

October 2016

Ensuring global equity in open research

Susan Bull



Cite this as: Bull, Susan (2016) Ensuring global equity in open research. Wellcome Trust.
<https://dx.doi.org/10.6084/m9.figshare.4055181>

Ensuring global equity in open research

Executive summary

Sharing research outputs beyond primary research teams requires a significant investment of resources to establish and maintain data curation standards, methods and infrastructure, and appropriate governance policies and processes. Policies mandating such sharing are becoming increasingly common, despite the costs involved. Reasons for promoting the sharing of research outputs include respecting the contributions of multiple stakeholders who directly and indirectly support research, by maximising the value and utility of their contributions. Specifically, it has been claimed that sharing can contribute to improving our scientific understanding of health and disease, and inform improvements in healthcare and the health of populations.

A lack of resources for appropriate data sharing has been identified as a hindrance in high income settings, and a very significant barrier in low and middle income settings. In such environments, initiatives including WWARN, H3Africa, INDEPTH, the ALPHA Network and MalariaGEN demonstrate that equitable sharing can be achieved, following a considerable investment in human resources, technology and infrastructure, for the curation and sustainable sharing of research outputs. Given the potential value of sharing research outputs, care is needed to determine how best to resource equitable and sustainable sharing on a broader scale within low and middle income settings.

To avoid exacerbating existing inequalities between higher and lower income settings, capacity to curate and share research outputs must be developed in conjunction with policies and processes to promote equitable sharing. Specialist expertise is needed to inform the development of policies and processes that promote transparent and equitable data sharing, address the rights and responsibilities of both data providers and data recipients, and provide for appropriate benefit sharing and management of intellectual property rights for datasets, algorithms and software. To inform the development of such policies, the interests of multiple research stakeholders must be recognised and balanced. The interests of two important groups of stakeholders, the primary researchers who develop research outputs, and the participants and communities which contribute to them, are discussed in more detail below.

Research outputs represent a very significant investment of time and effort on the part of primary researchers. Primary researchers' interests in conducting initial analyses of their research findings, as well as in receiving appropriate recognition and credit in secondary analyses of their data, have been widely recognised in higher and lower income settings. It has also been considered inequitable to develop researchers' capacity to share research outputs from low and middle income settings without also developing their capacity to benefit from sharing their research outputs and to analyse relevant datasets shared by others.

Participants and communities involved in research also have interests sharing in benefits arising from the sharing of research outputs. While there is no consensus that direct benefits are a requirement of ethical data sharing, stakeholders have discussed the importance of secondary research providing indirect benefits, by contributing to the knowledge base needed to address health issues of relevance to their communities and country. Where secondary research does not address issues of direct relevance to communities involved in primary research, it is considered important that it should have the potential to advance health more generally. The limited empirical data in this field suggest that while those communities contributing to research in low and middle income settings have an interest in data sharing contributing to scientific knowledge, their primary interest is in the translation of such knowledge into tangible health improvements for their communities and others. The extent to which sharing research outputs is translated into improvements in healthcare is discussed further below.

Gaps in the data

Most arguments in favour of sharing research outputs appeal to empirical claims that such sharing will promote science and contribute to advances in healthcare. Such claims can be tested and evaluated by tracking uses made of shared research outputs, by reviewing the outcomes of secondary analyses, and determining the extent to which such outcomes have influenced healthcare provision. Given the relative novelty of sharing health research outputs in low and middle income settings, there is very limited empirical evidence to date about the extent to which such practices have promoted science and advanced healthcare, although there is evidence that datasets, such as those made available by INDEPTH, have been extensively accessed by secondary researchers. Given the key role of potential advances in science and healthcare both in arguments in favour of data sharing, and in participant and community motivations for agreeing to share data, it is important to address this evidence gap.

Correspondingly, there is relatively little evidence that policies promulgating the sharing of research outputs have resulted in harm to the research enterprise in low and middle income settings. There is anecdotal evidence of ethics review boards in low and middle income settings declining to approve research protocols where research output sharing was an integral component of a study. There are also instances of data being submitted to repositories without the metadata required to promote their utility, so that researchers could satisfy data sharing requirements in form if not in substance.

As the sharing of research outputs becomes increasingly commonplace, empirical research is required to determine if researchers' concerns about sharing potentially adversely affecting their research capacity and career development are borne out in practice, and if so, how such effects can be addressed. Correspondingly, research should be conducted to determine if sharing results in any harms to participants and communities participating in research, for example by leading to stigmatisation. Empirical research is also needed to evaluate the extent and nature of any adverse effects that policies requiring the sharing of research outputs may have had on researchers' ability or willingness to conduct primary research.

A third important area in which there is little empirical evidence at present, relates to identifying which approaches to data sharing are currently being employed in low and middle income settings, their resource implications, and effectiveness in terms of promoting equitable use of shared resources.

Potential options and approaches for research funders

A research agenda

Funders have an important role in supporting the development, implementation and evaluation of models for sharing the outputs of research. In response to the gaps identified above, it may be valuable for funders **to support research to inform the development of models of sharing which promote global equity, including research:**

- **which maps and evaluates emerging practices for sharing research outputs - including assessing their impacts on levels of scientific activity and the outcomes of that activity in terms of career advancement and health promotion, the nature and extent of any harms that have resulted from sharing research outputs, and how these are distributed amongst stakeholders;**
- **to inform the development of models of sharing which aim to promote equity - including assessing the effects of policy measures such as the use of exclusive access periods, publishing embargoes, and requirements that data access requests include capacity building components.**

Supporting the development of policies and practices for equitable sharing

Researchers in low and middle income settings have requested clear guidance from funders about appropriate policies and practices for sharing outputs. It would be valuable for funders to **resource the development of, and promote access to, training programmes, exemplars and templates to build capacity in low and middle income settings to develop policies and practices for the equitable sharing of research outputs which are locally appropriate and based on recognised good practices.**

