals and publishers have also instituted data-sharing as a condition for publication.¹²

To date, genomic research and clinical trials seem to attract the most advanced obligatory thresholds for data-sharing.¹³ Requirements have generally been for scientists to commit part or complete research data to publicly accessible databases following stipulated periods after publication.¹⁴ The data to be shared are largely embedded as supplementary material in published articles, on institutional or project webpages, or deposited in repositories.¹⁵

The reach of influence of the burgeoning new culture of data-sharing, the compelling advantages outlined above, and others put forward in the literature¹⁶ make it reasonable to expect that the scientific community will adopt obligatory data-sharing across all health disciplines and perhaps beyond.¹⁷

Yet there are risks in data-sharing which are less deliberated,¹⁸ especially when considered in relation to the inherent equality assumed between data-producing scientists and user scientists. Some of the risks stem from lack of confidence in data-sharing; doubts concerning utility and quality of data; unwillingness to invest additional resources to make data sharable; and a general disconnect between data-originators and data-users.¹⁹ These risks are thought to be aggravated in research environments like sub-Saharan Africa which bears the greatest burden of global health problems and yet has the least dedicated resources for research.²⁰ This article is written to align our experiences in working in sub-Saharan Africa with relevant literature to explore the most typical risks facing science and scientists in the region. We contend that the risks we outline pose critical ethical hurdles that give moral grounds for giving obligatory data-sharing a

³Pisani et al., *op. cit.* note 2, p1; Taichman DB, Sahni P, Pinborg A et al. (2017). Data Sharing Statements for Clinical Trials-A Requirement of the International Committee of Medical Journal Editors. *NEJM*.376(23):2277-9; Merson et al., *op. cit.* note 1, p1; UK Medical Research Council. (2015). Data Sharing Policy. Available from: <https://www.mrc.ac.uk/documents/pdf/mrc-data-sharing-policy/>; Taichman DB, Backus J, Baethge C et al. (2016). Sharing Clinical Trial Data: A Proposal from the International Committee of Medical Journal Editors. *PLOS Medicine*; Pisani et al., *op. cit.* note 1, p.1.

⁴Ferguson L. (2014). How and why researchers share data (and why they don't). Available from: <https://www.wiley.com/network/researchers/licensing-and-open-access/how-and-why-researchers-share-data-and-why-they-dont>; Lang T. (2011). Advancing Global Health Research Through Digital Technology and Sharing Data. *Science*. 331(6018):714-717; Dallmeier-Tiessen S, Darby R, Gitmans K, et al. (2014). Enabling Sharing and Reuse of Scientific Data.16-43.

⁵Pisani et al., *op. cit.* note 2, p1; Brack M, Castillo T. (2015). Data Sharing for Public Health: Key Lessons from Other Sectors. Chatham House, London; Pisani et al., *op. cit.* note 1, p.1; Lang, *op. cit.* note 4, p.2; Parker M, Bull SJ. (2009). Ethics in collaborative global health research networks. *Bioethics Research Review*.4:165-168.

⁶Pisani E, Ghataure A, Merson L. (2018). Data sharing in public health emergencies: A study of current policies, practices and infrastructure supporting the sharing of data to prevent and respond to epidemic and pandemic threats. London: Wellcome Trust.

⁷United Nations Development Programme (UNDP). (2016). South-South Cooperation.

⁸Pisani et al., *op. cit.* note 2, p.1.

⁹US NIH, NIH Data Sharing Policy and Implementation Guidance. Available from: https://grants.nih.gov/grants/policy/data_sharing/data_sharing_guidance.htm; UK MRC, *op. cit.* note 3, p.2; The Royal Society. (2013). Principles and Guidelines for Access to Research Data from Public Funding, OECD, Editor.

¹⁰Pisani E, Whitworth J, Zaba B, Abou-Zahr C. (2010). Time for fair trade in research data. *The Lancet*. 375:703-705.

¹¹US NIH, *op. cit.* note 9, p.3; UK MRC, *op. cit.* note 3, p.2; Pisani et al., *op. cit.* note 2, p1.

¹²Taichman et al., *op. cit.* note 3, p.2; Ferguson, *op. cit.* note 4, p.2.

¹³Pisani et al., *op. cit.* note 6, p3.

¹⁴Ibid.

¹⁵Pisani et al., *op. cit.* note 6, p3; Ferguson, *op. cit.* note 4, p.2.

¹⁶Bezuidenhout L. (2019). To share or not to share: Incentivizing data sharing in life science communities. *Developing World Bioethics*. 19:18-24. <https://doi.org/10.1111/dewb.12118>; Taichman et al., *op. cit.* note 3, p.2; CIOMS. (2016). International Ethical Guidelines for Health-Related Research Involving Humans. CIOMS/WHO: Geneva; Pisani et al., *op. cit.* note 1, p.1; US NIH, *op. cit.* note 9, p.3; Pisani et al., *op. cit.* note 2, p1; Merson et al., *op. cit.* note 1, p1; UK MRC, *op. cit.* note 3, p.2; Pisani et al., *op. cit.* note 1, p.1

¹⁷Merson et al., *op. cit.* note 1, p1.

¹⁸Bezuidenhout L, Chakauya E. (2018). Hidden concerns of sharing research data by low/middle-income country scientists. *Glob Bioethics*. 29(1):39-54; Bull S. (2016). Ensuring global equity in open research. Available from: <https://doi.org/10.6084/m9.figshare.4055181>; Wellcome Trust; Merson et al., *op. cit.* note 1, p.1; Aallah G, Chantler T, Geissler PW. (2016). Global Health Research in an Unequal World: Ethics case studies from Africa. Croydon: CAB International; Pisani et al, *op. cit.* note 1, p.1; Pisani et al., *op. cit.* note 1, p.1.

¹⁹Pisani et al., *op. cit.* note 6, p3.

