

What constitutes equitable data sharing in global health research? A scoping review of the literature on low-income and middle-income country stakeholders' perspectives

Natalia Evertsz,¹ Susan Bull,^{2,3} Bridget Pratt⁴

To cite: Evertsz N, Bull S, Pratt B. What constitutes equitable data sharing in global health research? A scoping review of the literature on low-income and middle-income country stakeholders' perspectives. *BMJ Global Health* 2023;**8**:e010157. doi:10.1136/bmjgh-2022-010157

Handling editor Senjuti Saha

► Additional supplemental material is published online only. To view, please visit the journal online (<http://dx.doi.org/10.1136/bmjgh-2022-010157>).

Received 15 July 2022
Accepted 23 February 2023



© Author(s) (or their employer(s)) 2023. Re-use permitted under CC BY. Published by BMJ.

¹Royal Melbourne Hospital, Melbourne, Victoria, Australia

²Nuffield Department of Population Health, University of Oxford, Oxford, UK

³Department of Psychological Medicine, The University of Auckland, Auckland, New Zealand

⁴Queensland Bioethics Centre, Australian Catholic University, Banyo, Queensland, Australia

Correspondence to

Dr Bridget Pratt;
bridget.pratt@acu.edu.au

ABSTRACT

Introduction Despite growing consensus on the need for equitable data sharing, there has been very limited discussion about what this should entail in practice. As a matter of procedural fairness and epistemic justice, the perspectives of low-income and middle-income country (LMIC) stakeholders must inform concepts of equitable health research data sharing. This paper investigates published perspectives in relation to how equitable data sharing in global health research should be understood.

Methods We undertook a scoping review (2015 onwards) of the literature on LMIC stakeholders' experiences and perspectives of data sharing in global health research and thematically analysed the 26 articles included in the review.

Results We report LMIC stakeholders' published views on how current data sharing mandates may exacerbate inequities, what structural changes are required in order to create an environment conducive to equitable data sharing and what should comprise equitable data sharing in global health research.

Conclusions In light of our findings, we conclude that data sharing under existing mandates to share data (with minimal restrictions) risks perpetuating a neocolonial dynamic. To achieve equitable data sharing, adopting best practices in data sharing is necessary but insufficient. Structural inequalities in global health research must also be addressed. It is thus imperative that the structural changes needed to ensure equitable data sharing are incorporated into the broader dialogue on global health research.

INTRODUCTION

Data sharing is increasingly viewed as an important component of good science, due to its role in promoting research integrity, enabling review of research findings, increasing the statistical power of meta-analyses and making efficient use of resources.¹ Arguments that data sharing is an ethical imperative rest on its potential to

WHAT IS ALREADY KNOWN ON THIS TOPIC

- ⇒ Data sharing is thought to be an ethical imperative due to its potential to maximise the value and utility of collected data.
- ⇒ There is growing consensus that data sharing should be equitable; however, there is still no clear concept of equitable data sharing.

WHAT THIS STUDY ADDS

- ⇒ As a matter of procedural fairness and epistemic justice, the perspectives of low-income and middle-income country stakeholders must inform concepts of equitable health research data sharing.
- ⇒ This paper investigates their recent views in relation to how equitable data sharing in global health research should be understood.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

- ⇒ This review found that implementing best-practices in equitable data sharing is insufficient on its own.
- ⇒ Structural change is required in order for equitable data sharing to be possible.

maximise the value and utility of collected data, limiting unnecessary duplication of research, and thereby reducing the burdens and potential harms to research participants.^{2–4} It is generally assumed that data sharing will play an important role in improving health outcomes.^{5–11}

This paper focuses on the sharing of primary research datasets and metadata generated during health-related research in low-income and middle-income countries (LMICs). In LMICs, relatively limited research resources have resulted in disproportionately smaller amounts of data being collected relative to disease burdens.¹² Sharing research data collected from LMICs is expected to make significant

contributions to global health by maximising the available evidence to inform healthcare, health policy, health system funding and regulatory decisions.^{2-4 13 14} Thus, data sharing is increasingly mandated by journals, publishers and funders, and calls have been made to strengthen the current mandates, requiring scientists worldwide to commit research data to publicly accessible databases within specified timeframes.¹⁵ Powerful public and philanthropic stakeholders have profoundly influenced the development of international standards for such sharing and institutional policies requiring data sharing as a condition of research support.^{5 10 16}

Despite the potential advantages of sharing health research data to maximise secondary analyses, such sharing raises several challenges related to equity and justice. Where the curation and sharing of data is not meaningfully recognised or rewarded, it is likely that the benefits of data sharing will primarily accrue among researchers conducting and publishing secondary analyses.^{17 18} In such contexts, data sharing mandates may exacerbate inequities between researchers conducting primary research and sharing data, and researchers accessing that data and conducting secondary analyses. Due to structural constraints, primary researchers in LMICs can become entrenched in the role of data production, a role which is not meaningfully rewarded under the current system. Meanwhile, secondary researchers in high-income countries (HICs) analyse and publish findings and are rewarded with further funding and career opportunities.¹⁹⁻²¹

In recognition of these challenges, in 2011, the major funders of global health research jointly called for public health data to be shared in ways that are equitable, ethical and efficient.²² Yet, in 2022, a clear concept of equitable research data sharing has still not been defined.²³ This presents a key challenge to developing policies and processes for equitable data sharing in global health research. There is a need for a clear and rigorous concept of equitable data sharing that reflects and incorporates the voices of all stakeholders, including LMIC researchers and communities.

In this paper, we investigate LMIC stakeholders' published views in relation to what comprises equitable data sharing in global health research. Building on the work of Bull *et al*,¹ we conducted a scoping review of the most recent 6 years of published literature on LMIC stakeholders' experiences and perspectives of data sharing in global health research and thematically analysed the identified literature. We report views on how current data sharing practices may exacerbate inequities, what benefits accrue from equitable data sharing, what barriers exist to equitable data sharing, what foundations are required to create an environment conducive to equitable data sharing and what comprises equitable data sharing in global

health research. In light of our findings, we conclude that achieving equitable data sharing requires not only adopting best practices in data sharing but also addressing structural inequalities in global health research.

METHODS

Scoping reviews seek to identify literature relevant to the research objective and may include a variety of article formats.^{24 25} This scoping review sought to identify a published literature on LMIC stakeholders' experiences and perspectives of data sharing in global health research from 2015 onwards. An initial search strategy using a combination of subject headings and text words was developed in consultation with a reference librarian and run through MEDLINE (OvidSP). Papers discussing data sharing issues that specifically mentioned equity/justice/marginalisation and LMICs or global health in the abstract were identified. The keywords from these articles were analysed and incorporated into a revised search strategy, along with keywords from relevant articles. The revised search strategy was then finalised with the librarian (see online supplemental file 1). On the 23 September 2020, the following databases were searched: Embase (OvidSP), Global Health (OvidSP), Global Health (CAB Direct) and Web of Science Core Collection (Clarivate Analytics). The search was limited to studies about humans reported in English and published between 2015 and 23 September 2020.

A total of 1068 references were returned and imported into Covidence, which removed 396 duplicates. The remaining 672 articles were screened by title and abstract according to a matrix of inclusion and exclusion criteria (figure 1). The remaining 86 articles were screened by full text and 15 met the inclusion criteria. A hand search of the included literature identified a further 11 articles. All abstracts and full texts were reviewed by the first author. All full-text articles that were considered 'possible inclusions' by the first author were also reviewed by the last author and a mutual decision made on their inclusion.