Resourcing equitable sharing

Care is needed to ensure that the development of capacity for sharing research outputs does not adversely impact capacity to conduct relevant primary research, or to provide healthcare in low and middle income settings. Funders should **support the development of practices that meet the standards required for effective research output sharing, recognising that such standards are often above and beyond the data curation and analyses needed to address the primary research questions in a specific study.**

Development and implementation of data standards, metadata and interoperability

Large scale research consortia and collaborators addressing specific health conditions are developing standards and resources for the collection of data relevant to specific conditions, and for the subsequent sharing of relevant research outputs. Funders may promote equitable output sharing by **ensuring that researchers in low and middle income settings have an equal voice in discussions about the development and implementation of such standards and have access to the resources required to implement them.**

Capacity to access and use research outputs

Equity questions in output sharing arise not only in relation to capacity to share high quality curated outputs but also in relation to the capacity to make best use of others' research outputs. Funders should consider means of promoting the capacity of researchers in low and middle income settings to **identify, access and analyse shared research outputs to address both locally relevant and global research questions, and to develop equitable collaborations for such analyses where appropriate.**

Ensuring global equity in open research

Sharing data is not only about the technical dimension such as data management, repositories and libraries; developing countries are concerned about factors that impede data sharing, in particular, fairness. (Tangcharoensathien, Boonperm, and Jongudomsuk 2010)

Introduction

Policies mandating the sharing of health research outputs, including research data and coding, are becoming increasingly common and commanding widespread support. Practical instrumental rationales for such sharing are that it can maximise the value and utility of datasets, leading to scientific developments which can, in turn, advance health (see Annex A). Ethical principles of fairness and reciprocity have been considered to require that research outputs be shared to benefit communities that directly and indirectly contribute to them (Langat et al. 2011; Strech and Littmann 2012; Tangcharoensathien, Boonperm, and Jongudomsuk 2010). The principle of respect for persons has been interpreted to require that participants' contributions to research be maximised by making the best use of their data, to honour their expectations that results of research will be disseminated to advance scientific development. (Gotzsche 2011; Mello et al. 2013) In low and middle income settings commentators have suggested that developing appropriate responses to some complex adverse health outcomes may not be possible without the sharing of research outputs and involvement of multiple research groups (National Academies of Sciences 2015).

Significant concerns about sharing research outputs have also been raised, including the potential for sharing to hamper scientific development and healthcare provision, and to adversely affect the interests of stakeholders in research (see Annex A). Consequently there has been significant discussion in the literature about the need to develop and implement appropriate policies and practices for sharing research outputs (Eichler et al. 2013; Mello et al. 2013; Zarin 2013; Merson, Gaye, and Guerin 2016). In particular, the need for approaches to sharing to ensure the appropriate balancing of potential benefits and potential harms has been discussed (National Academies of Sciences 2015). Research funders have described equitable sharing as requiring the recognition and balancing of multiple stakeholders' interests including 'the needs of researchers who generate and use data, other analysts who might want to reuse those data, and communities and funders who expect health benefits to arise from research'.(Walport and Brest 2011).

The report below addresses discussions in the literature about promoting equity in data sharing, challenges that have been identified and potential ways forward. It focuses on sharing outputs of health-related research, including individual level datasets, particularly in low and middle income settings.

Background

The routine pre-publication sharing of health research data beyond established collaborative relationships is a relatively novel phenomenon. In the context of biomedical research, much of the drive to share such data has its origins in the Human Genome Project and in the subsequent development of genomics and research methods such as the genome-wide association study (GWAS). Open access models of genomic data release, drawing on the Bermuda Principles and the Fort Lauderdale agreement, are now commonplace (Human Genome Organisation 1996, 1997; The Wellcome Trust 2003). A recognition of the potential advantages of sharing data, accompanied by significant investment and developments in the infrastructure required to support such sharing, has led to policies mandating the sharing of data and other outputs from biomedical and public health research becoming increasingly widespread (OECD 2007; Toronto International Data Release Workshop et al. 2009; Wellcome Trust 2009; Higher Education Funding Council for England et al. 2016).

While research suggests that clinical researchers publishing in high impact journals increasingly support and undertake sharing of individual-level data with non-collaborating researchers (Rathi et al. 2012), such practices are still relatively uncommon in low and middle income settings where policies and infrastructure to support systematic data archiving and sharing are not well established (Rani, Bekedam, and Buckley 2011). When seeking to develop effective, efficient and equitable practices for sharing research outputs in low and middle income settings, multiple issues need to be addressed (Alter and Vardigan 2015; Carr and Littler 2015; Lötter and van Zyl 2015; Rani, Bekedam, and Buckley 2011; Sankoh and Ijsselmuiden 2011; Tangcharoensathien, Boonperm, and Jongudomsuk 2010; Walport and Brest 2011; Whitworth 2010; Pisani et al. 2010). Specifically, in addition to a considerable investment in building capacity in human resources, technology and infrastructure, models of good sharing practices capable of commanding the trust and confidence of relevant stakeholders are required. Well-resourced research initiatives sharing research outputs from low and middle income settings have demonstrated that these requirements can be met (Parker et al. 2009; de Vries et al. 2011; Herbst et al. 2015; H3Africa Consortium 2014; de Vries et al. 2015). The challenge is now to determine how best to support equitable open research that is responsive to the interest of multiple stakeholders on a broader scale in low and middle income settings.

Capacity building and equitable research

At its core, equitable sharing of research outputs requires the acquisition and maintenance of multiple interrelated capacities. The discussion below begins by focusing on the specific interests and capacity development needs of stakeholders who are involved in the production of research outputs that will be shared. It then addresses the capacities required to support governance policies and processes that are responsive to stakeholder interests and promote equitable sharing.