²⁰CIOMS, *op. cit.* note 16, p5; Rani M, Buckley, B. (2012). Systematic Archiving and Access to Health Research Data: Rationale, Current Status and Way Forward. *WHO Bulletin*. 90:12-12; Aallah et al, *op. cit.* note 18, p.5.

second look. It also gives a strong basis for treating the South unequally under the emerging culture. The issues are however, neither peculiar to sub-Saharan Africa nor solely pertaining to the Global North-South divide. They are also of interest and relevance to all collaborations that are characterized by substantial differences in expertise, financial, and technological capacities among scientists/researchers.²¹ For sub-Saharan Africa, the issues are generally underlain by systemic factors that are embedded in historically and politically rooted structural issues beyond solving by the scientific community. We raise them nonetheless because there exist well-established ethical principles that could help address or reduce their potential to perpetuate data-sharing risks in the region. We recommend six principles, three of which mirror the principles espoused by the Institute of Medicine (IOM) for sharing patients' data,²² and more broadly encapsulate the Nairobi Data Sharing Principles.²³ For clarity, the views expressed exclude data-sharing issues relating to data solicited for challenging results, exposing errors, or verifying manuscripts; data-sharing inhibitions related to research participants' welfare; and data-sharing related to commercial potential and intellectual property rights.

1.1 | Producing quality health data in sub-Saharan Africa

Large scale collection of health research data is limited in sub-Saharan Africa largely because of resource constraints.²⁴ Rather, models like the health and demographic surveillance system (HDSS) with global presence across Africa, Asia, Oceania, and Central America collect and aggregate data through small scale household surveys.²⁵ The aggregated data is used to report nationally representative data, support population health analysis, and inform national and international health decisions and policy. The introduction of projects like INDEPTH iSHARE and Data Documentation Initiative have encouraged data-sharing among sites and across the North-South divide to facilitate research.²⁶ With 88% of HDSSs in Africa, the successes of the model and its

global data-sharing ratings point to advantages that sub-Saharan Africa can realize from promoting data-sharing.²⁷ Yet, even among HDSS scientists, many are unconvinced and hesitant to share data.²⁸ Globally, some 54% of all authors do not share their data²⁹ while some 65% of those who publish peer reviewed articles also desist from sharing data or providing information that allows readers to discover or access data underlying their articles. Even for data storage in repositories, only 20% of authors deposit their own data while less than 9% of all authors share links to their data.³⁰

1.2 | To share or not to share data: the paradox of being production-rich and reward-poor

Because of sub-Saharan Africa's high burden of disease and other conditions of global health interest,³¹ the region plays host to essential data that are unavailable in other regions. Scientists are keenly aware that this ironically presents comparative advantages since many of the issues are less likely to threaten the North. The situation therefore presents international and scholarly interests for local scientists to collect data, access resources, produce scientific papers, and build research skills.³² The region's poor health indices also mean that given the right research, it has the greatest potential to make an impact on global health and the Sustainable Development Goals (SDGs).³³ Yet, limited essential skillsets in scientific productivity, scarcity of technological resources, and emigration of trained and experienced staff among other limitations stifle the region's capacity to create new knowledge and innovations for health.³⁴ Analyzing large datasets is sometimes too cumbersome for locally available analytical tools³⁵ while resources for searching and accessing data, linking and comparing, cross-referencing, aggregating, and merging datasets to identify patterns for generating insights are in limited availability.³⁶ Health institutions and scientists may therefore have rich data in terms of the critical nature of health issues still confronting the region, for which there are no existing solutions.³⁷

²¹Bezuidenhout, *op. cit.* note 16, p.5.

²²Institute of Medicine. (2015). Strategies for Responsible Sharing of Clinical Trial Data. Available from : <http://www.iom.edu/activities/research/sharingclinicaltrialdata.aspx>; Pisani et al., *op. cit.* note 6, p.3.

²³Committee on Data for Science and Technology of the International Council for Science. (2014). Data sharing principles in developing countries (the Nairobi Data Sharing Principles).

²⁴Pisani et al., *op. cit.* note 6, p.3; Rani, Buckley, *op. cit.* note 20, p.5.

²⁵INDEPTH Network. (2018). About Us. Available from: <http://www.indepth-network.org/about-us>; Herbst K et al. (2015). The INDEPTH Data Repository: An International Resource for Longitudinal Population and Health Data from Health and Demographic Surveillance Systems. *Journal of Empirical Research in Human Research Ethics*. 10(3):324-333.

²⁶Yazoume Y, Wamukoya M, Ezeh A, Emina JBO, Sankoh O. (2012). Health and demographic surveillance systems: a step towards full civil registration and vital statistics system in sub-Saharan Africa? *BMC Public Health*. 12(741); INDEPTH Network. (2016). iShare2. Available from: <http://www.indepth-network.org/projects/ishare2>; Pisani et al., *op. cit.* note 2, p.1.

²⁷Pisani et al., *op. cit.* note 1, p.1; Herbst et al., *op. cit.* note 25, p.7

²⁸Anane-Sarpong E, Wangmo T, Ward CL, Sankoh O, Tanner M, Elger, BS. (2017). "You cannot collect data using your own resources and go and put it on open access": Perspectives from Africa about public health data sharing. *Developing World Bioethics*. 00:1-12.

²⁹Ferguson, *op. cit.* note 4, p.2.

³⁰Pisani et al., *op. cit.* note 6, p.3.

³¹Global Forum for Health Research. (2008). Global Forum for Health Research. Equitable access: research challenges for health in developing countries. Geneva.

³²Pisani et al., *op. cit.* note 6, p.3.

³³United Nations. (2015). Transforming our world: the 2030 Agenda for Sustainable Development. Available from: http://www.un.org/ga/search/view_doc.asp?symbol=A/RES/70/1&Lang=E; UNDP, *op. cit.* note 7, p.3; Lang, *op. cit.* note 4, p.2.

³⁴Pisani et al., *op. cit.* note 6, p.3; Aallah et al., *op. cit.* note 18, p.5.

³⁵Serwadda D, Ndebele P, Grabowski MK, Bajunirwe F, Wanyenze RK. (2018). Open data sharing and the Global South-Who benefits? *Science*. 359:642-3; Boyd D, Crawford K. (2012). Critical questions for big data: Provocations for a cultural, technological, and scholarly phenomenon. 662-679.

³⁶Ibid; Aallah et al., *op. cit.* note 18, p.5.

