Background paper: the ethics of data sharing and biobanking in health research

Meeting in Cape Town, 13-14 November 2018

Global Forum on Bioethics in Research

Purpose of this document

Data sharing and biobanking have the potential to support scientific research and increase scientific efficiency. These activities could be particularly useful in resource limited settings such as low- and middle- income countries (LMICs) in so far as they maximise the utility of data and minimise unnecessary duplication. However, researchers working with and in LMICs face several ethical challenges in respect of both data sharing and biobanking. The purpose of this paper is to map out some of the key ethical issues associated with data sharing and biobanking in LMICs to prepare for participation in the 2018 Global Forum on Bioethics in Research (GFBR) meeting.[1] The meeting will focus on human subject research and includes both digitised and non-digitised data.[2]

The document outlines the scope of the meeting theme and covers the following areas:

1. Introduction and context
2. Respecting participants and communities
3. Promoting equity
4. Advancing good governance

There is a significant amount of literature on the ethics of data sharing and biobanking in high income countries (HICs). To date, much of the literature has focused on the issues of confidentiality, informed consent and the different models of consent that might be used for the storage and use of tissue and data. As data sharing and biobanking practices expand to the LMIC context, new ethical, legal and social concerns have emerged in LMICs. For example, global collaborative research projects are increasing in number and size and this has led to concerns about ownership, control, infrastructure and sustainability, particularly in LMIC settings.

Also, epidemic diseases remain a grave threat to the world and the timely sharing of data and knowledge is likely to be an essential part of responding to this threat. Stakeholders are currently drawing up frameworks for the collection and use of samples in emergency situations. (Delauney et al, 2016) This raises the question as to whether or not, there are fundamentally different ethical issues raised when conducting research in non-emergency versus emergency situations.

It is also important to recognise that the landscape of data ethics and biobanking continues to evolve, bringing with it new challenges and opportunities. Take for example two emerging issues in HICs. First, stakeholders in HICs are currently considering the impact of linking health and social data to genomic data and other existing research data. Second, there is growing concern about the ethical implications of an increasing move to link large databases and permit exploration with machine learning/AI approaches. While these issues may not currently be at the forefront of discussions in low resource settings there is a need to consider the value and impact of data linkage and the use of AI approaches in these contexts.

[1] The meeting will focus on ethics but consider broader issues e.g. social, economic and legal where they relate to ethical issues.

[2] Data sharing and biobanking for public health is out of the meeting scope as is the collection and sharing of pathogens, non-human animals, plants etc.

The upcoming GFBR meeting will examine the ethical issues of relevance and importance to LMICs and provide an opportunity for stakeholders (e.g. bioethicists, researchers, scientists, funders, policy-makers) to engage in rigorous critical assessment through discussion of real-life LMIC case studies.

This paper is being published with the call for case studies and proposals on guidance and policy issues. Case studies and proposals may relate to the issues below or other issues that present ethical challenges. They should focus on research in LMICs and could address (but are not limited to) one or more of the following general questions:

Respecting participants and communities (see part 2 for more specific questions)

- What constitutes 'genuine' community engagement for data sharing and biobanking research in LMICs?

Promoting equity (see part 3 for more specific questions)

- What are the key drivers and barriers to data sharing and biobanking in LMIC settings?

Advancing good governance (see part 4 for more specific questions)

- How do you ensure that governance processes are appropriate and fit for purpose in LMICs?
- What are the roles and responsibilities of stakeholders (e.g. researchers, funder, policymakers, etc.) in facilitating ethical and equitable data sharing practices?

1. Introduction and context

Data sharing is an important requirement for effective and efficient biomedical research. Researchers are increasingly required to share data so that it can have the greatest possible impact. To this end, many funding agencies and scientific journals are adopting data sharing policies. (Wellcome, 2017; NIH, 2003; Taichman, 2017)

In the context of LMICs, clinical and public health data have the potential to generate valuable datasets to address the challenge of disease burden that low income settings disproportionately face. (Bull, Roberts and Parker, 2015; Pisani and AbouZahr, 2010; Parker and Kwiatkowski, 2016) International collaborative platforms such as the WorldWide Antimalarial Resistance Network (WWARN), the Infectious Diseases Data Observatory (IDD0) and H3Africa,[3] can consolidate large datasets and generate reliable evidence that will enable research-driven responses to some of the major challenges in LMICs.

There is a growing movement toward having 'open data' – which is openly accessible via a public repository. (Groves, 2012) The Scholarly Publishing and Academic Resource Coalition defines open data as being "freely available on the internet permitting any user to download, copy, analyse… without financial, legal or technical barriers other than those inseparable from gaining access to the internet itself." (Serwadda et al, 2018) There is less support for the notion of open data sharing among health researchers in LMICs where concerns about resource inequities, social justice and historical contexts may lead to a mistrust of open data policies. (Serwadda et al, 2018; Denny et al, 2015) Researchers' responsibilities can also impact on views about how data should be shared – including their responsibilities to protect participants' privacy and confidentiality, and to ensure data is used for acceptable purposes by researchers with the relevant expertise. Views about what

[3] For more information, see the following websites: WorldWide Antimalarial Resistance Network (WWARN) at http://www.wwarn.org/ , Infectious Diseases Data Observatory (IDD0) at https://www.iddo.org/ and H3Africa at https://www.fic.nih.gov/Funding/Pages/collaborations-h3africa.aspx.

mechanisms are needed to ensure appropriate stewardship and custodianship of shared data and protection of data subjects' interests can impact on a researchers' willingness to share data openly. Often the more complex and detailed the dataset, the greater the concern about open access data.