Thematic analysis was performed on the included 26 articles in the following five stages: familiarisation with the data, initial coding framework creation, coding, finalisation of the coding framework and recoding of articles for consistency with the final framework.²⁶⁻²⁸ Given the limited pre-existing knowledge of LMIC stakeholders' views on equitable data sharing, an inductive approach to coding was used. An inductive approach allowed us to code the data in a way that was responsive to stakeholders' views, rather than seeking to fit them into an existing framework. The coding framework was coreviewed by the first and last authors before it was finalised. Coding of full-text articles was then performed by the first author using NVivo V.1.5. The five emergent themes identified were: threats to equity from unfair data sharing, benefits of equitable data sharing, barriers to equitable data

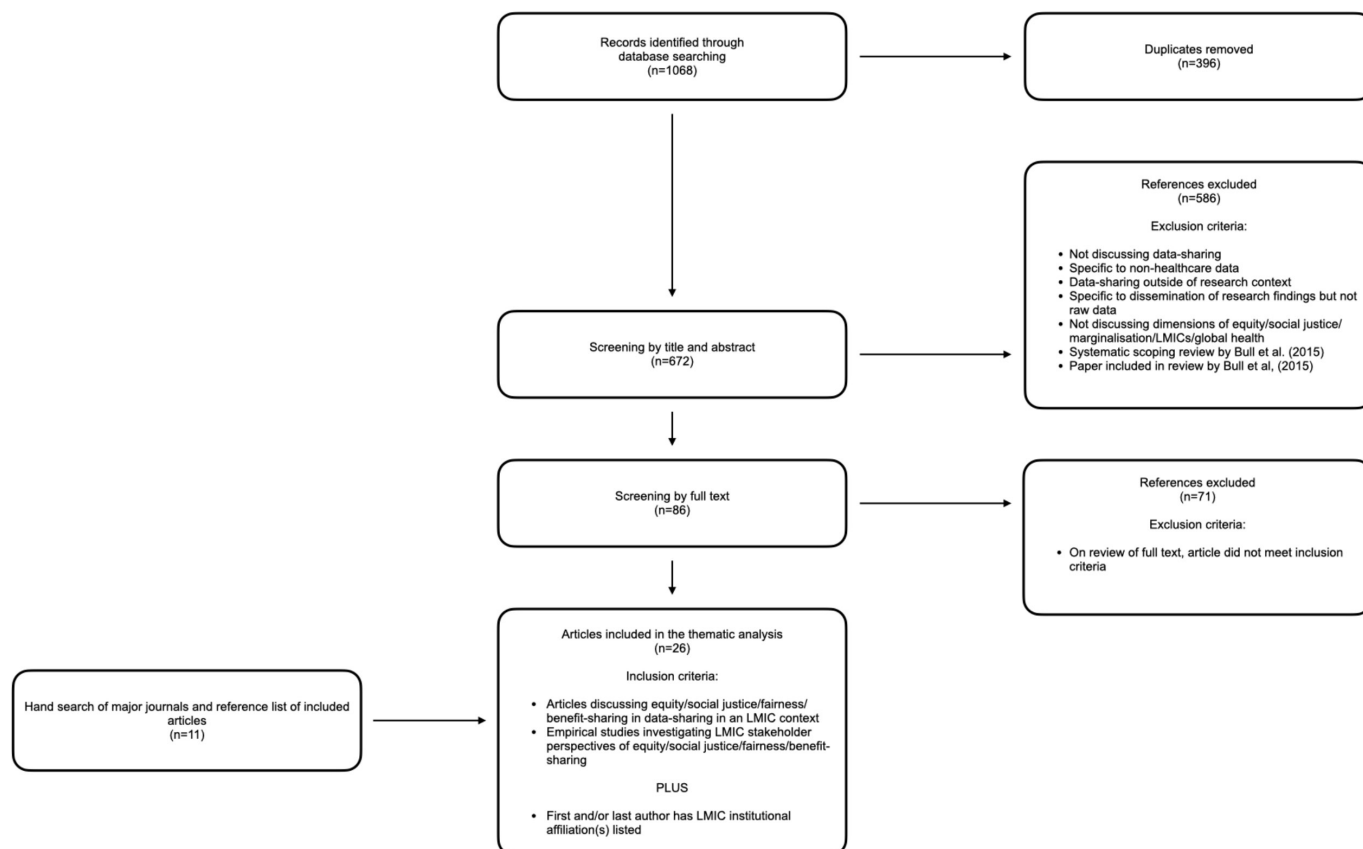


Figure 1 Preferred Reporting Items for Systematic Reviews and Meta-Analyses diagram. LMICs, low-income and middle-income countries.

sharing, foundations for equitable data sharing and features of best practice in equitable data sharing.

Authors' positionality

As in all research, it is important to understand our positionality and, therefore, how our professional and life experiences informed our understanding of the data. The first author conducted the search, coding and analysis of the literature. She is a medical doctor from Australia, with 4 years of prior experience working on empirical ethics projects related to global health research. Those projects raised themes of power, colonial relations and structural barriers to achieving equity.

The second author is a bioethicist currently based in New Zealand, whose research interests have focused on global health ethics since 1999. Over the past decade, she has conducted empirical ethics research on data and sample sharing and biobanking in collaboration with colleagues in Australia, Ghana, Hong Kong, India, Kenya, Malawi, South Africa, Thailand, UK and Vietnam.^{19–21 29–40} She has a longstanding interest in inequities and the exercise of power during the construction of global health data sharing landscapes and associated knowledge production processes.

The study was supervised by the last author, a bioethicist who grew up in the USA and has lived most of her adult life in Australia. She has over 10 years' experience conducting global health ethics research with HIC and

LMIC collaborators. This includes qualitative studies in Uganda, India, Thailand and the Philippines.^{41–44} In conducting her research, she is immersed in studying concepts and issues of equity, power, epistemic justice, social justice and global justice.

Our expertise and experience led us to see the data through a structural lens, which is reflected in our conclusions. However, we also undertook a reflexive approach to data analysis and made a conscious effort to be aware of our positionalities and remain grounded in the data when interpreting the results and the implications of the study. None of us are from LMICs, but we are committed to reducing epistemic injustice within global health.^{45 46} We undertook this study to highlight LMIC voices in relation to data sharing, not our own HIC voices, and that aim guided us throughout.

Patient and public involvement

This scoping review was conducted as a student research project at the University of Melbourne. No primary data was collected as part of this project and it was not appropriate or possible to involve patients or the public in the design, or conduct, or reporting, or dissemination plans.

RESULTS

Literature reviewed

Of the 26 articles, 10 reported the results of qualitative studies of LMIC stakeholders' perspectives on ethical data sharing; 1 reported a quantitative survey; 9 were editorials, commentaries or conceptual analyses; and 6 documented case studies of data sharing (online supplemental file 2). The 17 empirical studies collected views of senior and junior researchers, fieldworkers, research participants, community representatives, members of research ethics committees and members of local regulators and government.^{19–21 31 33 34 47–51} The case studies were based in Argentina, Guatemala, the Dominican Republic, India, Pakistan, Sri Lanka, Thailand, Kenya, Liberia, Sierra Leone, South Africa and Zambia (Some papers featured a number of case studies from different geographical regions.^{52 53}), in addition to 1 that focused on the World-wide Anti-Malarial Resistance Network, which comprises over 280 global collaborators.^{52–57} Of the 26 articles included, 16 were from institutions or projects associated with the Wellcome Trust.^{12 19–21 31 33 34 49 52–55 57–60} In 23 out of the 26 papers, at least 1 author had an HIC institutional affiliation, which potentially reinforces concerns about the dominance of HIC stakeholder perspectives in global health data sharing landscapes. None of the reviewed papers were exclusively authored by researchers with HIC institutional affiliations, although in some cases either the first or last author had such an affiliation.

Concerns about inequity in data sharing

The identified literature described numerous threats to equity that arise due to unfair data sharing (table 1).