³⁷CIOMS, *op. cit.* note 16, p.5.

Yet, rich data production may neither necessarily advantage local scientists in increasing publications, attaining professional reward systems which are heavily allied to publications,³⁸ or helping the region bail itself out of its problems. The paradox of being production-rich and reward-poor is that producing rich data becomes both a reason to share data and also for discouraging present data-sharing in the hope of potential future rewards. This paradox represents a complex issue underlain by inequalities and imbalances which posit a unified risk to data-sharing by the South that may be bigger than the sum and implications of data-sharing disadvantages espoused in the scholarly literature.³⁹

1.3 | Burden-benefit discrepancies

In producing and sharing data for health research, those who contribute data (burdens) must be given credit (benefits).⁴⁰ The fundamental principle of producing interoperable, reliable quality health data⁴¹ using basic technological tools goes beyond regular research processes which are often possible for one person to do within a short time. Yet they are critical because potential user-scientists have limited ways of checking the quality of those data.⁴² The burdens of running lengthy simulations, studying complex trends, designing and creating appropriate databases for data collection, and narrowing data to suit different research questions are demanding and expensive when resources are limited.⁴³ Yet, these investments are needed to increase the likelihood of utilizing data for new findings.⁴⁴ Moreover, the long manual processes benefit greatly from producing-scientists' aspirations, ideas, and intellectual goals which we deem substantial enough to compare in value with post-publication secondary analysis leading to new publications.⁴⁵ Meanwhile, scientific rigor is as central to data production as for the ultimate knowledge production.⁴⁶ However, prominence in recognition and rewards in science remain largely, if not solely, on publications.⁴⁷ Data production is rarely rewarded,⁴⁸ creating a burden-benefit discrepancy against scientists whose competitive advantages lie in data production,⁴⁹ but who for one reason or the other may fail to complete the continuum

from data to knowledge production. The discrepancy discourages data-sharing from the South.

1.4 | The “values-adherence gap”

It is reported in sub-Saharan Africa that the ethical values of data-sharing are generally embraced at institutional levels, but not as much at individual levels.⁵⁰ Apart from hints of underlining security and actuarial concerns, the reluctance stems from scientists' hopes of waiting to make the most of their data before sharing. Broader utilitarian losses arise. Figure 1 pictorially depicts a conceptualized phenomenon which we name the “values-adherence gap”. It shows good faith in data-sharing and a reluctance to adhere as a fear of it.

First, the values-adherence gap in data-sharing is fuelled by reluctance to share data in spite of good faith and trust in the benefits of doing so. The factors for reluctance which we sum in the model point to inequities in the environment, privileges, burdens, incentives, opportunities, and rewards.

Sub-Saharan Africa has the least gross domestic product globally and the least support for scientific infrastructure.⁵¹ With 85% of households not connected to the internet,⁵² the region arguably has the least access to scientific journals and publications, much less Open Access Journals.⁵³ Post graduate student earnings in the North are often three to four times higher than young PhD graduate professional earnings in the South. Young lecturers earn monthly salaries equivalent to about USD1200.00 immediately following PhD studies, a figure two to three times less than what they earned as PhD students in Europe (Personal communication with two returnee-lecturers from Tanzania and Ghana, 2018; see http://www.snf.ch/SiteCollectionDocuments/Annex_XII_Ausfuhrungsreglement_Beitragreglement_E.pdf). Computers and basic analytical tools and reference software such as EndNote which are easily available in Northern institutions at no cost to scientists must be bought by individual producer-scientists, out-of-pocket in the South.⁵⁴ The implications of these challenges include slow speeds to write and publish as well as limited potential to access and share other scientists' data. Besides, scientists who invest personal incomes in research (limited though they might be) face raised magnitudes of the perils of losing data if mandatory sharing periods elapse before they can maximize publications. For them, concerns about data-sharing override normative motivations towards taking the practice up as the right action.⁵⁵ Under-resourced scientific environments thus impede adherence to data-sharing in spite of scientists' faith in it.

³⁸Dallmeier-Tiessen et al., *op. cit.* note 4, p.2; Hodson S. (2013). Data-Sharing Culture Has Changed. *Research Information*.

³⁹Bezuidenhout L, Chakauya E. *op. cit.* note 18, p.5; Pisani et al., *op. cit.* note 1, p.1; Boyd & Crawford, *op. cit.* note 35, p.9; Dallmeier-Tiessen et al., *op. cit.* note 4, p.2; Pisani E, Abouzahr C. (2010). Sharing health data: good intentions are not enough. *Bulletin of the World Health Organization*. 88(6):462-466.

⁴⁰Committee on Data for Science, *op. cit.* note 23, p.7.

⁴¹Ibid.

⁴²Pisani et al., *op. cit.* note 2, p.1.

⁴³Hodson, *op. cit.* note 38, p.10.

⁴⁴Pisani et al., *op. cit.* note 2, p.1.

⁴⁵Boyd & Crawford, *op. cit.* note 35, p.9; Pisani et al., *op. cit.* note 6, p.3.

⁴⁶Boyd & Crawford, *Ibid.*

⁴⁷Dallmeier-Tiessen et al., *op. cit.* note 4, p.2; Hodson, *op. cit.* note 38, p.10; Pisani & Abouzahr, *op. cit.* note 39, p.10.

⁴⁸Pisani et al., *op. cit.* note 1, p.1; Pisani et al., *op. cit.* note 6, p.3.

⁴⁹Serwadda et al., *op. cit.* note 35, p.9; Pisani et al., *op. cit.* note 1, p.1.

⁵⁰Pisani et al., *op. cit.* note 2, p.1; Bezuidenhout, *op. cit.* note 16, p.5; Anane-Sarpong et al., *op. cit.* note 28, p.8; Pisani et al., *op. cit.* note 6, p.3.

⁵¹UNDP, *op. cit.* note 7, p.3; United Nations, *op. cit.* note 1, p.4; Aallah et al., *op. cit.* note 18, p.5.

⁵²International Telecommunication Union. (2017). ICT Facts and Figures 2017. Available from: <https://www.itu.int/en/ITU-D/Statistics/Pages/facts/default.aspx> ITU.