Biobanks are biorepositories that accept, process, store and distribute bio specimens and associated data for use in research and sometimes clinical care. They can be from university based repositories, institutional and government supported repositories, commercial biorepositories, population based biobanks and virtual biobanks. (De Souza, 2013)

Four developments in biobanking are worth noting:

- there has been a recent expansion of efforts by LMICs to build their own biobanks to collect, archive and re-use human biological samples and their related data for public health purposes and to support genetic studies. (Klingström et al, 2016)
- the breadth and complexity of data associated with – or that can be derived from – stored samples is steadily increasing to encompass many different aspects such as genetic and proteomic information, particularly if immortalised cell lines are generated. This in turn has led to some debate about the blurring of lines between data sharing platforms and biobanks. While many of the ethical challenges associated with data sharing are also encountered in biobanking, there are several questions that are specific to biobanking which are directly related to the collection and storage of samples. For example, biobanking raises questions about what counts as 'appropriate' re-use that is in line with communities' values[4] and the prioritisation of the use of samples given that they are a depletable resource.
- there is a recognition that the richness of biobank use over time will depend on linkages with health data, genomic data and other multi-omic data integration and analyses.
- there is move towards a more dispersed and multi-jurisdictional arrangement for biobanks in globalised research. How will sample and data protection principles and regulations from one jurisdiction, intended to promote sharing while also protecting the interests of subjects, impact on research in LMICs (e.g. the EU General Data Protection Regulation)?

The following sections of the background paper are predicated on the assumption that the challenges which are faced by data sharing and biobanking are similar; where biobanking raises specific ethical concerns, these are flagged and discussed in greater detail. For the purposes of the GFBR meeting the current ethical issues in both data sharing and biobanking in LMIC research will be divided into three broad categories:

- Respecting individual participants and communities
- Promoting equity
- Advancing good governance.

This paper concludes with an overview of relevant ethical guidelines.

2 Respecting individual participants and communities[5]

Data sharing and biobanking should be carried out in a way that protects the interests of individuals and affected communities while ensuring the maximum benefit to health using shared data and specimens.

[4] Although there are particular concerns in relation to the use of human samples, this is also relevant to data sharing.

[5] Respect is used here in its broadest sense, beyond respecting dignity and privacy.

2.1 Respecting individual participants

Consent: A properly executed consent process[6] is a vital part of respecting individual participants and maintaining public trust in research. Researchers engaged in data/sample sharing must ensure that effectively designed consent processes are put into place. However, this may be challenging given low literacy levels, unfamiliarity with health research, and language barriers that hinder the communication of terminology. (Traore et al, 2015; Simon et al, 2011, Tindana, Bull et al. 2012) There has also been much debate about appropriate models of consent to allow sharing, storage and future use of samples and data. Whilst there is no consensus, the model currently favoured by the research community is broad consent. (Simon et al, 2011, Bull, Cheah et al, 2015) One paper reviewing whether the use of broad consent is appropriate in LMICs suggested that whilst there are no a priori ethical reasons to prevent the use of this consent model in LMICs, there are some clear requirements that it should only be used in conjunction with a governance framework and genuine community engagement (Tindana et al, 2016, Tindana et al 2017). Given the complexity, unfamiliarity and abstract nature of data and sample sharing, providing accessible information about sharing can be challenging.

Questions relating to informed consent include:

- how information about storage, sharing and future use can be effectively communicated in the LMIC research setting.
- whether broad consent is acceptable in specific research contexts. (Please note that it is not the intention of this meeting to revisit conceptual debates about the acceptability of broad consent versus specific consent.)
- whether there are real possibilities for withdrawal of data or samples from future research
- whether children who provided samples for paediatric research – based on their parents' consent – should be re-contacted and consent when they are older for the continued use of their samples. (Giesbertz, 2016)
- whether it is appropriate to use legacy tissue in situations where adequate consent has not been obtained.

Privacy and confidentiality*:* Many concerns about data sharing and biobanking have been described in terms of privacy and confidentiality. (Mittlestadt and Floridi, 2016) De-identification is frequently offered as a solution to protecting individual privacy but risks of re-identification may remain depending on anonymization standards. This is particularly a concern in relation to the sharing of health records and for longitudinal biobanks, which need to link between personal data and the de-identified data (e.g. if they collect periodic updates from a person's health record). The lack of consistency in the applications of anonymization remains an unresolved issue. (Emam et al, 2015). Privacy protection remains a major issue and there is a need to examine whether existing regulatory and governance structures are responsive to individual privacy concerns, while at the same time, allowing for effective and efficient research. (Kaye, 2012)

2.2 Respecting communities – community engagement[7]

Community engagement (CE) can be time consuming and needs to be well-resourced as it is hard to explain some types of complex research. The rationale, methods and implications of data sharing and sample sharing tend to be complex and theoretical and harder to explain as compared to more standard trials. There are various approaches to community engagement and strategies should be in

[6]Participants should have capacity, be informed (provided with relevant information) and free to decide (i.e. voluntarily give their consent with no coercion/undue influence).