Stakeholder interests and capacity development needs

In discussions of equitable data sharing, particular attention has been paid to the interests of two groups of stakeholders in low and middle income settings, the researchers who generate research outputs, and the participants and communities that contribute directly to them. These are considered in turn below.

Primary researchers

The importance of respecting the interests of primary researchers producing research outputs to be shared has received considerable attention in the literature (Castellani 2013; Lopez 2010; Pisani and AbouZahr 2010; Tangcharoensathien, Boonperm, and Jongudomsuk 2010; Whitworth 2010). Research outputs are viewed as valuable resources that are collected and managed only with considerable effort. Stakeholders from higher and lower income settings have commented on the importance of ensuring that researchers have the opportunity to conduct and publish primary analyses of their research and to receive recognition for producing datasets in the subsequent publications by secondary analysts, in professional assessments, and in funding applications (Kuntz 2013; Manju and Buckley 2012; Piwowar et al. 2008; Rani, Bekedam, and Buckley 2011; Rathi et al. 2012). Researchers in low and middle income settings have also discussed the importance of having the capacity to curate datasets in a way that maximises their utility, and minimises the possibilities of flawed secondary analyses and of critiques being made of the primary research. Commentators have additionally noted that it would be unfair to develop capacity to share research outputs in low and middle income settings without also developing the capacity for the primary researchers who generated them to analyse relevant research outputs (Pisani et al., 2010a; Sankoh & Ijsselmuiden, 2011; Walport & Brest, 2011; Whitworth, 2010). Each of these points are considered in turn below.

Periods of exclusive access to research outputs and publication moratoria

Delays in the release of research data are not universally implemented and the value of releasing data prior to publication has been widely recognised (Toronto International Data Release Workshop Authors, 2009). However, many commentators have noted the value of exclusive fair use periods for researchers in both higher and lower income settings (Geller et al., 2004; Gotzsche, 2011b; Manju & Buckley, 2012; Pearce & Smith, 2011; Pisani & AbouZahr, 2010; Pisani et al., 2010a; Rathi et al., 2012; Ross, Lehman, & Gross, 2012; Savage & Vickers, 2009; Tangcharoensathien et al., 2010; Vickers, 2006). At present there are significant variations in the length of time before research data are, ranging from set periods such as 12 months from the end of data collection, or nine months from the dataset completing quality control, to unspecified lengths of time, which are, in some cases, linked to the publication of an article with primary findings. Where research outputs are released prior to publication there may be additional publication moratorium periods imposed, which are often for set periods of up to 9-12 months. (H3Africa 2014, July 2012.)

Particular concerns have been expressed about the abilities of some researchers in low and middle income settings to effectively compete with more highly resourced secondary researchers to publish initial and subsequent analyses of data. Exclusive fair use periods have been implemented in consortia such as MalariaGEN and H3Africa explicitly to 'assist in

balancing the significant differences in analytic capacity present in developed and developing countries' (Parker et al. 2009).

While such initiatives support researchers' abilities to publish primary analyses, journals are increasingly requiring the data underpinning research papers to be made available for sharing at the time of publication. Researchers in low and middle income settings have raised concerns that they may not have the opportunity to publish subsequent analyses, despite requesting that their future analysis and publication plans be respected. For example, primary researchers may have limited capacity to monitor secondary uses made of research outputs once released. Additionally, in competitive resource-limited environments where there are few mechanisms to promote research integrity, and significant mistrust amongst researchers, concerns arise that publication embargoes may not be respected. (Hate et al. 2015)

Recognition for the generation of research outputs

Researchers in low and middle income settings have reported experiences of receiving neither co-authorship nor the acknowledgement that was expected following secondary analyses of their data (Cheah et al. 2015). The need for those releasing datasets to receive appropriate recognition has been discussed widely in the literature, and calls have been made for such contributions to be appropriately recognised in secondary analyses, reviews of professional development and funding applications (Castellani 2013; Lopez 2010; Tangcharoensathien, Boonperm, and Jongudomsuk 2010; Whitworth 2010). Co-authorship on secondary analyses has been discussed as a method of explicitly acknowledging the scientific contribution of primary producers of datasets (Pearce and Smith 2011; Savage and Vickers 2009; Merson et al. 2015). However, the contribution of data creators may not be sufficient to meet journal standards in the absence of a collaborative analysis and significant contributions to a publication (Anderson and Merry 2009; Expert Advisory Group on Data Access 2015). Funders may agree on alternative forms of acknowledgement of the contribution of primary researchers and data curators that should be taken into account in funding applications. However, professional development reviews in low and middle income settings that currently focus on authorship may be slow to change, limiting incentives to devote resources to the curation and release of research outputs.

Building capacity to analyse and share research resources

Fair trade is usually governed by mutually binding agreements and measures. This implies achieving a balance between the rights and responsibilities of those who generate data and those who analyse and publish results using those data. Such a balance lies in ensuring that the means and capacity to share and actively participate in the analysis of those data are in the hands of those who generate the data and not only in those who want to analyse it. (Sankoh and Ijsselmuiden 2011)

The sharing of high quality research outputs in sustainable repositories requires significant investment in human resources, technology, and infrastructure. The need for training, mentoring, and career pathways for primary researchers and specialist support staff who curate research outputs for release are well recognised (Pisani et al. 2010). Where research outputs are shared from repositories hosted within low and middle income settings,

additional expertise is needed to sustain the resource and appropriately manage release. Supplementary resources are required to monitor secondary uses of research data, respond to queries from data accessors, and evaluate results of secondary analyses. Researchers with limited experience of sharing research outputs may find it challenging to accurately identify and develop budgets for the resources needed to support data management and equitable release. Funders have been enjoined to establish the costs of sharing research outputs both in the short term, and in relation to sustainable sharing beyond the duration of individual grants (Expert Advisory Group on Data Access 2015).