⁵³Bezuidenhout, *op. cit.* note 16, p.5.

⁵⁴Ibid.

⁵⁵Ibid.

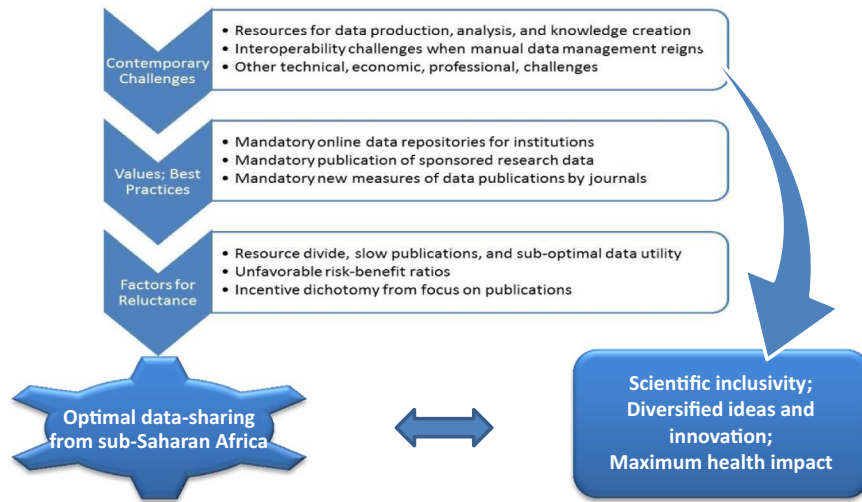


FIGURE 1 The Values-Adherence Gap [Colour figure can be viewed at wileyonlinelibrary.com]

Second, data which is under-utilized at first publication due to manual processes or use of sub-standard analytical tools holds the greatest potential for further analysis and new manuscript preparations because of their untapped knowledge-potential. Under-resourced scientists are the most likely to contribute such data because of the reasons earlier explained. When mandatory sharing times prove inadequate, producing scientists may get catapulted to a field of stiff competition for use of their data with better resourced competitors, without rights or privileges.⁵⁶ The fear of ideas being overtaken by more resourced competitors in the sub-consciously imposed and growing global professional race of “publish or perish” takes effect.⁵⁷ The corollary may include rushed thought processes and premature manuscript submissions as a means of reducing risks, perceived or real in adhering to data-sharing requirements. This values-adherence gap requires instituting protections such as data priority and exclusive user rights for producing-scientists.

Third, preserving data in ways that make for effective aggregation and third party analysis without the participation of producing scientists requires interoperability provisions which are generally beyond the competencies of under-resourced individuals. Additional expertise like biostatisticians may be sought at an extra intellectual and financial cost. This fuels the values-adherence gap: the involvement of additional expertise does not preclude need for the remaining work to benefit from the producing-scientists’ efforts; leadership, aspirations, and ideas. They must also remain linked to the initial intellectual goals of the project, demanding continuing attention from originating-scientists.⁵⁸ Ignoring such onerous data-sharing efforts may encourage the sharing of technically unusable data. Therefore, scientists who go the

extra mile to shoulder future scientists’ data needs by making data interoperable and accessible must be duly recompensed during secondary use. The scientific community could help bridge this adherence gap in two ways: institute data-sharing rewards that have comparative professional weights to the current ultimate of authorship; or make quality data production a sufficient criterion for it. The “publish [paper] or perish” paradigm must shift to encapsulate “publish [data] or perish”,⁵⁹ if the virtue of health research data can remain in its knowledge generation ability.

Fourth, because of the relatively lower availability of trained health expertise, sub-Saharan Africa like other poor regions is characterized by scientists who perform multiple roles and work under extreme pressures in congested health and or teaching facilities. There is fear that relatively better resourced researchers who are not simultaneously under similar conditions can quickly analyze and publish data before data-originators. A vicious cycle ensues (Figure 2). An intuitive response to this cycle may include the following: (a) holding on to data with plans of fuller use in manuscript writing; (b) delaying the submission of ready manuscripts until the highest number of manuscripts is ready for concurrent submission to retain control of data for as long as one is engaged in other roles; and (c) avoiding the publication of novel complex new ideas requiring release of copious data and perhaps metadata. Not only can quality, depth, urgency, and novelty in finding critical solutions be sacrificed in this cycle, but there may be reduced capacity for training and imparting scientific writing skills to upcoming scientists who get caught in the cycle.⁶⁰ This reluctance-factor against publishing novel findings would perpetuate the already low scientific productivity from the region, reduce aggregated data availability for new knowledge, and slow down the region’s reach of the SDGs.⁶¹

⁵⁶Dallmeier-Tiessen et al., *op. cit.* note 4, p.2; Hodson, *op. cit.* note 38, p.10.

⁵⁷Pisani & Abouzahr, *op. cit.* note 39, p.10; Ferguson, *op. cit.* note 4, p.2; Bezuidenhout, *op. cit.* note 16, p.5.

⁵⁸Merson et al., *op. cit.* note 1, p.1; Boyd & Crawford, *op. cit.* note 35, p.9; Pisani & AbouZahr, *op. cit.* note 39, p.10.

⁵⁹Pisani et al., *Ibid*; Pisani et al., *op. cit.* note 2, p.1; Merson et al., *op. cit.* note 1, p.1; Hodson, *op. cit.* note 38, p.10.

⁶⁰US NIH, *op. cit.* note 9, p.3; UK MRC *op. cit.* note 3, p.2; Aellah et al., *op. cit.* note 18, p.5.

⁶¹United Nations, *op. cit.* note 33, p.9.