[7] If you would like to find out more about community engagement, see https://mesh.tghn.org/

keeping with the nature of the research and goals of the engagement. (Tindana et al, 2015) Whilst the use of broad consent for LMICs genomic and biobank research incorporating sample and data sharing is generally accepted, it has been suggested that this is only permissible if accompanied by 'genuine' community engagement. (Tindana, 2014) There needs to be more discussion as to what constitutes 'genuine' community engagement for research incorporating data and sample sharing in LMICs. Community engagement may be a useful tool in the following instances:

***Ensuring trustworthiness**:* Typically, discussions about trust in research involve discussions about improving public trust in research. However, this approach has been criticised as presenting the public as passive recipients of knowledge while implying that science has the 'right' answers. Therefore, it is suggested that it may be more constructive to think about what it means for research to be trustworthy. This would require a deeper reflection on what it means for research or researchers to be trustworthy and on what bases public trust is founded. (Aitken et al, 2016) The importance of ensuring trustworthiness cannot be overstated as studies have shown that patients' trust in researchers was the most powerful determinant of the kind of control they want over their data. (ter Meulen et al, 2011; Damschroder et al, 2007; Bull, Cheah et al, 2015 and Jao, Kombe et al, 2015)

Community engagement is a valuable tool in raising awareness and providing information to stakeholders. It also provides opportunities for researchers to reflect on the concerns and preferences of community members and to build higher levels of trustworthiness into research practices and institutional arrangements. (Aitken, et al 2016)

What methods can be used to engage communities in a way that increases trustworthiness in data sharing and biobanking?

***Identifying risks of discrimination and marginalisation**:* Researchers should be aware of the potential risks of discrimination and marginalisation of communities, especially in situations where communities are vulnerable given their social, economic or political contexts. (Jao et al, 2015) Evidence suggests that participants themselves are more likely to fear discrimination when they are uncertain about the motivation of researchers. (Hate et al, 2015) There is a risk that genomics research could increase stigma in cases where groups or conditions are already stigmatised (De Vries et al, 2012), or where a genetic explanation of illness may extend stigma to biological relatives. (Tekola et al, 2009)

> *But if you collect some demographic data, for example, some initial or some region, it is possible that they can track back. Even though you didn't put the name, the 13-digit ID, something else may allow you to track them, who they are . . . maybe the insurance company wants to get this information and want to know if this population in this region want to buy insurance, they may want to get it from you and they may know that these people got these diseases regularly. (TH-CTSG-G01-R1, Clinical Trials Coordinator, female)* (Cheah et al, 2015)

Active solicitation of views on collection and use of data/samples: It may be appropriate in certain circumstances to consider engaging the community or community representatives about their views on how data and samples should be collected, stored, shared and used. This is particularly important in biobank research as there may be higher expectations when tissue (rather than data) is stored and used. Certain communities may have views on the removal, storage, disposal and even return of the tissue. Researchers should consider the extent to which community engagement can be used to support the development of access processes or to reach access decisions. It has been suggested that CE should be an ongoing process and should play a role in ensuring that ongoing sample and tissue use is culturally appropriate. (Tindana, 2014) Questions as to which CE approaches and

methods could be used to ensure community involvement in access decisions on an ongoing basis – and whether this is desirable – remain unanswered.

> *I think . . . this body should not just be left to people with expertise alone. I think we should also have . . . representation from the community where this data is generated, so that they can be part of these decisions and see how the data is used . . . I think it would bring more trust from the community that "I was there" . . .. (IDI12, nurse manager, female)* (Jao et al, 2015)

How can tissue removal, storage and disposal be managed in a way that is culturally sensitive?

Which community engagement approaches could be used to ensure community involvement in access decisions?

3 Promoting equity

Data sharing and biobanking should recognise and balance the needs of the different communities involved. This includes researchers who generate the data, secondary users of the data, the communities from which the data or specimens came, and funders of the collection effort. (National Academy of Sciences, 2015) Researchers who generate data have a valid interest in using the data for their own interests and should be empowered to ask questions that are relevant to their immediate environment. Promoting equity in data sharing and biobanking involves consideration of equitable data sharing and benefit sharing in a number of stakeholders[8], but for the purposes of this meeting, the focus will be on the interests of researchers who generate and use data and data/sample providers/communities.