Equitable output sharing requires that in addition to building capacity to curate and release research outputs, researchers in low and middle income settings should have the capacity and resources to conduct high quality analyses of their research data (de Vries et al. 2015; Parker et al. 2009). Experience and expertise in analysis may be developed with appropriate training, mentorship and participation in collaborative research. Researchers in low and middle income settings may also require resources to address technological barriers to analyses, including out of date computing facilities with insufficient processing power, a lack of access to relevant proprietary analysis software and insufficient internet access and data plans to allow data to be uploaded or downloaded. Researchers with less experience of research design, data curation and analyses, have raised concerns that in the absence of appropriate capacity development, sharing research outputs may lead to reputational damage. In particular, experienced and well-resourced secondary users may critique both the quality of the dataset and the primary research, affecting primary researchers' reputations, ability to attract research funding, and career development (Bull et al. 2015).

In addition to developing capacities to conduct analyses with their primary datasets, researchers in low and middle income settings should ideally have similar opportunities to researchers in high income settings to discover and analyse outputs shared by colleagues working in relevant areas, and to initiate and participate in research collaborations where research outputs will be shared (Committee on Data for Science and Technology of the International Council for Science 2014). While ensuring that research outputs are available to secondary users in low and middle income settings is an important first step, additional resources and expertise will often be required for those outputs to be analysed effectively (Bezuidenhout et al. 2016).

Research participants and their communities

A second core group of stakeholders with interests in the sharing of research outputs are research participants and the communities from which they are drawn. Potential benefits and harms that may accrue to participants and communities, and means of equitably distributing these, are discussed below.

Potential harms to research participants and communities

Research participants and the communities from which they are drawn have multiple interests which could potentially be adversely affected by inequitable sharing which does not seek to minimise potential harms. In addition to seeking to protect participants' privacy, it is important that processes for sharing data can be responsive to the context in which

data are collected, in order to identify datasets and potential uses of data that are likely to be sensitive, and to implement appropriate protections.

Equitable access to the benefits of research

When community stakeholders in low and middle income settings have discussed the possibility of sharing the outputs of research to which they have contributed, there was no general consensus that direct benefits to research participants were a requirement of ethical data sharing (Bull et al. 2015). However, they discussed the importance of secondary research providing indirect benefits, such as addressing health issues of relevance to their communities. If such issues were not to be addressed, then it was considered important that secondary research should have the potential to advance health more generally, a view echoed in higher income settings (Bull et al. 2015; Ipsos MORI 2016).

Commentators have noted that benefits of sharing research outputs may not always be realised in low and middle income settings. In one example, the World Health Organization's sharing of avian flu virus specimens led to the production of vaccines that were unaffordable in lower income countries that could be affected by a flu pandemic (Tangcharoensathien, Boonperm, and Jongudomsuk 2010) Research consortia have sought to encourage equitable sharing of results of secondary analyses that have the potential to promote health. For example, the MalariaGEN data access agreement states that if the results of secondary analyses could lead to affordable health solutions in low and middle income settings, data accessors must agree to:

- offer non-exclusive licenses to such results on a reasonable basis for use in low income and low-middle income countries
- on request, provide a non-exclusive, royalty-free licence with the right to sub-license to the Foundation of the National Institutes of Health, USA solely for uses in low income and low-middle income countries
- provide preferential access to such results to the countries that contributed the samples from which the data in the MalariaGEN database is derived (MalariaGEN 2008).

Similarly, when reviewing applications for biospecimens the H3Africa Data and Biospecimen Access Committee will assess the potential for research to be published, lead to patents, or aid in discovery and development of new therapies (H3Africa Consortium 2015).

To support appropriate benefit sharing and management of intellectual property, primary researchers and organisations in low and middle income settings will require the power and capacity to negotiate equitable agreements. Resources such as Chatham House's *Model agreement for the sharing of public health surveillance data beyond national borders (in press)* and exemplars of data sharing policies, data access agreements and material transfer agreements developed by funders and research consortia can provide useful resources to inform such negotiations.

Potential harms to the research enterprise

Concerns have been raised that requirements for data sharing could have an adverse effect on the conduct of primary research. Preparing research outputs for sharing requires human, financial and infrastructure resources which may then not be available for relevant primary research in a setting. Policies mandating sharing of research outputs could result in participants not taking part in studies, ethics committees not giving approval, and other studies not having local resources to be conducted.

The need for appropriate metadata to accompany research outputs is well recognised (Merson, Gaye, and Guerin 2016; de Vries et al. 2014; Committee on Data for Science and Technology of the International Council for Science 2014). When researchers in low and middle income settings are insufficiently resourced to curate, share and analyse their data, they may have little incentive, or choice, but to release raw data without appropriate metadata, to comply with data release requirements associated with their research funding or with journal publication. When research outputs are released without appropriate metadata, or there are insufficient resources to respond to secondary users' queries, there is a high risk that they will be of little utility, or resources might be wasted in secondary analyses which are then unable to generate valid answers to the research question.

Governance processes for sharing research outputs

The final section of this report focuses on the policies and processes designed to promote the equitable and ethical sharing of research outputs. In low and middle income settings, capacity is needed to develop, implement and maintain good governance processes, which are key to promoting equitable sharing of research outputs. Researchers in higher and lower income settings have emphasised the importance of data sharing processes being accountable, efficient, fair and proportionate. Additionally the importance of developing policies to inform good governance processes has been widely recognised (Bull et al. 2015; Bull, Roberts, and Parker 2015). Funders are now routinely requiring explicit data sharing and management plans to be incorporated into grant applications.