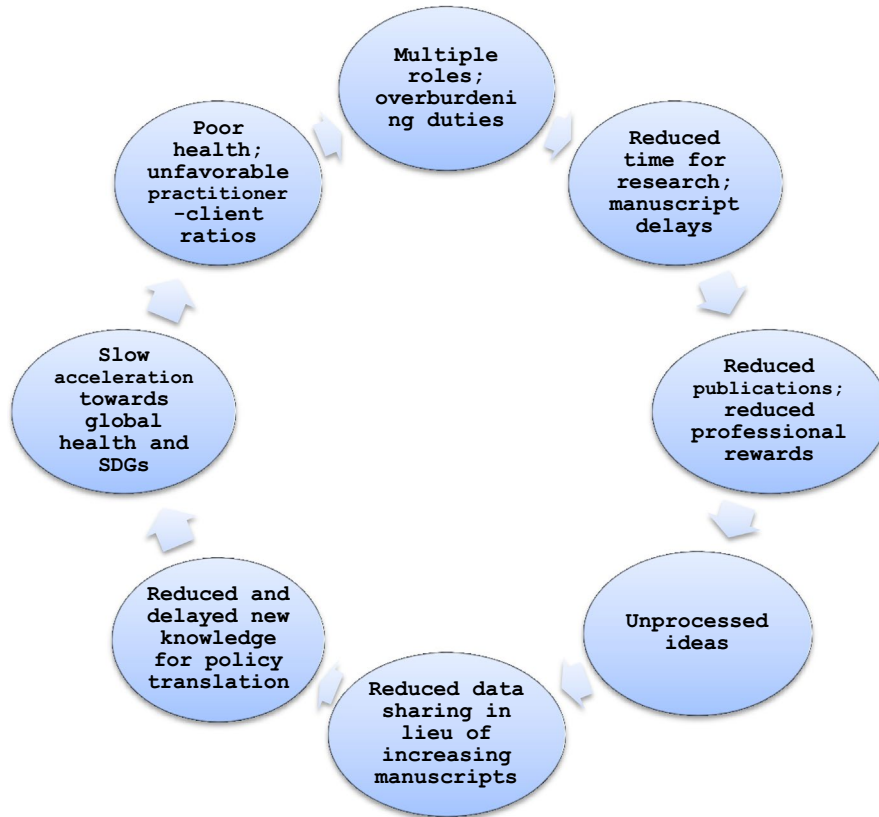


FIGURE 2 The vicious cycle of multiple roles and their effect on scientific productivity [Colour figure can be viewed at wileyonlinelibrary.com]

1.5 | Closing the values-adherence gap through principles-based solutions

It is clear that the global health vision of optimizing health research data-sharing for the common good is not by itself sufficient for promoting data-sharing from under-resourced sections of the scientific community. Attention to the potential safeties and perils for data-producing scientists need consideration.⁶² It is in the very nature of knowledge to be a public good, but data is not necessarily a public good, except when needed for health emergencies.⁶³ Mandatory sharing of what is not necessarily a public good attracts diverse individual valuing options. Calls for obligatory data-sharing should therefore be considered within a broader context beyond promoting science. It must necessarily target areas of intervention that focus on the individual: Attention is necessary to promote the interests of scientists who make both the science and data for making this science possible.⁶⁴

Although there may be much good will in data-sharing, many scientists do not wish to simply share data. They wish to benefit from

other scientists' data, the journals publishing them, and to contribute constructively to matters arising.⁶⁵ They thus want their scientific interests protected and to make positive net gains. The data-sharing aspirations will develop lopsidedly until and unless the scientific community formally acknowledges this.

A key strategy for addressing the issues raised and promoting data-sharing with sub-Saharan Africa is through a framework of ethical principles, rules that are universal, desirable, and feasible for general implementation. The relevant actors include scientists, research institutions, funders, open-science advocates, regulatory bodies, and journals/publishers. We propose six principles to help reflect on how the burgeoning new culture could be grown differently. The principles are (1) Justice; (2) Respect for scientists whose data are shared; (3) Minimizing risks; (4) Maximizing benefits; (5) Collaborative partnership; and (6) Transparency.

1.6 | Justice

Synonymous with fairness and equity,⁶⁶ this principle could be expressed in the new data-sharing culture via respect and reciprocity among scientists. Applying the principle requires the general

⁶²Pisani, et al., *op. cit.* note 1, p.1; Pisani et al., *op. cit.* note 6, p.3.

⁶³Ibid (note 6, p.3).

⁶⁴Merson L, Phong T, Nhan L, et al. (2015). Trust, respect and reciprocity: Informing culturally appropriate data sharing practice in VietNam. *Journal of Empirical Research on Human Research Ethics*. 10(3):251-63.

⁶⁵Ibid; Schuklenk, U. (2018). How can we ensure that the global south benefits from and contributes to the field of bioethics? *Developing World Bioethics*. 18(1).

⁶⁶Beauchamp TL, Childress JF. (2001). *Principles of Biomedical Ethics*. New York, NY: Oxford University Press.

benchmark of treating unequals unequally. Scientists receive different levels of support in the research environment that they work in. Until such support becomes comparable, scientists will remain unequal as far as data-sharing governance is concerned. Research environments characterized by limited sponsorship necessitate support to reduce personal individual-level research investments. Personal investments encourage perceptions of data as “owned property”, a concept that is far from data-sharing ideals.⁶⁷ Yet, ownership claims constitute genuine entitlements that should not be ignored from the data-sharing discourse either. This is because when a scientist finances, resources, generates, and preserves data in formats useable by others, they are entitled to some basic rights to “ownership”. So then, how much tolerance can be given to such entitlements, especially when perpetuated by global inequalities and most impacting on individual scientists? What role can the different actors play in addressing it? We can borrow protocols from the pharmaceutical arena that direct that well-meaning ethical science should not encourage free-riding or disproportionate benefiting off the investments of others.⁶⁸ Under-resourced scientists risk losing their “investments” or utility of their data too soon. Their data may also be under-utilized because of technological limitations to data analysis and usability.⁶⁹ They need protections.⁷⁰

We propose that the principle, Justice be actualized via formal attempts by journals/publishers, regulatory bodies and where applicable, funders to vary the periods required for mandatory sharing⁷¹ and incorporate negotiable levels of privileges including data exclusivity. Protected periods will grant temporary exclusive user rights to data-origimators as is used to protect drug originators against unfair commercial use.⁷²

As regards the systemic challenge of limited access to journals and articles “published out of reach” for many under-resourced scientists, we take cognizance of the efforts of some publishers and the WHO to increase accessibility and affordability for scientists from the South.⁷³ However, the imbalances that remain are enough to still render the sharing culture unfair. In light of this, a sub-principle of Justice, Reciprocal Justice is implicated in requiring actions that make appropriate return on gains made.⁷⁴ It can be applied in data-sharing if requests for data are made to proceed on case-by-case or solicited-basis rather than uncontrolled, unrestricted, open, and free

access.⁷⁵ This gives opportunity to prospective users to negotiate data-sharing with data-origimators. The involvement of the latter will likely increase obligations to reciprocate efforts. This will promote inclusivity in knowledge creation.⁷⁶

The foregoing proposals do not preclude other forms of compensation outlined below to compensate for secondary use of other scientists’ data. This is particularly warranted if users require datasets in whole to address new research questions.