3.1 Promoting equitable data and sample sharing – researchers in LMICs

Data sharing practices are still relatively uncommon in LMICs where the necessary policies, expertise and infrastructure to ensure the meaningful use of publicly available data are not well established. Knowledge and infrastructure gaps exist between HICs and LMICs. Moreover, the development of new approaches and technologies in HICs such as data linkage and AI technology, could compound existing inequalities given the disparities in infrastructure and knowledge in LMICs. To avoid exacerbating existing inequalities, steps need to be taken to promote the collection and use of data in ways that achieve equitable outcomes. (Bull, 2016)

> *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.* (Bull, 2017)

Researchers from LMICs may generate data and/or collect samples in a range of contexts. For example, as part of local or national studies with no immediate intention of sharing, to engaging in international collaborations where processes for sample and data sharing have been negotiated and

[8] Multiple stakeholder groups have been identified as having potential interests in data sharing, including funders, regulators, research reviewers, policy developers, the broader scientific community, and the public.

established prior to sample and data collection. In both situations, LMIC researchers have a valid interest in using their data effectively but they may need time to conduct initial analyses of their findings. In addition to this, they may miss out on career advancement opportunities if they are not acknowledged in secondary analyses of the data. (Parker, Bull et al 2009) In such cases, data sharing may be of greater benefit to researchers in HICs who have higher analytical capabilities and can use data generated from LMICs more quickly and more efficiently to publish articles in high impact journals, while researchers in LMICs struggle to get published in the same journals. Concerns about potential commercial exploitation of data are also not uncommon.

> *We are all in the business, profit and business. So researchers, they don't produce anything that you can sell. I am not making mobile phones or I am not making plastic ware. I am making data and knowledge; I cannot sell them. The only thing I can do is produce the result that will convince the sponsor to give me money to continue to produce results… so, because we are now living in the world of the economic model like that, if people are using my work to make money for themselves, because if they use the data they publish paper, their rank goes higher, they get more funding and they get money, not me. (TH-SR-1-15, Senior Researcher) (Cheah, et al, 2015)*

International collaborations may provide good opportunities for researchers to work with partners to develop mutually beneficial data sharing arrangements. Examples of good practice initiatives from LMICs include INDEPTH, MalariaGEN, H3Africa, and WWARN.[9] They 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. (Bull, 2017) However, while international collaborations have the potential to support equitable sharing – both between the collaborators and with external researcher – this may not always be promoted. Evidence suggest that researchers in LMIC settings may face challenges in negotiating equitable contractual relationships with researchers in high resource settings. (Sankoh et al, 2011)

What constitutes ethical and equitable sharing of data and samples in international collaborations?

What are the challenges faced by researchers in LMICs in developing and managing international collaborations in data sharing/biobanking research?

Data sharing for secondary analyses: Researchers in LMICs may want to access data generated by others for their own secondary analyses. There is added value to nurturing strong research capacity in LMICs, and to having LMIC researchers lead research including secondary data analyses, as they have contextual information that helps them articulate good (context-specific) research questions and interpret data. Currently, researchers in LMICs rarely make requests for data for secondary analyses as they lack the capacity to analyse datasets. In the case of genomic research, many researchers struggle to even download datasets they have applied for and may need support to help them through the process. Concerns have been raised that researchers merely have theoretical access to data but are unable to utilise the data in a practical manner.

[9] For more information see websites: INDEPTH at http://indepth-network.org/ , MalariaGEN at https://www.malariagen.net/ , H3Africa at https://www.fic.nih.gov/Funding/Pages/collaborations-h3africa.aspx and WWARN at http://www.wwarn.org/.

> *Although the majority of (Mahidol Oxford Tropical Medicine Research Unit) MORU research data is generated in LMICs, to date, no requests for access to MORU data have been received from institutions in LMICs. Instead, applicants tend to be from well-resourced groups in higher-income settings who have good IT infrastructure and the capacity to conduct complicated statistical analyses and mathematical modelling*. (Cheah et al, 2017)

If these gaps in knowledge and capacity are left unmet, researchers in LMICs might be consigned to the role of data collection, thus exacerbating current inequalities.

> *… researcher and several others expressed concerns that without more exposure to global datasets and training in complex meta-analysis, scientists from non-endemic countries would be unable to join the "big data" era. They would be increasingly consigned to the data collection end of the research spectrum, their involvement in analytic collaborations such as study groups merely tokenistic. (125, WWARN science group head)* (Pisani and Botchway, 2017)

It is important that researchers in LMICs can use their data effectively and in a timely manner as well as utilise existing datasets. Such use has the potential to translate into research that is of value to LMICs as these researchers are in the best positions to ask questions that are relevant to their immediate environments and to curate data in ways that maximises their utility and minimises the possibilities of flawed secondary analyses. (Bull 2016) Ethical research would therefore require promoting fairness and building capacity.

What are the opportunities to translate theoretical access to data and samples into practical access to data and samples for researchers in LMICs?

Promoting fairness: Data sharing should be conducted in a way that does not adversely affect the careers of researchers or impede their ability to conduct research that is relevant to the needs of the communities in which they work. (Bull, Cheah et al, 2015) At present, professional recognition and progression is determined by one's capability to publish research, (Walport and Brest, 2011) particularly in high impact journals. There remains a need to consider the ways in which the contribution of data sharers and technicians can be recognised and acknowledged with a view to safeguarding career paths for data scientists and technicians. Also, the current debate in LMICs should consider how LMIC researchers can be empowered to lead research projects, data analysis and the write-up of manuscripts that get published in high impact journals.

What are effective ways to address researchers' concerns that mechanisms for data sharing may adversely affect their career development?