Data sharing was described as unfair when it potentially exacerbated inequities between HIC and LMIC researchers, between LMIC research participants and communities and HIC populations and between HIC and LMIC health research systems. The most frequently raised concern centred around the capacity of unfair data sharing practices to reinforce unequal power relations and neocolonial behaviours, where HIC researchers mined data from LMICs without sharing the benefits with LMIC researchers or populations.^{12 31 47 49–53 60–63} Global inequities in the research enterprise and knowledge production mean that many LMIC researchers do not have the time, human or financial resources to devote to secondary analyses. Consequently, the benefits of such sharing will primarily accrue to HIC analysts and institutions in the form of future funding opportunities, career progression and reputation, thereby further compromising opportunities to develop LMIC research infrastructure and personnel.^{20 31 62} Without such opportunities, disparities between HIC and LMIC research systems will remain. Concerns were also expressed that unfair data sharing practices risk widening the research-output gap between HIC and LMIC researchers, perpetuating a dynamic where LMIC researchers are entrenched in the role of data production while HIC researchers

Table 1 Summary of potential benefits and threats to equity from data sharing

Advancing equity through fair data sharing	Threats to equity from unfair data sharing
<p>Benefits to LMIC researchers</p> <ul style="list-style-type: none"> ▶ Researchers receive fair rewards for collection and sharing of data ▶ Fostering equitable collaborations ▶ Increased publications, leading to reputational/career benefits and further funding opportunities 	<p>Threats to LMIC researchers</p> <ul style="list-style-type: none"> ▶ Worsening of unfair power relations with HIC researchers ▶ Opportunity and fiscal cost unfairly burden poorly resourced researchers and institutions; and the rewards are insufficient to offset the burdens ▶ Exclusion from participation in advancing science (LMIC researchers are stuck in a cycle of data production only) ▶ Career and reputational damage leading to loss of funding opportunities ▶ Lack of recognition hampers career progression
<p>Benefits to LMIC participants and communities</p> <ul style="list-style-type: none"> ▶ Secondary research is actualised to local or national public health benefit, leading to improved individual and community health and reduction of the global inequality in disease burdens ▶ Reduced duplication of research reduces the burden of repeated participation in studies on marginalised participants ▶ Education and empowerment of participants during the consent process and inclusion in the creation of data sharing policies 	<p>Threats to LMIC participants and communities</p> <ul style="list-style-type: none"> ▶ Commercial exploitation ▶ Stigmatisation ▶ Benefits of research are not shared with originating communities
<p>Benefits to local health research systems</p> <ul style="list-style-type: none"> ▶ Capacity building of local researchers and research institutions 	<p>Threats to local health research systems</p> <ul style="list-style-type: none"> ▶ Reduced opportunity to develop local research infrastructure and personnel
<p>HIC, high-income country; LMIC, low-income and middle-income country.</p>	

analyse and publish findings.^{19–21} The costs of data sharing were also thought to unfairly burden poorly resourced researchers and institutions without returning commensurate benefits, furthering exacerbating inequities between LMIC and HIC researchers.^{21 31 49 50 52 53 61 62}

Where secondary researchers are removed from the context of the primary research, it was noted that secondary analyses can result in inaccurate conclusions and stereotyped reporting, thus undermining the integrity of the primary research and interrupting its translation to health outcomes in LMICs.^{20 52 59 63} Concerns were also raised that where translational benefits to health outcomes accrued from secondary analysis, they will not be made available in the LMIC communities in which the data were collected.^{34 53 64} These factors were ultimately thought to obstruct the equitable distribution of benefits to originating communities that would otherwise be derived from data sharing.⁵³

Views on benefits of equitable data sharing

The identified literature described numerous benefits of equitable data sharing in global health research (table 1), which, in turn, advance equity between HIC and LMIC researchers, between LMIC populations and HIC populations and between HIC and LMIC health research systems. Key potential benefits of equitable data sharing for LMIC researchers related to capacity, collaboration and recognition. They included fair rewards for LMIC researchers' efforts and equal opportunities for career development, including increased publications and

research funding.^{31 61} Where systemic barriers obstructed capacities to access and analyse shared data (eg, limited analytic capacity due to scarce technological resources, lack of trained staff and funding), their removal was also linked to enhanced productivity and competitiveness of LMIC scientists. This, in turn, can ultimately help to reduce global disparities in research output.^{53 61} Collaborative partnerships between data sharers and data accessors were seen to have a number of benefits for LMIC researchers, including increased opportunities for authorship, skill and knowledge sharing, upskilling in data sharing and improvements to the quality and rigour of research.^{19 31 47} This was thought to reduce the global divergence in research outputs and activities.⁴⁸ Potential benefits of equitable sharing to LMIC participants and communities rested on the translation of secondary research into local or national public health benefits,^{19 31 34 52 59} which can help reduce health disparities between LMIC and HIC populations.

Views on barriers to equitable data sharing

Numerous environmental, personal and relational barriers to equitable data sharing were discussed (table 2).

Environmental barriers, which comprise the structural conditions that prevent equitable data sharing, were the most prominent theme in the literature. All identified sources described a lack of resources and capacity as a barrier to equitable data sharing. Many researchers conduct research in their personal time, aided by personal finances, leaving little time or money for data

Table 2 Summary of barriers to and foundations of equitable data sharing

Barriers preventing equitable data sharing	Foundations of equitable data sharing
Environmental barriers:	Environmental foundations:
Background structural conditions preventing equitable data sharing	Creating robust governance structures
<ul style="list-style-type: none"> ▶ Unequal research systems, resources and capacity between HICs and LMICs ▶ Lack of fiscal resources for primary research and data curation and sharing ▶ Limited local policy or legal frameworks to govern and enforce measures to make data sharing equitable ▶ Lack of enforceability of data sharing agreements and policies ▶ Competitive scientific research environment, and rewards grounded in publication 	<ul style="list-style-type: none"> ▶ Establish frameworks and policies that support equitable data sharing, backed by enforceable governance mechanisms ▶ Governance mechanisms should be: <ul style="list-style-type: none"> – Fair, ethical and accountable – Implemented at all levels: from institutional to international levels – Harmonised across levels to ensure equitable relations across borders – Developed in consultation with community members – Flexible to account for technology advances – Accountable and subject to external review
Relational and personal barriers	Redressing inequities in capacity
<ul style="list-style-type: none"> ▶ Fear of imbalance between burdens and reward ▶ Experience in data sharing ▶ Trust between primary researchers and secondary researchers ▶ Concerns surrounding data sovereignty 	<ul style="list-style-type: none"> ▶ Capacity building of LMIC research systems; provision of resources for sustainable and long-term data sharing
	Relational foundations: Building trust
	<ul style="list-style-type: none"> ▶ Alignment of interests between primary and secondary researchers ▶ Institutional policies supporting equitable data sharing ▶ Long-term relationships and collaborative partnerships with secondary researchers
HICs, high-income countries; LMICs, low-income and middle-income countries.	

sharing.^{47 48 50 50 61} Lack of secure access to the internet and computers, inability to access library resources off-site, insufficient funds for software and the professional memberships and subscriptions required for effective participation in Open Science, and poor digital and data science literacy exclude LMIC researchers from meaningfully participating in data sharing.^{31 48 51 61 63} A lack of core infrastructure funding from institutions, governments and funders means researchers have little way to tackle these barriers.⁵⁰ Furthermore, managing large datasets and access requests often requires the employment and training of dedicated staff, and long-term funding for the maintenance of data and biobanks is scarce.^{53 55}

Another frequently discussed barrier to equitable data sharing was the scientific environment, where rewards are predominantly grounded in publication but curating and sharing data alone is not grounds for authorship. LMIC stakeholders further reported additional pressures from working multiple roles across teaching and/or medicine, with little time or institutional support for research.⁵⁰ Thus, the burden of data collection is high, while publication outputs are low.^{51 61} Data sharing was seen to exacerbate these existing pressures. LMIC researchers also feared that requirements to publish data before they could perform analyses put them at risk of being scooped by HIC researchers.^{31 50 51 62 63}