Another application of the principle of Justice is for the public and funders to extend sponsorship to all quality research thereby reducing demands on individuals to fund research out-of-pocket. This is because compliance with data-sharing is significantly dependent on available financial and technical resources.⁷⁷ Justice must also support the protection of young and or busy (slow) scientists in the critical post-first-publication period when perceived and real risks of losing data are greatest. Even where this risk is barely perceived, its interpretation could be informed by traditions and concerns which nonetheless influence actions and reactions to mandatory data-sharing.⁷⁸ Justice will increase the generation of new ideas from otherwise under-producing scientists, assure equity in the new culture, and show empathetic recognition of effort and the stumbling blocks that handicap scientists’ own exploitation of data.

1.7 | Respect for scientists whose data are shared

The principle of respect is generally underlain by values of dignity and considerations for the welfare of the “other”.⁷⁹ In relation to data-sharing, it relates best to a conscious recognition of the efforts of all who share data and considerations that benefit them. At the point of first publication when health data is intellectually prepared, analyzed, and ready for secondary use, user-scientists have much effort and cost taken off their work. The now smaller (not necessarily inclusive of metadata) datasets require minimal processing and fewer resources to translate into new knowledge. By this principle therefore, we advocate for recognition and subtle rights of producing scientists to benefit from all potential arising from the data they share.

Many cite acknowledgements and citations for such effort, but other forms of recognition are more commensurate. In agreement with Pisani and colleagues⁸⁰ that prioritizing the recognition of publications need not preclude other deserving efforts in science, we appeal for international backing from the global academic/scientific community, including institutions to the inclusion of quality data production in the assessment of scientists’ suitability for research

⁶⁷Ferguson, *op. cit.* note 4, p.2; Capron AM, Mauron A, Elger BS, Boggio A, Ganguli-Mitra A, Biller-Andorno N. (2019). Ethical norms and the international governance of genetic databases and biobanks: findings from an international study. *Kennedy Institute of Ethics*. 19; Brack & Castillo, *op. cit.* note 5, p.3; Anane-Sarpong et al., *op. cit.* note 28, p.8.

⁶⁸Diependaele L, Cockbain J, Sterckx S. (2017). Raising the Barriers to Access to Medicines in the Developing World – The Relentless Push for Data Exclusivity. *Developing World Bioethics*. 17:11-21; Pisani et al., *op. cit.* note 8, p.3; Hodson, *op. cit.* note 38, p.10.

⁶⁹Pisani et al., *op. cit.* note 1, p.1.

⁷⁰Bezuidenhout, *op. cit.* note 16, p.5.

⁷¹Pisani et al., *op. cit.* note 2, p.1.

⁷²Diependaele et al., *op. cit.* note 68.

⁷³Schuklenk, *op. cit.* note 65, p.18.

⁷⁴Capron et al., *op. cit.* note 67, p.19; Beauchamp & Childress, *op. cit.* note 66, p.19.

⁷⁵Committee on Data for Science and Technology, *op. cit.* note 29, p.9.

⁷⁶Aallah et al., *op. cit.* note 18, p.5.

⁷⁷Merson et al., *op. cit.* note 64, p.18.

⁷⁸Bezuidenhout, *op. cit.* note 16, p.5.

⁷⁹Metz T. (2010). African and Western Moral Theories in a Bioethical Context. *Developing World Bioethics*.10.

⁸⁰Pisani et al., *op. cit.* note 1, p.1 ; Pisani et al., *op. cit.* note 2, p.1.

career progression. Other professional incentives can be created and given international recognition. Respect for scientists who share data should therefore revolve around recognition, incentivization, and motivation⁸¹ to ensure continuity of quality data production, especially from the most unreached sections of the research community.

1.8 | Minimizing risks to scientists who share data

It is our view that some of the risks outlined stem from sheer unfamiliarity with the data-sharing concept and its potential benefits. For these, training and resourcing may be adequate.⁸² Participation would help change perceptions and the fear factor too. Our focus is however, on risks revolving around unfair competition. Ethical sharing necessitates the removal of anything unfair that is preventable and in this case outweighs potential benefits.⁸³ Scientists are generally not a population who needs protection. The literature rarely discusses risks they face and where mentioned are downplayed in order not to sound patronizing of the noble profession. Mandatory data-sharing in spite of systemic and professional inequalities is a recipe for vulnerabilities in research and sharing relationships. Yet, even apart from major research funders in the region that are overt with their sanctions,⁸⁴ mandatory data-sharing is inherently punitive when compliance is non-negotiable. As the under-resourced regroup, rethink, and re-plan additional manuscripts, they would require protections including periods longer than the proposed six months after publication.⁸⁵ Institutions, publishers, and data repositories may further grant priority access and exclusive use during these periods. In the absence of such considerations, losses on data use to more resourced user-scientists are unjustifiable. The pressures and genuine strain on producing scientists should compel flexibilities in mandatory data-sharing periods. This recommendation is especially relevant in minimizing risks following a project's first publication.

Other ways of relaxing the risk of "use or lose" include requirements for data-originators to be notified about other scientists' intentions to use their data. At systems' levels, funders could also consider adding substantial investments into building analytical capacities and infrastructure across academic institutions to ease researchers' efforts and support those for whom analyzing data is problematic.⁸⁶ Incentives could also be made available for private researchers whose data are consciously and consistently made usable. As perceived and real risks are reduced, hesitations to share will decrease.⁸⁷

⁸¹Pisani et al., *op. cit.* note 2, p1 ; Merson et al., *op. cit.* note 64, p.18.