To promote equitable research output sharing in low and middle income settings, it is important that examples of existing good governance practice are readily available, and resources are available to support the appropriate tailoring of these to specific contexts. Consortia and research networks focusing on research data from low and middle income settings, including INDEPTH, MalariaGEN, H3Africa, WWARN and the Mahidol Oxford Tropical Medicine Research Unit have developed and publicised policies and processes for curating and sharing research outputs which have been developed in consultation with a wide range of relevant stakeholders (<http://www.indepth-ishare.org/index.php/home>; <https://www.malariagen.net/data/our-approach-sharing-data>; <http://h3africa.org/consortium/documents>; <http://www.wwarn.org/working-together/sharing-data>; <http://www.tropmedres.ac/data-sharing>). Appropriate policies and processes can also support proportionate review, distinguishing research outputs which may be released with minimal oversight, from those where a more substantial curation process is required to appropriately manage potential benefits and harms (Hrynaszkiewicz

and Altman 2009; Vallance and Chalmers 2013; Toronto International Data Release Workshop et al. 2009).

When research outputs requiring higher levels of curation are being released, views differ about the most appropriate mechanisms for curation to manage potential harms and benefits. Where resources are limited, questions arise, for example, about whether capacity should be built so that existing research ethics committees can form part of the research output governance process in low and middle income settings, or whether establishing dedicated data access committees may be a more appropriate mechanism to protect stakeholders' interests (Bull et al. 2015). Funders have also been recommended to consider whether consolidation of data access committees may be appropriate in some contexts. Doing so could consolidate expertise and assist to ensure sustained access to research outputs independently of funding cycles (Expert Advisory Group on Data Access 2015). Irrespective of the curation method chosen, capacity building, to enable the promotion of equitable sharing is necessary.

Research ethics committees play an important role in ensuring that sharing is equitable and ethical, whether they are involved solely in reviewing the primary protocol from which research outputs are to be shared, or additionally in sharing curation processes. Unfortunately, collaborative studies where data sharing is planned often raise ethical concerns for which research ethics committees have limited guidance (Dove et al. 2016). Resources have been developed to provide training on ethics and best practices in data sharing to research ethics committees in low and middle income settings (<https://globalhealthtrainingcentre.tghn.org/ethics-and-best-practices-sharing-individual-level-data-clinical-and-public-health-research/>). Further training, and the opportunity for ethics committees to discuss data sharing issues in regional meetings, is desirable.

Effective governance processes for research output data sharing have multiple potential benefits, including increasing patient privacy, minimising the likelihood of poor quality secondary research, promoting compliance with legislation and regulation, promoting researchers' abilities to adhere to commitments made during recruitment and fulfil their responsibilities to ensure data are used ethically (Bull, Roberts, and Parker 2015). Governance processes can contribute to promoting the interests of primary researchers, and safeguarding the interests of research participants and the communities from which they come, preventing an inequitable distribution of harms and benefits. Of particular importance are the need to ensure consent processes are appropriate, and the need to protect research participants from potential harms of research, including re-identification, and stigmatisation of individuals, communities and countries (Cheah et al. 2015; Denny et al. 2015; Hate et al. 2015; Jao, Kombe, Mwalukore, Bull, Parker, D, et al. 2015; Jao, Kombe, Mwalukore, Bull, Parker, Kamuya, et al. 2015; Merson et al. 2015).

Although views differ about which stakeholder interests should be represented in governance processes to promote equitable sharing, the role of local representation is often acknowledged (Bull et al. 2015). The *H3Africa Data and Biospecimen Access Committee Guidelines*, for example, advocate that the committee be comprised primarily of people working within African countries (H3Africa Consortium 2015). Such representation aims to

promote trust and counter potential concerns about exploitation, and to ensure that secondary uses are not offensive to African populations and are relevant to their health needs.

Collaboration

Scientists from developed countries often take the following approach with researchers in developing countries: “Share your data with me, you do not have analytical capacities. I will analyse and publish papers for global public good.” Instead, their approach should be: “We can analyse the data together and learn from each other for the benefit of all people.” This approach would gradually create equal partnerships, a level playing field, goodwill and trust for collaborations beyond simply sharing data. (Tangcharoensathien, Boonperm, and Jongudomsuk 2010)

The advantages of sharing clinical and public health data through collaborative data-sharing arrangements in a range of low- and middle-income settings have been recognized in the literature (Manju and Buckley 2012; Tangcharoensathien, Boonperm, and Jongudomsuk 2010; Whitworth 2010). When discussing preferred approaches to sharing research outputs, researchers in multiple low and middle income settings have expressed a strong preference for collaborative approaches to data sharing, irrespective of the presence of appropriate governance processes (Cheah et al. 2015; Denny et al. 2015; Hate et al. 2015; Jao, Kombe, Mwalukore, Bull, Parker, D, et al. 2015; Merson et al. 2015). Collaborative research output sharing was the most familiar form of sharing for most researchers, and considered to be a key method of minimising risks of sharing and promoting participants’ and communities’ interests. Sharing data within collaborations was thought to support trust building and capacity development, and to increase the benefits that primary researchers would receive in terms of acknowledgment and authorship. Secondary analyses would be improved as relevant contextual elements of the research outputs could be discussed and any errors or misunderstandings addressed. Collaboration also has the potential to increase the relevance of secondary research to health needs of participants and communities, and increase the likelihood that findings with implications for healthcare provision could be translated into practice.

However, requiring that secondary data analyses can only be conducted within collaborative relationships has the potential to significantly limit the utility of datasets. Primary researchers can only be substantially involved in a finite number of collaborative studies, which limits the nature and range of potentially valuable secondary uses of research outputs. The Expert Advisory Group on Data Access has recommended that ‘funders should be clear that using collaboration with the study team as the sole means through which to allow data to be accessed is not appropriate, other than in exceptional circumstances in which it can be justified’. (EAGDA 2015) However, given the strong support for, and multiple perceived advantages of, collaborative data sharing, it is important that this issue be given careful consideration when developing policies and processes for data sharing in low and middle income settings.