In some LMIC settings, there is a lack of robust regulatory and governance structures to support equitable data sharing.⁵⁸ Existing ethical frameworks and case studies were thought to primarily address HIC contexts and did not reflect the difficulties LMIC researchers faced.⁵³ A lack of enforceable agreements leaves LMIC researchers exposed to exploitation and fails to safeguard participants from harms.^{53 64} The lack of relevant guidance on fair data sharing processes and the ethical dilemmas raised by unfair processes was thought to stagnate action to implement equitable data sharing.³⁴

Personal and relational barriers discouraged researchers from sharing data but did not preclude it. Such barriers included a lack of experience with data sharing,^{19 21 61} lack of trust in secondary researchers,^{21 50 51 61} data sovereignty concerns^{21 31 47 49 50} and fear of the imbalance between burdens and rewards created by inequitable data sharing practices (table 2).^{21 49 52 61 62}

Views on foundations for equitable data sharing

The identified literature revealed three key environmental and relational foundations for equitable data sharing: redressing capacity inequities, building trust and creating robust governance and policy structures (table 2). The first included unequal research system capacity between HICs and LMICs across technological and human resources, including expertise in data curation, management and analysis.^{31 60} Significant investment is required to build both technological and human capacity in LMICs to enable equitable data sharing.^{31 51 53 63 65} Dedicated funding for building capacity for data sharing at institutional or national levels,

rather than at the project level, resources for sustainable and long-term data sharing (such as funding data access committees and data managers) as well as flexible microfinancing programmes (Microfinancing programs provide very small sums, are easily applied for, and flexible in how they might be spent.) were among the strategies put forward to address this gap.^{31 50 51 59} Building trust occurs through developing long-term relationships and establishing fair and collaborative partnerships between primary and secondary researchers.^{19 21 31 34} Alignment of interests between primary and secondary researchers and accessing institutions that had their own policies supporting equitable data sharing were seen as important factors that built trust.^{21 31}

Most authors called for governance mechanisms to be implemented to ensure fair processes and to protect the interests of LMIC researchers and participants.^{12 19 21 49 53 55 55 63 64} They recommended data sharing policies and agreements be enforceable to encourage adherence and to foster trust between primary and secondary researchers.^{19 33 34 47 53 61} Governance structures are required at all levels, from the institutional level to the international level, and are largely lacking in the current data sharing environment.^{12 19–21 31 47 49 53 55 57 58 60 63 64} Community engagement, LMIC researcher engagement, transparency and accountability to external review were considered important in ensuring that governance structures operated ethically and did not perpetuate an exploitative or neocolonial dynamic.^{33 58 61 63}

Views on best practices to promote equitable data sharing

Views about best practices to promote equitable data sharing spanned four themes: advancing recognition and reciprocity (via fair benefit and burden sharing); providing protections for researchers, participants and communities; upholding obligations; and data sovereignty. Data sharing policies and practices should be appropriate to the content and type of data collected, curated and shared. This analysis encompasses the breadth of contexts addressed in the reviewed literature and not all recommendations will be relevant to all types of data sharing.

Benefit sharing with primary researchers was thought to be an important dimension of equitable data sharing.^{19 34 47 51 57 61 63} Most stakeholders called for authorship or another, equal form of academic recognition to balance the burdens of data sharing.^{12 21 31 34 47 49 50 52 53 58 59 61 63} Acknowledgement of primary researchers in publications by secondary researchers was not considered sufficient recognition or reward. Other benefits to be shared with primary researchers included professional promotions, partnerships for mutual exchanges of data, training and skills sharing.^{52 53 61}

Benefit sharing with participants or originating communities was also frequently identified as an important aspect of equitable data sharing,^{19–21 31 33 47 53 58 60 61 63 64} as was prioritising secondary research with the potential

to translate to local health benefits for originating communities.^{19–21 33 61 63} In recognition of the inequities in access to health resources in originating communities, stakeholders felt that data sharing practices should not contribute to widening these differences, but instead should work to promote global health equity.³³ Accordingly, it was often felt that secondary research responsive to the needs of originating communities and/or that has translational potential should be prioritised.^{19–21 33 61 63} Developing secondary research agendas in partnership with communities and local health policy-makers was seen as a way to promote research with the potential to benefit communities.^{19 53} Where this was not possible, stakeholders concluded that benefits should at least be shared with originating communities by making the results of secondary analyses available to local researchers, communities and health ministries.^{58 60 63} Some community representatives were supportive of health benefits accruing to the wider public, rather than directly back to their communities.¹⁹ This was in recognition of the fact that health research data, especially when publicly funded, constitutes a public good. Stakeholders further asserted that commercial outputs of secondary research (such as vaccines) should be made available to originating communities.^{60 64}

As secondary researchers benefit from data sharing, some authors felt they should also share the burdens created by such sharing in line with the principle of reciprocity.^{12 21 47} Sharing the cost of managing the dataset was the main form of burden sharing identified. This was seen as particularly relevant in the case of legacy datasets, where costs incurred in the preparation and sharing of data for secondary use are not provided for by the initial funding.¹²

Protections for LMIC stakeholders focused on minimising harms from data sharing. For researchers, these included implementing managed access to data and embargo periods to allow lower-resourced LMIC researchers to publish articles from their data first.^{12 19–21 31 33 47 49 51 52 54 61 62} For participants and communities, minimising harms focused on protecting privacy, preventing stigma and reducing misinterpretation of data during secondary analyses.^{19 20 49 52 53 58 60 64} Proposed means for protecting the community and participant interests were twofold:

1. Placing additional regulations on non-local data-access requests. (We note that some stakeholders preferred to encourage local collaborations to build regional research capacity, while others preferred international collaborations for the opportunity of skills sharing and to enhance their reputation.)^{19 20 49 52}
2. Community engagement to determine what constitutes sensitive data, what secondary uses might lead to harm or stigma and what regulations/limitations should be in place for secondary access.^{52 53}

Although some stakeholders called for open access to data, many considered this a potentially harmful policy.^{20 21 31 33 47 49 61 63 64} While open access approaches

allow for maximum utility and transparency, authors and most stakeholders felt that the potential harms to LMIC researchers and participants (table 1) outweighed this benefit.

Secondary researchers, journals and funders were thought to have obligations to facilitate equitable data sharing processes. In relation to secondary researchers, this included a duty to contribute to building LMIC research capacity.^{19 21 31 33 34 47 53 61} Where possible, secondary researchers were expected to proactively invite primary researchers to collaborate in designing and conducting secondary analyses.^{20 47 52 53} It was acknowledged, however, that collaboration is not always needed nor possible and, in such instances, secondary researchers should make efforts to fairly share benefits with primary researchers in terms of recognition and with originating communities by returning results.^{47 63} Journals were also seen to carry obligations, including making it a condition of publication that results of secondary analyses are disseminated to originating communities, primary researchers, other researchers and ministries of health where the data was collected.^{53 63} Lastly, if funders are to mandate data sharing, it was suggested they should invest in the required technology and human resources to manage the dataset, as well as building analytical capacity, so that LMIC researchers can also perform secondary analyses of shared data.^{31 47 52 59 63}

Some models of data sovereignty were thought to advance equitable data sharing, whereas others were thought to obstruct it. A custodianship model places an onus on the custodian to honour their ethical obligations to data subjects as well as to the best interests of research and the broader public, thus balancing the interests of data producers, data subjects and the role of data for public good.^{20 31 47 64} Other stakeholders preferred a model based on ownership and intellectual property rights. Ownership rights were viewed as a foundation for academic recognition and as the basis to restrict access to data if required.^{20 21 49}

Collectively, the reviewed literature suggests that building enabling foundations and implementing practices to promote equitable sharing have the potential to address the barriers and concerns raised by LMIC stakeholders (online supplemental file 3). However, none of the enabling foundations combat the environmental barrier presented by the competitiveness of the research environment, where rewards are predominantly grounded in publication.