Building capacity: The inability to analyse and publish data quickly disproportionately impacts low resource settings. (Bull, Roberts and Parker, 2015) To promote long term sustainable research and collaboration, the capacity to curate, share and analyse high-quality data sets needs to be built and fostered in LMIC settings. International collaboration may provide good opportunities for capacity building.

The CIOMS International Ethical Guidelines for Health-Related Research Involving Humans states that 'Health-related research often requires international collaboration and some communities lack the capacity to assess or ensure the scientific quality or ethical acceptability of health-related research proposed or carried out in their jurisdictions. Researchers and sponsors who plan to conduct research in these communities should contribute to capacity-building for research and

review.' Capacity in this context includes 'research infrastructure building and strengthening research capacity.' (CIOMS, 2016)

There is also some concern that funders and sponsors may focus on efficient delivery outputs and are reluctant to support capacity building as it is time consuming. (Pisani and Botchway, 2017)

Is the use of data sharing and biobanking an important and useful way of doing research in LMICs and why? Is developing these capacities important and appropriate in resource limited settings?

3.2 Promoting equitable sharing of benefits – participants and communities

Participants and communities involved in research have interests in sharing the benefits of research arising from their contributions.[10] However, it is still unclear as to what would constitute a benefit and who this should be shared with. (Ramsay et al, 2014) In discussions, stakeholders have discussed the importance of both direct and indirect benefits. (Bull, Roberts and Parker, 2105) Indirect benefits are particularly relevant in the context of secondary research which may not address health issues of relevance to participants and communities. In such cases, indirect benefits such as the ability to advance health more generally may be of interest to the community. (Bull, Cheah et al, 2015) For example, the AWI-Gen project in Africa aims to identify genetic factors that contribute to body composition, including among other factors, obesity. Suggestions were made at a workshop that AWI-Gen could provide additional benefit through public education on obesity and link to existing patient organisations to provide relevant information. (Ramsay et al, 2014)

What benefits should data sharing and biobanking research aim to deliver and who should be the beneficiaries? Who decides on these matters and how can the realisation of benefits be promoted (e.g. through good governance)?

Commercialisation and sharing of benefits: The use of data sharing and biobanking platforms for commercial gain can be a sensitive issue. Commercialisation raises concerns about Intellectual property and ownership rights. (Petrini, 2012) Community expectations and views may also vary considerably depending on historical, political and cultural contexts. For example, in Vietnam, commercialisation is said to be welcomed because it is viewed as the best likelihood to advance health. (National Academies of Sciences, Engineering, and Medicine, 2015) Stakeholders in Mumbai however, were warier about the objectives of researcher. (Hate et al, 2015)

> *The recent sharing of avian flu virus specimens by developing countries through the World Health Organization resulted in the production of avian influenza vaccines at a price of US$ 10–20 per dose. This is unaffordable in low-income countries where total health expenditure is less than US$ 30 per person. Should an avian flu pandemic occur, there would be huge death tolls in countries without access to vaccines; while rich countries' populations would be fully protected, literally from any moral obligations to countries that shared their specimens. Such unilateral benefit inhibits data sharing.* (Tangcharoensathien et al, 2010)

What are the opportunities and challenges of commercialisation?

[10] For an overview of the discourse on benefit sharing, see Dauda and Dierckx (2013).

4 Advancing good governance

Data sharing and biobanking have the potential to improve the quality and value of research. Data sharing allows for the independent scrutiny of research results, which increases their reliability and reproducibility as well as enhancing accountability. As data management tools become more sophisticated, complex analyses can be carried out more efficiently, generating new and valuable knowledge. However, research that is inappropriately prepared or inappropriately shared may hamper rather than promote health if misinterpreted or the subject of biased, inappropriate or poorly designed projects. In the context of LMICs, it may be pertinent to question whether data sharing and biobanking research is responsive to the health problems of people living in LMICs. (Rudan, 2011)

Advancing effective and efficient research involves considerations of governance and sustainability of research platforms.

Accountable, efficient, fair and proportionate governance frameworks: Robust governance processes are needed to ensure that research is carried out efficiently, effectively and ethically. Many of the earlier processes described above such as informed consent, community engagement and equitable sharing are necessary steps in promoting ethical research but are only effective if incorporated into ethically appropriate governance frameworks.

Governance frameworks typically include Data Access Committees (DAC) and Ethics Review Boards. DACs are responsible for data release to external requestors based on legal, ethical and scientific eligibility. (Mulder et al, 2017) A recent study involving interviews with DAC members and experts from North American and Europe, observed they had concerns about the effectiveness and consistency of current access review procedures and oversight processes. (Shabani, 2016) DACs in LMICs are likely to face similar challenges. There is a need for more discussion about issues faced by DACs in LMIC settings.

Who has the capacity or authority (e.g. DAC, Ethics Review Board) to evaluate data sharing applications?

The literature suggests that there is a lack of appropriate regulation and governance mechanisms for biobanking or data sharing in LMICs. (Bull, Roberts et al, 2015) Regulations are often absent, outdated, conservative, or inefficient and difficult to navigate. (de Vries et al, 2017) In some instances, regulations may not apply to small biobanks and there is some concern about the lack of oversight of the activities of these small biobanks, particularly in relation to commercial activities.