Concluding thoughts

Sharing the outputs of research offers the potential to address complex health challenges. Ground-breaking consortia and research units focused on addressing such challenges in low and middle income settings have demonstrated that their research outputs can be equitably released and can contribute to scientific knowledge and health policy. Funders and advocates of open research now have the challenge of promoting the equitable sharing of research outputs on a broader scale. Empirical data from low and middle income settings suggest that a key element in the success of such endeavours will be the development of trustworthy and trusted policies and processes for such sharing (Bull et al. 2015).

Acknowledgements

My sincere thanks to colleagues for their review of the executive summary in this report within a very tight timeframe: Professor Michael Parker, Dr Jantina de Vries, Associate Professor Vicki Marsh, Professor David Osrin, Professor Douglas Wassenaar, Laura Merson, Dr Angeliki Kerasidou and Dr Louise Bezuidenhout. Their very helpful comments have been incorporated, however this report does not purport to represent their views.

Susan Bull,
Ethox Centre
Nuffield Department of Population Health
University of Oxford

References

- Alter, G.C., and M. Vardigan. 2015. 'Addressing global data sharing challenges ', *Journal of Empirical Research on Human Research Ethics*, 10: 317-23.
- Anderson, Brian J., and Alan F. Merry. 2009. 'Data sharing for pharmacokinetic studies', *Paediatric Anaesthesia*, 19: 1005-10.
- Bezuidenhout, L. M., S. Leonelli, A. H. Kelly, and B. Rappert. 2016. 'Beyond the digital divide: Towards a situated approach to open data', *Science and Public Policy*, In press.
- Bull, S, P. Y. Cheah, S. Denny, I Jao, V. Marsh, L Merson, L.T.N. Nhan, D Osrin, N Shah More, D Tangseefa, D Wassenaar, and M Parker. 2015. 'Ethics and best practices for sharing individual-level health research data from low and middle income settings', *Journal of Empirical Research on Human Research Ethics*, 10: 302-13.
- Bull, S, N Roberts, and M Parker. 2015. 'Stakeholder perspectives on sharing de-identified individual-level data from medical and public health research: a scoping review', *Journal of Empirical Research on Human Research Ethics*, 10: 225-38.
- Carr, D. , and K. Littler. 2015. 'Sharing research data to improve public health: A funder perspective', *Journal of Empirical Research on Human Research Ethics*, 10: 314-16.
- Castellani, J. 2013. 'Are clinical trial data shared sufficiently today? Yes', *BMJ (Online)*, 347 (7916).
- Cheah, PY, D Tangseefa, A Somsaman, T Chunsuttiwat, F Nosten, NPJ Day, S Bull, and M Parker. 2015. 'Perceived benefits, harms and views about how to share data responsibly: a qualitative study of experiences with and attitudes towards data-sharing among research staff and community representatives in Thailand', *Journal of Empirical Research on Human Research Ethics*, 10: 278-89.
- Committee on Data for Science and Technology of the International Council for Science. 2014. "Data sharing principles in developing countries (the Nairobi Data Sharing Principles)." In.: Committee on Data for Science and Technology of the International Council for Science,.
- de Vries, J, Susan J. Bull, Ogobara Doumbo, Muntaser Ibrahim, Odile Mercereau-Puijalon, Dominic Kwiatkowski, and Michael Parker. 2011. 'Ethical issues in human genomics research in developing countries', *BMC Medical Ethics*, 12: 5.
- de Vries, J., P. Tindana, K. Littler, M. Ramsay, C. Rotimi, A. Abayomi, N. Mulder, and B. M. Mayosi. 2015. 'The H3Africa policy framework: negotiating fairness in genomics', *Trends Genet*, 31: 117-9.
- de Vries, J., T. N. Williams, K. Bojang, D. P. Kwiatkowski, R. Fitzpatrick, and M. Parker. 2014. 'Knowing who to trust: exploring the role of 'ethical metadata' in mediating risk of harm in collaborative genomics research in Africa', *BMC Med Ethics*, 15: 62.
- Denny, S.G., B. Silaigwana, D. Wassenaar, S.J. Bull, and M. Parker. 2015. 'Developing ethical practises for public health research data sharing in South Africa: The views and experiences from a diverse sample of research stakeholders', *Journal of Empirical Research on Human Research Ethics*, 10: 290-301.
- Dove, E. S., D. Townend, E. M. Meslin, M. Bobrow, K. Littler, D. Nicol, J. de Vries, A. Junker, C. Garattini, J. Bovenberg, M. Shabani, E. Levesque, and B. M. Knoppers. 2016. 'RESEARCH ETHICS. Ethics review for international data-intensive research', *Science*, 351: 1399-400.
- Eichler, Hans-Georg, Frank Petavy, Francesco Pignatti, and Guido Rasi. 2013. 'Access to patient-level trial data--a boon to drug developers', *New England Journal of Medicine*, 369: 1577-9.
- Expert Advisory Group on Data Access. 2015. "Governance of data access." London: Expert Advisory Group on Data Access.
- Gotzsche, Peter C. 2011. 'Why we need easy access to all data from all clinical trials and how to accomplish it', *Trials*, 12: 249.
- H3Africa Consortium. 2014. 'Research capacity. Enabling the genomic revolution in Africa', *Science*, 344: 1346-8.
- H3Africa Consortium 2015. "Data and Biospecimen Access Committee Guidelines." H3Africa Consortium