DISCUSSION

This review found that data sharing under existing mandates to share data (with minimal restrictions) risks perpetuating inequities between researchers and countries.^{50 61} Rich data production in LMICs does not necessarily translate to rich rewards, due to the lack of local research capacity and infrastructure.⁶¹ A lack of adequate protections risks perpetuating a neocolonial

dynamic whereby raw data is mined from LMICs by HIC researchers, who then receive the rewards in terms of publications and improved public health in their own countries.²⁰ The current data sharing regime, which gives primacy to utility, is thus reinforcing asymmetries in power and privilege and relations of coloniality that are inherent in global health. In recent years, such unfair power dynamics have been the focus of renewed attention in global health research and practice more broadly.^{45 66–68} In such a climate, mandatory data sharing can be inherently punitive to LMIC researchers who must comply with funder requirements or risk losing access to their research support.⁶¹ Anane-Sarpong further posits that, given the inequities between LMIC and HIC researchers, creating a norm of data sharing without addressing the inequities between LMIC and HIC researchers that is greater than so far laid out in the literature.⁶¹

The way forward, as proposed in the reviewed literature, is to prioritise equity in data sharing. There was an overarching call for equitable rather than open data sharing.^{20 21 31 33 47 49 61 63 64} Yet, there is a continued lack of understanding and guidance on equitable data sharing. Current data sharing policies and guidelines also often marginalise or do not reflect LMIC voices and fail to address structural inequities in knowledge production landscapes. Moving forward, as a matter of procedural fairness and epistemic justice, we must ensure that any concept of equitable data sharing that is developed ultimately reflects and incorporates the voices of LMIC researchers and communities (as well as other stakeholders).

Published views suggest that LMIC stakeholders believe that, while equitable data sharing should encompass specific practices to implement at the research collaboration/project level, it cannot be achieved without addressing structural inequities within global health research. Equitable data sharing requires fixing the grossly unequal system of global health knowledge production. Key long-term approaches to address structural inequities include: redressing capacity inequities in data curation, data sharing and the analysis of shared data to provide greater opportunities for LMIC researchers to lead secondary analyses and become lead authors on resulting publications; creating robust governance and policy structures at international, national and institutional levels; and building trust, which is consistent with findings of the earlier review by Bull *et al.*¹ Significant resources are needed to support the collection, curation and effective sharing of high-quality data to advance global health.^{69 70} Beyond these approaches, structural changes are needed to alter how researchers are rewarded in the research environment, including rewards predominantly grounded in peer-reviewed publications.

At the research collaboration/project level, LMIC stakeholders identified key practices to promote equity in data sharing, including appropriate recognition of the work by primary researchers, reciprocity involving fair sharing of benefits and burdens, capacity building collaborations

between primary and secondary researchers, and appropriate funding for data sharing activities.⁵³ Some of these key practices may also be best supported or made easier by structural changes. For example, adequate recognition could be facilitated by reform of how funders and research institutions evaluate and reward researchers.^{71–73} In this review, LMIC researchers overwhelmingly called for authorship of publications reporting the results of secondary analyses of shared data in order to appropriately recognise the burdens of gathering and maintaining a dataset.^{12 21 31 34 47 49 50 52 53 58 59 61 63} Globally, current promotion criteria and funding for future research remain closely tied to publication in peer-reviewed journals.^{31 50 51 61} Hence, authorship credit is seen as an important element of benefit sharing. Other forms of recognition and acknowledgement in publications were considered insufficient reward for data sharing.

However, we suggest that the needed reform here may not be for journals to alter authorship criteria. Funders and research institutions could incorporate additional metrics beyond quantity, quality and citation impact of publications such as the number of studies using shared data, similar to benchmarks for citations. Such approaches have the potential to directly recognise and reward the sharing of high-quality data,^{71 72} and highlight the responsibilities of multiple stakeholders to promote effective mechanisms for appropriate recognition throughout data sharing cycles.

Building enabling foundations is associated with significant financial costs and can take considerable time. Yet, not pursuing structural change also poses a risk to the utility of research and to public health gains in LMICs: heavy burdens and lack of capacity perpetuates low scientific productivity in the region, hampering the production of new knowledge and slowing progress toward improved health outcomes. In such contexts, onerous data sharing obligations encourage sharing technically unusable data that contributes little or nothing towards reducing health burdens and reduces the utility of shared data.^{19 57 59 61 70} Supporting a longer term effective global health research enterprise and LMIC knowledge production is thus best served by addressing structural inequities and promoting capacity to address local disease burdens via primary and secondary research.

This review was limited to academic peer-reviewed literature and English-language publications. Views from LMICs in South America and the Middle East were lacking. Further exploration is required into views of stakeholders from these locations and from non-English language speakers. The review was also biased by the predominance of contributions to the identified literature being affiliated with the Wellcome Trust, a major funder of global health research. Future research could usefully explore best practices in equitable data sharing in global health research funded by other entities. It could also review and analyse the opinions and conceptions of equitable data sharing expressed by LMIC stakeholders in grey literature and other formats. This review

only captured a published, peer-reviewed literature. Going beyond, the peer-reviewed literature is required to ensure LMIC stakeholders' experiences, perspectives and interests inform the development of equitable data sharing frameworks. Qualitative studies may also be important to capture diverse and less heard viewpoints.

This study did not include LMIC collaborators. Working with them would likely have enriched our analysis. The study was conducted as part of a 3-month student research project to carry out a scoping review of the ethics literature on data sharing from 2015 onwards. The first author chose to focus data analysis on literature within that dataset written by LMIC authors because she wanted to concentrate on less heard voices in data sharing. This occurred mid-way through the project and, at the time, due to a lack of resources to compensate an LMIC collaborator and the project's design being complete, we felt uncomfortable bringing another collaborator on board. However, an LMIC researcher's perspective could have further contextualised our analysis and thus provided further insights, including a more nuanced understanding of the structural challenges facing LMIC researchers in relation to equitable data sharing. During analysis and writing up, the second author drew on her experience of conducting research in this field to identify areas where the initial interpretation of findings warranted additional consideration and discussion by all authors, including reflections on positionality.

Ultimately, the study draws attention to the importance of structural change in enabling equitable data sharing. Similarly, there are growing calls for making structural changes in the broader global health enterprise in order to decolonise the field and address unfair power dynamics and epistemic injustices.⁷⁴ As such, it is imperative that future discussions and efforts to make structural changes in global health identify and encompass structural changes that are needed to ensure researchers worldwide are able to generate, curate, share and access data for secondary analyses to address pressing global health problems. To promote epistemic justice and procedural fairness, such work should meaningfully engage relevant data sharing actors, including those who are often marginalised. To achieve equitable data sharing, it must be part of a broader ongoing and inclusive conversation about redressing structural inequities in global health.

A key area for further investigation is mapping out the next steps to obtain a clear conception of equitable data sharing. When it comes to secondary uses of data, there are a number of actors, including LMIC stakeholders, with potentially competing priorities and interests that need to be recognised and respected. A conception of equitable data sharing that fairly balances these interests needs to be elucidated in order to advance progress toward equitable data sharing in global health.

Acknowledgements The authors would like to thank Jim Berryman at the University of Melbourne Brownless Biomedical Library for his assistance with creating the search strategy.

Contributors BP and NE conceived of the study and its methods. NE collected and analysed the data, under the supervision of BP and with assistance from SB. NE wrote the first draft of the manuscript. BP and SB revised the work critically for intellectual content. All authors gave final approval of the version to be published. BP accepts full responsibility for the work and/or the conduct of the study, had access to the data, and controlled the decision to publish.

Funding This research was partially supported by a Wellcome Trust Strategic Award (096527) and a Department for International Development/Wellcome—Epidemic Preparedness—Coronavirus Grant (221559/Z/20/Z). For the purpose of open access, SB has applied a CC BY public copyright licence to any Author Accepted Manuscript version arising from this submission. BP was supported by a University of Melbourne R Douglas Wright Research Fellowship.