⁸²Merson et al. *op. cit.* note 64, p18; Bezuidenhout, *op. cit.* note 16, p.5

⁸³Emanuel EJ, Wendler D, Killen J, Grady C. (2004). What Makes Clinical Research in Developing Countries Ethical? The Benchmarks of Ethical Research. *Journal of Infectious Disease*. 189(5):930-937.

⁸⁴Pisani et al., *op. cit.* note 6, p.3.

⁸⁵Tiachman et al., *op. cit.* note 4, p2; Pisani et al., *op. cit.* note 2, p.1.

⁸⁶Pisani et al., *op. cit.* note 2, p1; Bezuidenhout, *op. cit.* note 16, p.5.

⁸⁷Bezuidenhout, *Ibid.*

1.9 | Maximizing benefits to scientists who share data

For an endeavor to be ethical, its benefit to risk or cost evaluations must necessarily be positive.⁸⁸ Benefits necessarily form the flip side of risks or costs. The value of data-sharing is therefore justified by its benefits.

We earlier explained data-sharing hesitations that stem from dissatisfaction with having producing and using scientists on two sides of an incentive-dichotomy that disproportionately benefits the latter. We also suggested support for rewards and recognition to be equitably spread across data production through to knowledge creation, regardless of whether the research process is a continuum or separated by different actors. We have also made the case that quality data production already has intellectual properties from contributing scientists that deserve high valuing. These arguments point to the persuasion that sharing data must be accompanied by reward. Need for a benefit model is thus a matter of both Justice and Beneficence (moral obligation to act for the benefit of others).⁸⁹ We therefore suggest the following beneficent options as critical for promoting quality data production: co-authoring opportunities, global recognition, professional promotions, partnering for mutual exchanges of data, cost-sharing, training, and skills strengthening. We wish to emphasize that closing the research-output gap can be helped if quality data that solely supports successful peer-reviewed secondary publications can also be considered as containing adequate intellectual content to justify originating-scientists' authorship status.⁹⁰ Additional contributions may be warranted in accordance with the ICMJE's authorship requirements,⁹¹ but the discussion should at that point be left to a matter of author ranking than possibility. Ethics committees and research institutions should give this recommendation their backing to make data-sharing more attractive for the South.

1.10 | Collaborative Partnership

The complexities of contemporary health issues, the uncertainties surrounding data from unfamiliar contexts, and the need for diversity of ideas necessitate collaboration.⁹² When scientific teams of diverse backgrounds collaborate on research, the quality, quantity, and rigor improves. Team effort, networking, and large scale analysis help build critical pillars for future research. The principle of Collaborative Partnership, especially across the North-South divide is important for data-sharing in the following

⁸⁸Emanuel et al., *op. cit.* note 83, p.24.

⁸⁹Hurst DJ. (2016). Benefit Sharing in a Global Context: Working Towards Solutions for Implementation. *Developing World Bioethics*; doi.org/10.1111/dewb.12118; Beauchamp & Childress, *op. cit.* note 66, p.19.

⁹⁰Anane-Sarpong et al., *op. cit.* note 28, p.8.

⁹¹International Committee of Medical Journal Editors. Uniform requirements for papers submitted to biomedical journals: writing and editing for biomedical publication. Available from: <http://www.icmje.org>.

⁹²Bull et al. *op. cit.* note 18, p.5.

ways: (1) it is effective for diversifying, respecting, and encouraging different types of knowledge and processes of their creation; (2) minimizing “data ownership” claims for smoother and early sharing of data; (3) encouraging the formation of formal and informal sharing networks in which mutual analysis of one another's data can increase scientific productivity (e.g., closed consortia, trust-based networks, and small-scale internally-funded institutional repositories); (4) improving the responsiveness of new analysis to the health needs of communities from whom data originated; and (5) strengthening attachments, impact, and commitment to translate findings into policy and tangible health products. These advantages in-turn strengthen collaboration within and across the sub-region's institutions and scientists as seen in the INDEPTH experience, the Global Health Network which shares research data across many low and middle income countries,⁹³ and several other research facilities for instance in Kenya and South Africa.⁹⁴

Emanuel and colleagues⁹⁵ outline several benchmarks of Collaborative Partnership relevant in application to data-sharing with sub-Saharan Africa. Key among them is the determination of research value, responsibilities, equality in partnership, respect, and benefit sharing. The Council on Health Research for Development's Research Fairness Index⁹⁶ also provides guidelines, tools, checklists, and agreement templates that can complement the implementation of the principle. What remains is for the international regulatory organs to reflect on adapting the relevant provisions into a globally accepted principles-based data-sharing framework. Collaboration in data-sharing will however, not always be possible or practical. Prospective producer-scientists could be controversial or in disagreement with new research plans.⁹⁷ Therefore, while the principle remains largely desirable,⁹⁸ producing-scientists should be left to make good faith efforts to work effectively with user-scientists who express interest in their data. Much however, depends on user-scientists to notify and make opportunities for collaboration available and discoverable.

Collaborative partnership is also strengthened through Reciprocal Justice. This requires setting obligations for reciprocation on the part of user-scientists corresponding to whether data is required to be released or shared in partial or complete forms. In anticipation of unequal intellectual contributions, written agreements are helpful. We strongly recommend that except where the aim of secondary analysis and new manuscript writing from shared data is to challenge original results or where major conflicts of interest exist,

Collaborative Partnership should be promoted and subtly mandated by Journals and data repositories.