Multiple different approaches and systems risk creating a fragmented and confusing landscape that fails to realise the full benefits of data sharing. To address this, a number of international initiatives[11] have offered governance structures and guidelines to promote the internationalization and standardization of biobanks. There is a need to further consider the attributes of an ideal model of good governance of international consortia which does not disadvantage certain partners in the consortia.

Chen and Pang (2014) have called for further progress through the consolidation of existing guidelines into a single global governance framework for biobanks. However, who should co-ordinate this work, its feasibility and the legitimacy of the resulting framework are critical questions.

[11] Public Population Project in Genomics and Society (http://www.p3g.org/), the International Society for Biological and Environmental Repositories (http://www.isber.org/), the International Agency for Research on Cancer (https://www.iarc.fr/)and the Global Alliance for Genomics and Health (https://www.ga4gh.org/).

How can we develop and implement accountable, efficient, fair and proportionate governance frameworks to support ethical best practice in biobanking and data sharing, with harmonisation across borders?

Sustainability of data sharing and biobanking: Setting up and managing biobanks and data sharing platforms requires sophisticated technology and laboratory infrastructure as well high levels of expertise. While initial funding may be made available for the setting up and maintenance of repositories for a period, sustaining these repositories in the long term remains a challenge even in high income settings. Surveys among biobanks operating in a LMIC setting indicate that limited resources and short term funding tied to specific projects threaten the sustainability of the biobanks. (Klingström et al, 2016) Cost recovery and commercialisation models are being considered, which may create barriers for researchers in LMICs. Further discussions are needed to develop ethically appropriate sustainability models. Questions have also been raised as to whether the best path forward is to build capacity for repositories in LMICs or whether it would be better to focus on ensuring access, credit and benefit while storing samples where capacity already exists.

What would sustainability models for biobanks and data repositories look like?

Is the best path forward to build capacity for repositories in LMICs or would it be better to focus on ensuring access, credit and benefit while storing samples where capacity already exists?

Relevant Ethical Guidelines

CIOMS, (2016), International Ethical Guidelines for Health-related Research Involving Humans, Fourth Edition. Geneva. Council for International Organizations of Medical Sciences at https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf	Guideline 1 – Scientific and social value and respect for rights Guideline 2 – Research conducted in low resource settings Guideline 3 – Equitable distribution of benefits and burdens in the selection of individuals and groups of participants in research Guideline 4 – Potential individual benefits and risks of research Guideline 6 – Caring for participants' health needs Guideline 7 – Community engagement Guideline 8 – Collaborative partnership and capacity building for research and research review Guideline 10 – Modifications and waivers of informed consent Guideline 11 – Collection, storage and use of biological materials and related data. Guideline 12 – Collection, storage and use of data in health related research Guideline 13 – Reimbursement and compensation for research participants
Convention for the Protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine:	A framework Convention that aims to protect the dignity and identity of all human beings and guarantee everyone, without

Convention on Human Rights and Biomedicine (Oviedo Convention) at https://rm.coe.int/168007cf98	discrimination, respect for their integrity and other rights and fundamental freedoms with regard to the application of biology and medicine. It sets out fundamental principles applicable to daily medical practice and is regarded as such at the European treaty on patient's rights. It also deals specifically with biomedical research, genetics and transplantation of organ and tissues.
Convention on Biological Diversity at https://www.cbd.int/convention/text/default.shtml	The Convention has 3 main objectives: 1. The conservation of biological diversity 2. The sustainable use of the components of biological diversity 3. The fair and equitable sharing of the benefits arising out of the utilization of genetic resources
Nagoya Protocol at https://www.cbd.int/abs/text/default.shtml	This is a supplementary agreement to the Convention on Biological Diversity (CBD). It provides a transparent legal framework for the effective implementation of one of the three objectives of the CBD: the fair and equitable sharing of benefits arising out of the utilization of genetic resources
Universal Declaration on Human Genome and Human Rights at http://www.unesco.org/new/en/social-and-human-sciences/themes/bioethics/human-genome-and-human-rights/	

References

Aitken, M., Cunningham-Burley, S., & Pagliari, C. (2016). Moving from trust to trustworthiness: Experiences of public engagement in the Scottish Health Informatics Programme. *Science and Public Policy*, *43*(5), 713-723.

Bobrow, M., Banks, J., Burton, P., Smith, G. D., & Eeles, R. (2014). Establishing incentives and changing cultures to support data access. Wellcome Trust http://www. wellcome. ac. uk/stellent/groups/corporatesite/@ msh_peda/documents/web_document/wtp056495. pdf.

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 Ethic*s 10(3): 302-313.

Bull, S., Roberts, N., and Parker, M. (2015) Views of Ethical Best Practices in Sharing Individual-Level Data from Medical and Public Health Research: A Systematic Scoping Review. *Journal of Empirical Research on Human Research Ethics*, 10(3).