Competing interests None declared.

Patient and public involvement Patients and/or the public were not involved in the design, or conduct, or reporting, or dissemination plans of this research.

Patient consent for publication Not applicable.

Ethics approval Not applicable.

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Data are available upon reasonable request from the corresponding author.

Supplemental material This content has been supplied by the author(s). It has not been vetted by BMJ Publishing Group Limited (BMJ) and may not have been peer-reviewed. Any opinions or recommendations discussed are solely those of the author(s) and are not endorsed by BMJ. BMJ disclaims all liability and responsibility arising from any reliance placed on the content. Where the content includes any translated material, BMJ does not warrant the accuracy and reliability of the translations (including but not limited to local regulations, clinical guidelines, terminology, drug names and drug dosages), and is not responsible for any error and/or omissions arising from translation and adaptation or otherwise.

Open access This is an open access article distributed in accordance with the Creative Commons Attribution 4.0 Unported (CC BY 4.0) license, which permits others to copy, redistribute, remix, transform and build upon this work for any purpose, provided the original work is properly cited, a link to the licence is given, and indication of whether changes were made. See: <https://creativecommons.org/licenses/by/4.0/>.

REFERENCES

- 1 Bull S, Roberts N, Parker M. Views of ethical best practices in sharing individual-level data from medical and public health research: a systematic scoping review. *J Empir Res Hum Res Ethics* 2015;10:225–38.
- 2 Hughes S, Wells K, McSorley P, et al. Preparing individual patient data from clinical trials for sharing: the GlaxoSmithKline approach. *Pharm Stat* 2014;13:179–83.
- 3 Rathi V, Dzara K, Gross CP, et al. Sharing of clinical trial data among trialists: a cross sectional survey. *BMJ* 2012;345:e7570.
- 4 Rodwin MA, Abramson JD. Clinical trial data as a public good. *JAMA* 2012;308:871–2.
- 5 Open Access Policy. Bill & melinda gates foundation. Available: <https://www.gatesfoundation.org/about/policies-and-resources/open-access-policy> [Accessed 4 Nov 2021].
- 6 EFPIA-PhRMA. Principles successfully enable responsible clinical trial data sharing. Available: <https://www.efpia.eu/news-events/the-efpia-view/blog-articles/29112017-efpia-pharma-principles-successfully-enable-responsible-clinical-trial-data-sharing/> [Accessed 4 Nov 2021].
- 7 NIH data sharing policies. Available: https://www.nlm.nih.gov/NIHbmic/nih_data_sharing_policies.html [Accessed 4 Nov 2021].
- 8 Nature Portfolio. Reporting standards and availability of data, materials, code and protocols. Available: <https://www.nature.com/nature-portfolio/editorial-policies/reporting-standards> [Accessed 4 Nov 2021].
- 9 PLoS One. Data availability. Available: 2019. <https://journals.plos.org/plosone/s/data-availability> [Accessed 4 Nov 2021].
- 10 Taichman DB, Sahni P, Pinborg A, et al. Data sharing statements for clinical trials. *JAMA* 2017;317:2491.
- 11 Wellcome. Our new policy on sharing research data: what it means for you. Available: <https://wellcome.org/news/our-new-policy-sharing-research-data-what-it-means-you> [Accessed 4 Nov 2021].