1.11 | Transparency

For shared data to be Findable, Accessible, Interoperable, Reusable (FAIR), Equitable, Ethical, and Efficient as advocated by public health research funders in 2010,⁹⁹ the principle of Transparency is necessary. Transparency encapsulates trust and accountability¹⁰⁰ in an intricately woven fashion. Transparency enhances trust in assuring fairness; trust motivates accountability by assuring that data sharers and users take cognizance of each other's risks and benefits; and the process altogether enhances trust and motivates further transparency towards accountability. If data shared can be trusted, scientists who share them must be trustworthy; trustworthy scientists are likely transparent and accountable. A virtuous cycle ensues. Yet, Transparency will likely not come naturally to the culture of freely available data after first publication, especially among unequal scientists. Experiences reported in the literature of researchers' low commitment to research participants following completion of research¹⁰¹ are suggestive that commitments to originating-scientists will generally wane once data is available. This is because commitment generally wanes when the object of attraction is reduced or moved; if data can be gotten without its originator, any commitments to the latter will reduce or vanish. The situation worsens with distance, non-familiarity, and the absence of guidelines.¹⁰² Only international regulation can elicit the kind and scope of adherence needed to remove this challenge.

Overall, in noting that mandatory data-sharing requirements presuppose a certain level of equality towards compliance, transitioning to the new culture should necessarily be guided by further reflections around inequalities, opportunities, privileges, benefits, and incentives. Voices from relevant authorities' in sub-Saharan Africa are critical because of successes and challenges in the region that may not be familiar to the North, where most global scientific actors are based.¹⁰³ Besides, there has been a marked absence of empirical engagement with scientists from the region on data-sharing.¹⁰⁴ Any planned international guidelines will thus benefit from the early involvement of research ethics committees and scientists in the South before the data-sharing rules get established.

⁹³INDEPTH. (2013). The five year strategic plan (2013-2017). Accra: INDEPTH; Pisani et al., *op. cit.* note 2, p.1; Anane-Sarpong et al., *op. cit.* note 28, p.8.

⁹⁴Bezuidenhout, *op. cit.* note 16, p.5.

⁹⁵Emanuel et al. *op. cit.* note 83, p.24

⁹⁶Council on Health Research for Development (COHRED). (2017). Research Fairness Initiative. Geneva. Available from: http://rfi.cohred.org/wp-content/uploads/2017/01/RFI_ReportingGuide_20170112_V2.pdf.

⁹⁷Taichman et al., *op. cit.* note 3, p.2.

⁹⁸Ibid; Dallmeier-Tiessen et al., *op. cit.* note 4, p.2.

⁹⁹Pisani et al., *op. cit.* note 2, p.1; Pisani et al., *op. cit.* note 6, p.3

¹⁰⁰Ibid; Merson et al., *op. cit.* note 64, p.18.

¹⁰¹Purvis R, Abraham T, Long C, Stewart M, Warmack T, McElfish P. (2017). Qualitative Study of Participants' Perceptions and Preferences Regarding Research Dissemination. *American Journal of Bioethics Empirical Bioethics*. 8:69-74; Anane-Sarpong E, Wangmo T, Sankoh O, Tanner M, Elger BS. (2018). Application of Ethical Principles to Research using Public Health Data in the Global South: Perspectives from Africa. *Developing World Bioethics*. 18(2):99-108; Hurst, *op. cit.* note 89, p.26.

¹⁰²Molyneux S, Sofa B, Barasa E et al. (2016). Research Involving Health Providers and Managers: Ethical Issues Faced by Researchers Conducting Diverse Health Policy and Systems Research in Kenya. *Developing World Bioethics*. 16(3):168-177.

¹⁰³Pisani et al., *op. cit.* note 2, p.1.

¹⁰⁴Bezuidenhout, *op. cit.* note 16, p.5.

2 | CONCLUSION

Mandatory data-sharing signifies the future standard for best ethical science and is critical for the growing technological dispensation and the generation of new knowledge. It offers hope for new opportunities, innovations, relationships, and products that can improve health and save lives at minimum costs and optimum speed. Yet, it is clear that the strong global health vision of optimizing data-sharing for the common good alone is not sufficient for good faith adherence. Particularly for sub-Saharan Africa where several generative issues impede the realization of favorable risk-benefit ratios in data-sharing, the culture may not as yet be favorable relative to their expected outcomes. It is our view that the absence of established guidance to correct existing imbalances also makes acceptance, adherence, and promotion difficult. Motivating appetite for data-sharing under unequal circumstances will therefore not come naturally; the change must be spurred by technology, new beliefs and norms, and incentives. It requires transformative steps that are persuasive of increasing scientific productivity from the South, maximizing benefits and minimizing risks, respecting stakeholders, reciprocating effort, encouraging collaboration, and exhibiting transparency. The six ethical principles proposed will help address these by providing protections for the under-resourced scientists in the South; improving the realization of various scientific aspirations and access to technological infrastructure; helping close the global research-output gap and accelerating the South's reach of the SDGs through research. On our part, conducting a future empirical study on the application of the principles would provide additional insights into the discussion to complement this paper.

ACKNOWLEDGMENTS

This paper was written as part of an academic project made possible with support from the Institute for Biomedical Ethics (IBMB) and the Swiss Tropical and Public Health Institute of the University of Basel. Partial funding for the project came from the Basel-Stadt Commission

for Scholarships for Young Professionals from Developing Countries. We are grateful. We wish to also thank the anonymous reviewers of this journal for their critique and comments on the earlier version of the article which has made the final product better.

ORCID

Evelyn Anane-Sarpong  <https://orcid.org/0000-0002-9960-9080>

AUTHOR BIOGRAPHIES

Evelyn Anane-Sarpong is an immediate past student of the Institute for Biomedical Ethics (IBMB), University of Basel, and currently teaches topics in bioethics at the University of Cape Coast, Ghana. Her research interests include exploring social and professional relations underlying the ethics of conducting health research in Africa.

Tenzin Wangmo is a senior researcher at the Institute for Biomedical Ethics (IBMB), University of Basel. She holds a PhD and has research interests in bioethics and the vulnerable, with special focus on the health of older prisoners and the ethical care of older adults in medical institutions.

Marcel Tanner is President of the Swiss Academy of Sciences and Professor emeritus of Epidemiology and Medical Parasitology at the University of Basel. His research, teaching, and health planning are based on long term work in sub-Saharan Africa and Asia.

How to cite this article: Anane-Sarpong E, Wangmo T, Tanner M. Ethical principles for promoting health research data sharing with sub-Saharan Africa. *Developing World Bioeth.* 2020;20:86–95. <https://doi.org/10.1111/dewb.12233>