- Hate, K, S Meherally, N Shah More, A Jayaraman, S Bull, M Parker, and D Osrin. 2015. 'Sweat, skepticism and uncharted territory: A qualitative study of opinions on data sharing among public health researchers and research participants in Mumbai, India', *Journal of Empirical Research on Human Research Ethics*, 10: 239-50.
- Herbst, K., S. Juvekar, T. Bhattacharjee, M. Bangha, N. Patharia, T. Tei, and O. Sankoh. 2015. 'The INDEPTH data repository: an international resource for African and Asian researchers', *Journal of Empirical Research on Human Research Ethics*, 10: 324-33.
- Higher Education Funding Council for England, Research Councils UK, Universities UK, and Wellcome. 2016. "Concordat on Open Research Data." Higher Education Funding Council for England, Research Councils UK, Universities UK, and Wellcome
- Hrynaszkiewicz, Iain, and Douglas G. Altman. 2009. 'Towards agreement on best practice for publishing raw clinical trial data', *Trials*, 10: 17.
- Human Genome Organisation. 1996. "Summary of Principles Agreed Upon at the First International Strategy Meeting on Human Genome Sequencing (Bermuda, 25-28 February 1996) as reported by HUGO." Human Genome Organisation
- Human Genome Organisation 1997. "Summary of the Report of the Second International Strategy Meeting on Human Genome Sequencing (Bermuda, 27th February - 2nd March, 1997) as reported by HUGO." Human Genome Organisation
- Ipsos MORI. 2016. "The One-Way Mirror: Public attitudes to commercial access to health data: report prepared for the Wellcome Trust." Ipsos MORI.
- Jao, I, F Kombe, S Mwalukore, S Bull, M Parker, D Kamuya, S Molyneux, and V Marsh. 2015. 'Research Stakeholders' Views on Benefits and Challenges for Public Health Research Data Sharing in Kenya: The Importance of Trust and Social Relations', *PLoS ONE*, DOI: 10.1371/journal.pone.0135545.
- Jao, I, F. Kombe, S Mwalukore, S Bull, M Parker, Kamuya D, S. Molyneux, and V. Marsh. 2015. 'Involving research stakeholders in developing policy on sharing public health research data in Kenya: Views on fair process for informed consent, access oversight and community engagement', *Journal of Empirical Research on Human Research Ethics*, 10: 264-77.
- Kuntz, Richard E. 2013. 'The changing structure of industry-sponsored clinical research: pioneering data sharing and transparency', *Annals of Internal Medicine*, 158: 914-5.
- Langat, P., D. Pisartchik, D. Silva, C. Bernard, K. Olsen, M. Smith, S. Sahni, and R. Upshur. 2011. 'Is there a duty to share? Ethics of sharing research data in the context of public health emergencies', *Public Health Ethics*, 4: 4-11.
- Lopez, A. D. 2010. 'Sharing data for public health: where is the vision?', *Bulletin of the World Health Organization*, 88: 467.
- Lötter, L., and C. van Zyl. 2015. 'A reflection on a data curation journey', *Journal of Empirical Research on Human Research Ethics*, 10: 338-43.
- MalariaGEN. 2008. "MalariaGEN Data Access Agreement." MalariaGEN
- Manju, Rani, and B. S. Buckley. 2012. 'Systematic archiving and access to health research data: rationale, current status and way forward', *Bulletin of the World Health Organization*, 90: 932-39.
- Mello, Michelle M., Jeffrey K. Francer, Marc Wilenzick, Patricia Teden, Barbara E. Bierer, and Mark Barnes. 2013. 'Preparing for responsible sharing of clinical trial data', *New England Journal of Medicine*, 369: 1651-8.
- Merson, L, TV Phong, LTN Nhan, NT Dung, NV Kinh, TTD Ngan, M Parker, and S Bull. 2015. 'Trust, respect and reciprocity: Informing culturally appropriate data sharing practice in Viet Nam', *Journal of Empirical Research on Human Research Ethics*, 10: 251-63.
- Merson, L., O. Gaye, and P. J. Guerin. 2016. 'Avoiding Data Dumpsters--Toward Equitable and Useful Data Sharing', *New England Journal of Medicine*, 374: 2414-5.

- National Academies of Sciences, Engineering, and Medicine. 2015. *Sharing Research Data to Improve Public Health in Africa: A Workshop Summary*. The National Academies Press: Washington, DC.
- OECD. 2007. "OECD Principles and Guidelines for Access to Research Data from Public Funding." OECD.
- Parker, M., S. J. Bull, J. de Vries, T. Agbenyega, O. K. Doumbo, and D. P. Kwiatkowski. 2009. 'Ethical data release in genome-wide association studies in developing countries', *PLoS Medicine*, 6: e1000143.
- Pearce, Neil, and Allan H. Smith. 2011. 'Data sharing: not as simple as it seems', *Environmental Health: A Global Access Science Source*, 10: 107.
- Pisani, Elizabeth, and Carla AbouZahr. 2010. 'Sharing health data: good intentions are not enough', *Bulletin of the World Health Organization*, 88: 462-6.
- Pisani, Elizabeth, James Whitworth, Basia Zaba, and Carla Abou-Zahr. 2010. 'Time for fair trade in research data', *Lancet*, 375: 703-5.
- Piowar, H. A., M. J. Becich, H. Bilofsky, R. S. Crowley, B. I. G. Data Sharing ca, and Workspace Intellectual Capital. 2008. 'Towards a data sharing culture: recommendations for leadership from academic health centers', *PLoS Med*, 5: e183.
- Rani, Manju, Hendrik Bekedam, and Brian S. Buckley. 2011. 'Improving health research governance and management in the Western Pacific: A WHO Expert Consultation', *Journal of Evidence-Based Medicine*, 4: 204-13.
- Rathi, Vinay, Kristina Dzara, Cary P. Gross, Iain Hrynaszkiewicz, Steven Joffe, Harlan M. Krumholz, Kelly M. Strait, and Joseph S. Ross. 2012. 'Sharing of clinical trial data among trialists: a cross sectional survey', *BMJ*, 345: e7570.
- Sankoh, Osman, and Carel Ijsselmuiden. 2011. 'Sharing research data to improve public health: a perspective from the global south', *Lancet*, 378: 401-2.
- Savage, Caroline J., and Andrew J. Vickers. 2009. 'Empirical study of data sharing by authors publishing in PLoS journals', *PLoS ONE*, 4: e7078.
- Strech, Daniel, and Jasper Littmann. 2012. 'Lack of proportionality. Seven specifications of public interest that override post-approval commercial interests on limited access to clinical data', *Trials*, 13: 100.
- Tangcharoensathien, Viroj, Jirawan Boonperm, and Pongpisut Jongudomsuk. 2010. 'Sharing health data: developing country perspectives', *Bulletin of the World Health Organization*, 88: 468-9.
- The Wellcome Trust. 2003. "Sharing Data from Large-scale Biological Research Projects: A System of Tripartite Responsibility. Report of a meeting organized by the Wellcome Trust and held on 14–15 January 2003 at Fort Lauderdale, USA." The Wellcome Trust.
- Toronto International Data Release Workshop. 2009. 'Prepublication data sharing', *Nature*, 461: 168-70.
- Vallance, Patrick, and Iain Chalmers. 2013. 'Secure use of individual patient data from clinical trials', *Lancet*, 382: 1073-4.
- Walport, Mark, and Paul Brest. 2011. 'Sharing research data to improve public health', *Lancet*, 377: 537-9.
- Wellcome Trust. 2009. "Policy on Data Management and Sharing. ." The Wellcome Trust.
- Whitworth, Jimmy. 2010. 'Data sharing: reaching consensus', *Bulletin of the World Health Organization*, 88: 467-8.
- Zarin, Deborah A. 2013. 'Participant-level data and the new frontier in trial transparency', *New England Journal of Medicine*, 369: 468-9.