- 12 Waithira N, Mutinda B, Cheah PY. Data management and sharing policy: the first step towards promoting data sharing. *BMC Med* 2019;17:80.
- 13 Hrynaskiewicz I, Altman DG. Towards agreement on best practice for publishing raw clinical trial data. *Trials* 2009;10:17.
- 14 Ross JS, Lehman R, Gross CP. The importance of clinical trial data sharing: toward more open science. *Circ Cardiovasc Qual Outcomes* 2012;5:238–40.
- 15 Sim I, Stebbins M, Bierer BE, et al. Time for NIH to lead on data sharing. *Science* 2020;367:1308–9.
- 16 Global Research Collaboration for Infectious Disease Preparedness (GloPID-R). Principles of data sharing in public health emergencies. 2018. Available: <https://www.glopid-r.org/wp-content/uploads/2018/06/glopid-r-principles-of-data-sharing-in-public-health-emergencies.pdf>
- 17 Langat P, Pisartchik D, Silva D, et al. Is there a duty to share? ethics of sharing research data in the context of public health emergencies. *Public Health Ethics* 2011;4:4–11.
- 18 Zarin DA. Participant-level data and the new frontier in trial transparency. *N Engl J Med* 2013;369:468–9.
- 19 Jao I, Kombe F, Mwalukore S, et al. Research stakeholders' views on benefits and challenges for public health research data sharing in Kenya: the importance of trust and social relations. *PLOS ONE* 2015;10:e0135545.
- 20 Denny SG, Silaigwana B, Wassenaar D, et al. 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* 2015;10:290–301.
- 21 Hate K, Meherally S, Shah More N, et al. Sweat, skepticism, and uncharted Territory: a qualitative study of opinions on data sharing among public health researchers and research participants in Mumbai, India. *J Empir Res Hum Res Ethics* 2015;10:239–50.
- 22 Carr D, Littler K. Sharing research data to improve public health: a funder perspective. *J Empir Res Hum Res Ethics* 2015;10:314–6.
- 23 Kalkman S, Mostert M, Gerlinger C, et al. Responsible data sharing in international health research: a systematic review of principles and norms. *BMC Med Ethics* 2019;20:21.
- 24 Arksey H, O'Malley L. Scoping studies: towards a methodological framework. *Int J Soc Res Methodol* 2005;8:19–32.
- 25 Armstrong R, Hall BJ, Doyle J, et al. Cochrane update. "scoping the scope" of a cochrane review. *J Public Health (Oxf)* 2011;33:147–50.
- 26 Braun V, Clarke V. Using thematic analysis in psychology. *Qualitative Research in Psychology* 2006;3:77–101.
- 27 Gale NK, Heath G, Cameron E, et al. Using the framework method for the analysis of qualitative data in multi-disciplinary health research. *BMC Med Res Methodol* 2013;13:117.
- 28 Hsieh H-F, Shannon SE. Three approaches to qualitative content analysis. *Qual Health Res* 2005;15:1277–88.
- 29 Bull S, Cheah PY, Denny S, et al. Best practices for ethical sharing of individual-level health research data from low- and middle-income settings. *J Empir Res Hum Res Ethics* 2015;10:302–13.
- 30 Bull S. *Ensuring global equity in open research*. London: Wellcome Trust, 2016.
- 31 Cheah PY, Tangseefa D, Somsaman A, et al. 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. *J Empir Res Hum Res Ethics* 2015;10:278–89.
- 32 Cheah PY, Jatupornpimol N, Hanboonkunapakarn B, et al. Challenges arising when seeking broad consent for health research data sharing: a qualitative study of perspectives in thailand. *BMC Med Ethics* 2018;19:86.
- 33 Jao I, Kombe F, Mwalukore S, et al. 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. *J Empir Res Hum Res Ethics* 2015;10:264–77.
- 34 Merson L, Phong TV, Nhan LNT, et al. Trust, respect, and reciprocity: informing culturally appropriate data-sharing practice in Vietnam. *J Empir Res Hum Res Ethics* 2015;10:251–63.
- 35 Mwapasa M, Gooding K, Kumwenda M, et al. "are we getting the biometric bioethics right?" -the use of biometrics within the healthcare system in Malawi. *Glob Bioeth* 2020;31:67–80.
- 36 Parker M, Bull SJ, de Vries J, et al. Ethical data release in genome-wide association studies in developing countries. *PLoS Med* 2009;6:e1000143.
- 37 Tindana P, Molyneux S, Bull S, et al. Ethical issues in the export, storage and reuse of human biological samples in biomedical research: perspectives of key stakeholders in Ghana and Kenya. *BMC Med Ethics* 2014;15:76.
- 38 Tindana P, Molyneux S, Bull S, et al. "it is an entrustment": broad consent for genomic research and biobanks in sub-Saharan Africa. *Dev World Bioeth* 2019;19:9–17.
- 39 Vaz M, Warriar P, Wai-Loon Ho C, et al. Respecting values and perspectives in biobanking and genetic research governance: outcomes of a qualitative study in Bengaluru, India [version 1; peer review: awaiting peer review]. *Wellcome Open Res* 2022;7:78.
- 40 Warriar P, Ho C-L, Bull S, et al. n.d. Engaging publics in biobanking and genetic research governance - a literature review towards informing practice in India [version 2; peerreview: 1 approved]. *Wellcome Open Res*;6:5.
- 41 Pratt B, Hyder AA. Linking participatory action research on health systems to justice in global health: A case study of the maternal and neonatal implementation for equitable health systems project in rural Uganda. *J Empir Res Hum Res Ethics* 2018;13:74–87.
- 42 Pratt B, Zion D, Lwin KM, et al. Linking international clinical research with stateless populations to justice in global health. *BMC Med Ethics* 2014;15:49.
- 43 Pratt B. Sharing power in global health research: an ethical toolkit for designing priority-setting processes that meaningfully include communities. *Int J Equity Health* 2021;20:127.
- 44 Pratt B, Seshadri T, Srinivas PN. Overcoming structural barriers to sharing power with communities in global health research priority-setting: lessons from the participation for local action project in Karnataka, India. *Glob Public Health* 2022;17:3334–52.
- 45 Atuire CA, Bull S. COVID-19 heightens the imperative to decolonize global health research. *GJN* 2022;13:60–77.
- 46 Pratt B, de Vries J. Where is knowledge from the global south? an account of epistemic justice for a global bioethics. *J Med Ethics* 2023;medethics-2022-108291.
- 47 Anane-Sarpong E, Wangmo T, Ward CL, et al. "You cannot collect data using your own resources and put it on open access": perspectives from Africa about public health data-sharing. *Dev World Bioeth* 2018;18:394–405.
- 48 Bezuidenhout L, Kelly AH, Leonelli S, et al. \$100 is not much to you': open science and neglected accessibilities for scientific research in Africa. *Critical Public Health* 2017;27:39–49.
- 49 Kaewkungwal J, Adams P, Sattabongkot J, et al. Issues and challenges associated with data-sharing in IMICS: perspectives of researchers in Thailand. *Am J Trop Med Hyg* 2020;103:528–36.
- 50 Rappert B, Bezuidenhout L. Data sharing in low-resourced research environments. *Prometheus* 2016;34:207–24.
- 51 Bezuidenhout L, Chakauya E. Hidden concerns of sharing research data by low/middle-income country scientists. *Glob Bioeth* 2018;29:39–54.
- 52 Barnes KI, Canario JA, Vernekar SS, et al. Equitable data sharing: challenges and suggestions for ways forward. *Wellcome Open Res* 2019;4:172.
- 53 Fernando B, King M, Sumathipala A. Advancing good governance in data sharing and biobanking-international aspects. *Wellcome Open Res* 2019;4:184.
- 54 Cheah PY, Day NPJ. Data sharing: experience from a tropical medicine research unit. *Lancet* 2017;390.
- 55 Cheah PY, Day NPJ, Parker M, et al. Sharing individual-level health research data: experiences, challenges and a research agenda. *Asian Bioeth Rev* 2017;9:393–400.
- 56 Gorina Y, Redd JT, Hersey S, et al. Ensuring ethical data access: the Sierra Leone Ebola database (SLED) model. *Ann Epidemiol* 2020;46:1–4.
- 57 Humphreys GS, Tinto H, Barnes KI. Strength in numbers: the WWARN case study of purpose-driven data sharing. *Am J Trop Med Hyg* 2019;100:13–5.
- 58 Bull S, Bhagwandin N. The ethics of data sharing and biobanking in health research. *Wellcome Open Res* 2020;5:270.
- 59 Merson L, Gaye O, Guerin PJ. Avoiding data dumpsters -- toward equitable and useful data sharing. *N Engl J Med* 2016;374:2414–5.
- 60 Vaz M, Palmero AG, Nyangulu W, et al. Diffusion of ethical governance policy on sharing of biological materials and related data for biomedical research. *Wellcome Open Res* 2019;4:170.
- 61 Anane-Sarpong E, Wangmo T, Tanner M. Ethical principles for promoting health research data sharing with sub-Saharan Africa. *Dev World Bioeth* 2020;20:86–95.
- 62 Hayes R, Ayles H, Binka F, et al. Sharing clinical trial data. *Lancet* 2016;387.
- 63 Serwadda D, Ndebele P, Grabowski MK, et al. Open data sharing and the global south—who benefits? *Science* 2018;359:642–3.
- 64 Andanda P. Towards a paradigm shift in governing data access and related intellectual property rights in big data and health-related research. *IJC* 2019;50:1052–81.

- 65 Akintola SO. Legal implications of data sharing in biobanking research in low-income settings: the Nigerian experience. *S Afr J BL* 2018;11:15.
- 66 Abimbola S, Asthana S, Montenegro C, *et al*. Addressing power asymmetries in global health: imperatives in the wake of the COVID-19 pandemic. *PLoS Med* 2021;18:e1003604.
- 67 Büyüm AM, Kenney C, Koris A, *et al*. Decolonising global health: if not now, when? *BMJ Glob Health* 2020;5:e003394.
- 68 Parker M, Kingori P. Good and bad research collaborations: researchers' views on science and ethics in global health research. *PLoS One* 2016;11:e0163579.
- 69 Open research [Wellcome]. Available: <https://wellcome.org/what-we-do/our-work/open-research> [Accessed 19 May 2022].
- 70 Pisani E, Aaby P, Breugelmans JG, *et al*. Beyond open data: realising the health benefits of sharing data. *BMJ* 2016;355:i5295.
- 71 NEJM. Data authorship as an incentive to data sharing. Available: <https://www.nejm.org/doi/full/10.1056/NEJMs1616595> [Accessed 2 Feb 2022].
- 72 Kalager M, Adami H-O, Bretthauer M. Recognizing data generation. *N Engl J Med* 2016;374:1898.
- 73 FORCE11. 2013. Available: <https://force11.org/info/joint-declaration-of-data-citation-principles-final/>
- 74 Bhakuni H, Abimbola S. Epistemic injustice in academic global health. *Lancet Glob Health* 2021;9:e1465–70.

Supplement 1. Final search strategy

Global Health Database

	Searches	Results
1	(title:(justice or fairness or fair or ethics or ethical or equit*) OR ab:(justice or fairness or fair or ethics or ethical or equit*) OR ab:(benefit and sharing) AND yr:[2016 TO 2020]) AND (title:(health or medical) OR ab:(health or medical) AND yr:[2016 TO 2020]) AND (((title:((data or database* or dataset*) and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)) OR title:(information and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)) OR title:(secondary analysis) OR title:(finding* and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)) OR title:(result* and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)) OR title:(data and research))) AND yr:[2016 TO 2020])	68

Embase (Ovid SP) and Medline (Ovid SP)

	Searches	Results
1	((data or database* or dataset*) and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)).ti.	24500
2	((information and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)) or "secondary analysis").ti.	19385
3	(finding* and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)).ti.	5202
4	(result* and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership)).ti.	21244
5	(data or database or research or science).ti.	1165568
6	((data or database* or research or science) and ("open access" or open)).ti,ab.	339103