Bull, S. (2017). Ensuring global equity in open research at https://wellcome.ac.uk/sites/default/files/ensuring-global-equity-open-research-oct16.pdf

Cheah, P. Y., Tangseefa, D., Somsaman, A., Chunsuttiwat, T., Nosten, F., Day, N. P., ... & Parker, M. (2015). Perceived benefits, harms, and views about how to share data responsibly: A qualitative study of experiences with and attitudes toward data sharing among research staff and community representatives in Thailand. *Journal of Empirical Research on Human Research Ethics*, 10(3), 278-289.

Cheah, P. Y., Day, N. P., Parker, M., & Bull, S. (2017). Sharing Individual-Level Health Research Data: Experiences, Challenges and a Research Agenda. *Asian Bioethics Review*, 9(4), 393-400.

Chen, H., & Pang, T. (2014). A call for global governance of biobanks. *Bulletin of the World Health Organization*, *93*, 113-117.

CIOMS, (2016), International Ethical Guidelines for Health-related Research Involving Humans, Fourth Edition. Geneva. Council for International Organizations of Medical Sciences at https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf

Damschroder, L. J., Pritts, J. L., Neblo, M. A., Kalarickal, R. J., Creswell, J. W., & Hayward, R. A. (2007). Patients, privacy and trust: Patients' willingness to allow researchers to access their medical records. *Social science & medicine*, *64*(1), 223-235.

Dauda, B., & Dierickx, K. (2013). Benefit sharing: an exploration on the contextual discourse of a changing concept. *BMC Medical Ethics*, 14(1), 36.

Delaunay, S., Kahn, P., Tatay, M., & Liu, J. (2016). Knowledge sharing during public health emergencies: from global call to effective implementation. *Bulletin of the World Health Organization*, *94*(4), 236.

Denny, S. G., B. Silaigwana, D. Wassenaar, S. Bull and M. Parker (2015). "Developing Ethical Practices for Public Health Research Data Sharing in South Africa: The Views and Experiences From a Diverse Sample of Research Stakeholders." J Empir Res Hum Res Ethics **10**(3): 290-301.

De Souza, Y. G., & Greenspan, J. S. (2013). Biobanking past, present and future: responsibilities and benefits. AIDS (London, England), 27(3), 303.

de Vries, J., et al. (2012). "Investigating the potential for ethnic group harm in collaborative genomics research in Africa: Is ethnic stigmatisation likely?" *Social Science & Medicine* 75(8): 1400-1407.

de Vries, J., Munung, S., Matimba, A., McCurdy, S., Oukem-Boyer, O., Staunton, C., Yakubu, A., Tindana, P., and the H3Africa Forum. (2017) Regulation of genomic and biobanking research in Africa: a content analysis of ethics guidelines, policies and procedures from 22 African countries. *BMC Medical Ethics*, 18(8).

El Emam, K., Rodgers, S., & Malin, B. (2015). Anonymising and sharing individual patient data. bmj, 350, h1139.

Giesbertz, N. A., Bredenoord, A. L., & van Delden, J. J. (2016). When children become adults: should biobanks re-contact*?. PLoS medicine*, 13(2), e1001959.

Groves, T., & Godlee, F. (2012). Open science and reproducible research, *BMJ* 344, e4383

Hate, K., Meherally, S., Shah More, N., Jayaraman, A., Bull, S., Parker, M., & Osrin, D. (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(3), 239-250.

Jao, I., Kombe, F., Mwalukore, S., Bull, S., Parker, M., Kamuya, D., ... & Marsh, V. (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(3), 264-277.

Jao, I., F. Kombe, S. Mwalukore, S. Bull, M. Parker, K. D, C. Molyneux and M. V (2015). Research stakeholders' views on benefits and challenges for public health research data sharing: The importance of trust and social relations. *Plos One* in press.

Kaye, J. (2012). The tension between data sharing and the protection of privacy in genomics research, *Annual review of genomics and human genetics*, *13*, 415-431.

Klingström, T., Mendy, M., Meunier, D., Berger, A., Reichel, J., Christoffels, A., ... & Bongcam-Rudloff, E. (2016, May). Supporting the development of biobanks in low and medium income countries. In *IST-Africa Week Conference*, 2016 (pp. 1-10). IEEE at https://ieeexplore.ieee.org/document/7530672/#full-text-section

MalariaGen, Our approach to sharing data at https://www.malariagen.net/data/our-approach-sharing-data

Merson, L., Phong, T. V., Nhan, L. N. T., Dung, N. T., Ngan, T. T. D., Kinh, N. V., ... & Bull, S. (2015). Trust, respect, and reciprocity: informing culturally appropriate data-sharing practice in Vietnam. *Journal of Empirical Research on Human Research Ethics*, 10(3), 251-263.

Mittelstadt, B. D., & Floridi, L. (2016). The ethics of big data: Current and foreseeable issues in biomedical contexts. *Science and Engineering Ethics,* 22(2), 303-341.

MORU Tropical Network Data Sharing Policy 6th January 2016 version 1.0 at http://www.tropmedres.ac/_asset/file/data-sharing-policy-v1-0.pdf

Mulder, N., Adebamowo, C. A., Adebamowo, S. N., Adebayo, O., Adeleye, O., Alibi, M., ... & Ghedira, K. (2017). Genomic Research Data Generation, Analysis and Sharing–Challenges in the African Setting. *Data Science Journal*, 16.