Annex A

Table 1. Summary of Potential Benefits of and Concerns About Data Sharing.

Reasons to share individual-level data	Concerns about sharing individual-level data
<p>To improve science</p> <ul style="list-style-type: none"> • Enable verification, replication, and expansion of research results • Address biases, deficiencies, and dishonesty in research • Enable novel analyses and increase study power • Improve meta-analyses • Maximize data use, particularly for datasets that cannot be replicated • Inform research design and research funding • Improve teaching resources • Increase primary data producers' academic profiles and collaboration opportunities 	<p>May hamper science</p> <ul style="list-style-type: none"> • Reputational harms of critical secondary analyses • Consequences of flawed/poor quality secondary analyses • Reduction of incentives for primary research • Increased incentives to conduct short-term research rather than long-term research • Opportunity costs of curating and sharing data
<p>To improve health</p> <ul style="list-style-type: none"> • Inform health care planning and allocation • Inform regulatory review • Improve evidence base for clinical decision making • Improve use of health care resources • Improve patient care 	<p>May hamper health</p> <ul style="list-style-type: none"> • Effects of flawed secondary analyses on scientific evidence base • Burden of evaluating validity of secondary analyses • Effects of second-guessing regulatory procedures, policies, and processes
<p>Explicit moral claims</p> <ul style="list-style-type: none"> • Importance of maximizing the value and utility of data • Promotion of scientific values • Promotion of best practices in research conduct, analysis, and reporting • Demonstration of respect for research participants • Promotion of the public good 	<p>Explicit ethical issues</p> <ul style="list-style-type: none"> • Protection of participants' privacy and confidentiality • Validity of consent, including broad consent • Potential harms of secondary research for research participants including discrimination and stigma • Researchers' ability to fulfill commitments made to research participants during data collection • Effects of moral distance and limited awareness of the context in which data were collected • Potential impacts on public trust and confidence of conflicting analyses • Balancing the interests of differing stakeholders in data sharing • Making best use of limited research resources
	<p>Barriers to sharing</p> <ul style="list-style-type: none"> • Costs of developing and maintaining appropriate expertise and infrastructure • Curation costs • Ownership, intellectual property rights, and commercial confidentiality • Lack of policies and processes

Reproduced from Bull, S, N Roberts, and M Parker. 2015. 'Stakeholder perspectives on sharing de-identified individual-level data from medical and public health research: a scoping review', *Journal of Empirical Research on Human Research Ethics*, 10: 225-38.

The first part of the document discusses the importance of maintaining accurate records of all transactions. This includes not only sales and purchases but also any other financial activities that may occur. It is essential to ensure that all entries are properly documented and supported by appropriate evidence.

In addition, the document emphasizes the need for regular reconciliation of accounts. This process involves comparing the company's internal records with external statements, such as bank statements or supplier invoices, to identify any discrepancies. Regular reconciliation helps to prevent errors and ensures that the financial statements are accurate and reliable.

Furthermore, the document highlights the significance of maintaining up-to-date financial statements. These statements provide a clear and concise overview of the company's financial performance over a specific period. They are essential for internal decision-making and for providing information to external stakeholders, such as investors and creditors.

Finally, the document stresses the importance of transparency and accountability in financial reporting. This means that all financial transactions should be recorded and reported in a clear and honest manner. It is crucial to provide accurate and complete information to all relevant parties and to be open to scrutiny and audit.

October 2016

Version 1

Wellcome exists to improve health for everyone by helping great ideas to thrive. We're a global charitable foundation, both politically and financially independent. We support scientists and researchers, take on big problems, fuel imaginations and spark debate.

**Wellcome Trust, 215 Euston Road,
London NW1 2BE, UK
T +44 (0)20 7611 8888, F +44 (0)20 7611 8545,
E contact@wellcome.ac.uk, wellcome.ac.uk**

The Wellcome Trust is a charity registered in England and Wales, no. 210183. Its sole trustee is The Wellcome Trust Limited, a company registered in England and Wales, no. 2711000 (whose registered office is at 215 Euston Road, London NW1 2BE, UK).