National Academies of Sciences, Engineering, and Medicine. (2015). Sharing Research Data to Improve Public Health in Africa: A Workshop Summary. National Academies Press.

National Institutes of Health (NIH) (2003), Final NIH Statement on Sharing Research Data at https://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html.

Parker, M., & Kwiatkowski, D. P. 2016. The ethics of sustainable genomic research in Africa. *Genome Biology*, 17(1), 44. http://doi.org/10.1186/s13059-016-0914-3

Petrini, C. (2012). Ethical and legal considerations regarding the ownership and commercial use of human biological materials and their derivatives. *Journal of Blood Medicine* (3), 87.

Pisani, E., and AbouZahr, C. (2010). Sharing health data: good intentions are not enough. *Bulletin of the World Health Organization*, 88

Pisani, E., & Botchway, S. (2017). Sharing individual patient and parasite-level data through the WorldWide Antimalarial Resistance Network platform: A qualitative case study. *Wellcome open research*, 2.

Ramsay, M., De Vries, J., Soodyall, H., Norris, S. A., & Sankoh, O. (2014). Ethical issues in genomic research on the African continent: experiences and challenges to ethics review committees.

Rudan, I., Marušić, A., & Campbell, H. (2011). Developing biobanks in developing countries. *Journal of Global Health*, 1(1), 2.

Shabani, M., Thorogood, A., & Borry, P. (2016). Who should have access to genomic data and how should they be held accountable? Perspectives of Data Access Committee members and experts. *European Journal of Human Genetics*, 24(12), 1671.

Sankoh, O., & IJsselmuiden, C. (2011). Sharing research data to improve public health: a perspective from the global south. *The Lancet*, 378(9789), 401-402.

Serwadda, D., Ndebele, P., Grabowski, M. K., Bajunirwe, F., & Wanyenze, R. K. (2018). Open data sharing and the Global South—Who benefits?. *Science*, *359*(6376), 642-643.

Simon, C. M., L'heureux, J., Murray, J. C., Winokur, P., Weiner, G., Newbury, E., … Zimmerman, B. (2011). Active choice but not too active: public perspectives on biobank consent models. *Genetics in Medicine : Official Journal of the American College of Medical Genetics*, 13(9), 821–31.

Taichman, D. B., Sahni, P., Pinborg, A., Peiperl, L., Laine, C., James, A., ... & Frizelle, F. A. (2017). Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. *Annals of Internal Medicine*, 167(1), 63-65.

Tangcharoensathien, V., Boonperm, J., & Jongudomsuk, P. (2010). Sharing health data: developing country perspectives. *Bulletin of the World Health Organization*, 88(6), 468-469.

Tekola, F., et al. (2009). "Impact of social stigma on the process of obtaining informed consent for genetic research on podoconiosis: a qualitative study." *BMC Medical Ethics* 10(1): 13.

ter Meulen, R. H., Newson, A. J., Kennedy, M. R., & Schofield, B. (2011). *Genomics, health records, database linkage and privacy*. Nuffield Council on Bioethics.

Tindana, P., S. Bull, L. Amenga-Etego, J. de Vries, R. Aborigo, K. Koram, D. Kwiatkowski and M. Parker (2012). "Seeking consent to genetic and genomic research in a rural Ghanaian setting: a qualitative study of the MalariaGEN experience." *BMC Medical Ethics* 13: 15.

Tindana, P., Molyneux, C. S., Bull, S., & Parker, M. (2014). Ethical issues in the export, storage and reuse of human biological samples in biomedical research: perspectives of key stakeholders in Ghana and Kenya. *BMC Medical Ethics*, *15*(1), 76.

Tindana, P., De Vries, J., Campbell, M., Littler, K., Seeley, J., Marshall, P., ... & Parker, M. (2015). Community engagement strategies for genomic studies in Africa: a review of the literature. *BMC Medical Ethics*, *16*(1), 24.

Tindana, P. and J. De Vries (2016). "Broad Consent for Genomic Research and Biobanking: Perspectives from Low- and Middle-Income Countries." *Annual Review of Genomics and Human Genetics* 17: 2.1-2.19.

Traore, K., Bull, K., Niare, A., Konate, S., Thera, M., Kwaitowski, D., Parker, M., and Doumbo, O. (2015). Understandings of genomic research in developing countries: a qualitative study of the views of MalariaGEN participants in Mali. *BMC Medical Ethics*, 16(42). http://doi.org/10.1186/s12910-015-0035-7

Walport, M., and Brest, P. (2011). Sharing research data to improve public health. *The Lancet*, 377
Watson, P. H., Nussbeck, S. Y., Carter, C., O'Donoghue, S., Cheah, S., Matzke, L. A., ... & Johnston, R. N. (2014). A framework for biobank sustainability. *Biopreservation and biobanking,* 12(1), 60-68.

Wellcome Trust (2017) Policy on data, software and material management and sharing at https://wellcome.ac.uk/funding/managing-grant/policy-data-software-materials-management-and-sharing

This background paper was written by Sharon Kaur, Malaya University with input from the Planning Committee and GFBR secretariat. The paper draws on an earlier, shorter piece written by Matt Sample (Policy intern, Wellcome).

16 April 2018