7	(justice or fairness or fair or ethics or ethical or equit* or "benefit sharing").ti,ab.	472661
8	1 or 2 or 3 or 4 or 6	405201
9	(health or medical).mp.	10488669
10	5 and 7 and 8 and 9	1319
11	10	1319
12	limit 11 to english language	1268
13	limit 12 to humans	1035
14	limit 13 to yr="2015 - 2021"	616

Web of Science

	Search	Results
# 1	ti= ((data or database* or dataset*) and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	53,518
# 2	ti=(information and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership) or "secondary analysis") Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	33,586
# 3	ti=(finding* and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	3,611

# 4	ti= (result* and (share* or sharing or release* or releasing or disseminat* or distribut* or export* or recycl* or access* or reuse* or management or ownership) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	15,654
# 5	ti= (data or database or research or science) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	1,797,781
# 6	ts= ((data or database* or research or science) and ("open access" or open) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	426,849
# 7	ts=(justice or fairness or fair or ethics or ethical or equit* or "benefit sharing") Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	571,310
# 8	#1 or #2 or #3 or #4 or #6 Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	526,569
# 9	ts= (health or medical) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	3,256,404
# 10	#5 and #7 and #8 and #9 Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=All years	711
# 11	(#10) AND LANGUAGE: (English) Indexes=SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC Timespan=2016-2020	417

#12	<p data-bbox="357 297 1141 331">(#10) AND LANGUAGE: (English)</p> <p data-bbox="357 365 1141 790">Refined by: [excluding] WEB OF SCIENCE CATEGORIES: (LINGUISTICS OR MATHEMATICAL COMPUTATIONAL BIOLOGY OR ENVIRONMENTAL SCIENCES OR CRIMINOLOGY PENOLOGY OR BIODIVERSITY CONSERVATION OR CELL TISSUE ENGINEERING OR REGIONAL URBAN PLANNING OR CHEMISTRY ANALYTICAL OR TELECOMMUNICATIONS OR CHEMISTRY APPLIED OR TRANSPORTATION OR URBAN STUDIES OR ENGINEERING INDUSTRIAL OR ENVIRONMENTAL STUDIES OR GEOGRAPHY PHYSICAL OR ENGINEERING ELECTRICAL ELECTRONIC OR ENGINEERING MULTIDISCIPLINARY OR GREEN SUSTAINABLE SCIENCE TECHNOLOGY)</p> <p data-bbox="357 824 1141 925">Timespan: 2016-2020. Indexes: SCI-EXPANDED, SSCI, A&HCI, CPCI-S, CPCI-SSH, BKCI-S, BKCI-SSH, ESCI, CCR-EXPANDED, IC.</p>	384
------------	--	-----

Supplement 2. Articles included in the review

Year	Author	Article type	Location of institutional affiliation: first author	Location of institutional affiliation: last author
2015	Cheah PY, Tangseefa D, Somsaman A, Chunsuttiwat T, Nosten F, Day NPJ, Bull S, Parker M.	Qualitative study	United Kingdom/Thailand	United Kingdom
2015	Denny SG, Silaigwana B, Wassenaar D, Bull S, Parker M.	Qualitative study	South Africa	United Kingdom
2015	Hate K, Meherally S, More NS, Jayaraman A, Bull S, Parker M, Osrin D.	Qualitative study	India	United Kingdom
2015	Jao I, Kombe F, Mwalukore S, Bull S, Parker M, Kamuya D, Molyneux S, Marsh V.(a)	Qualitative study	Kenya	Kenya/United Kingdom
2015	Jao I, Kombe F, Mwalukore S, Bull S, Parker M, Kamuya D, Molyneux S, Marsh V.(b)	Qualitative study	Kenya	Kenya/United Kingdom
2015	Merson L, Phong TV, Nhan LNT, Dung NT, Ngan TTD, Kinh NV, Parker M, Bull S..	Qualitative study	United Kingdom/Vietnam	United Kingdom

2016	Hayes R, Ayles H, Binka F, Cowan F, Kamali A, Kapiga S, Kleinschmidt I, Mayaud P, Patel V, Smith P, Weiss H.	Commentary/Conceptual analysis	United Kingdom	United Kingdom ¹
2016	Merson L, Gaye O, Guerin PJ	Commentary/Conceptual analysis	United Kingdom/Vietnam	United Kingdom
2016	Rappert B, Bezuidenhout L	Qualitative study	United Kingdom	United Kingdom/South Africa
2017	Bezuidenhout L, Kelly AH, Leonelli S, Rappert B	Qualitative study	United Kingdom/South Africa	United Kingdom
2017	Cheah PY, Day NPJ	Case study	United Kingdom/Thailand	Thailand/United Kingdom
2017	Cheah PY, Day NPJ, Parker M, Bull S	Case study	United Kingdom/Thailand	United Kingdom
2018	Akintola SO. (2018)	Commentary/Conceptual analysis	Nigeria	n/a
2018	Anane-Sarpong E, Wangmo T, Ward CL, Sankoh O, Tanner M, Elger BS.	Qualitative study	Ghana/Switzerland	Switzerland
2018	Bezuidenhout L, Chakauya E	Quantitative study	United Kingdom/South Africa	South Africa
2018	Serwadda D, Ndebele P, Grabowski MK, Bajunirwe F, Wanyenze RK	Commentary/Conceptual analysis	Uganda	Uganda

¹ These authors were affiliated with the London School of Hygiene & Tropical Medicine, and are embedded in projects in Uganda, Tanzania and India

2019	Andanda P	Commentary/Conceptual analysis	South Africa	n/a
2019	Barnes KI, Canario JA, Vernekar SS, Goudar SS, Espinal R, Merson L, Cheah PY.	Case study	South Africa	United Kingdom/Thailand
2019	Fernando B, King M, Sumathipala A	Case study	Sri Lanka	United Kingdom/Sri Lanka
2019	Humphreys GS, Tinto H, Barnes KI.	Case study	United Kingdom	South Africa
2019	Vaz M, Palmero AG, Nyangulu W, Diallo AA, Ho CWL	Commentary/Conceptual analysis	India	Singapore/Hong Kong
2019	Waithira N, Mutinda B, Cheah PY	Commentary/Conceptual analysis	Thailand/United Kingdom	Thailand/United Kingdom
2020	Anane-Sarpong E, Wangmo T, Tanner M.	Commentary/Conceptual analysis	Ghana/Switzerland	Switzerland
2020	Bull S, Bhagwandin N	Editorial	United Kingdom	South Africa
2020	Gorina Y, Redd JT, Hersey S, Jambai A, Meyer P, Kamara AS, Kamara A, Harding JD, Bangura B, Kamara MAM.	Case study	United States of America	Sierra Leone

2020 Kaewkungwal J, Adams P,
Sattabongkot J, Lie RK, Qualitative interviews Thailand United States of America
Wendler D.

Supplement 3. Summary of how barriers to equitable data-sharing and potential threats of unfair data-sharing are addressed by foundations of and best practices in equitable data-sharing

<i>Barriers to equitable data-sharing and potential threats of unfair data-sharing</i>	<i>Foundations of equitable data-sharing</i>	<i>Practices to promote equitable data-sharing</i>
<i>Structural barriers</i>	<ul style="list-style-type: none"> • Unequal research systems, resources and capacity between HICs and LMICs leads to greater capacity for HIC researchers to perform timely secondary analysis, whereas LMIC researchers have inadequate funding and expertise to curate, share and maintain data sets 	<ul style="list-style-type: none"> • Capacity building of LMIC research systems; provision of resources for sustainable and long-term data-sharing
	<ul style="list-style-type: none"> • Lack of fiscal resources limits capacity to undertake original research 	<ul style="list-style-type: none"> • Funders to provide funding for sustainable data-sharing • Funders to provide funding for sustainable data-sharing
		<ul style="list-style-type: none"> • Establish frameworks and policies that support equitable

- | | | |
|--|--|---|
| <ul style="list-style-type: none"> Limited local and national policy and regulatory frameworks to govern data-sharing | <ul style="list-style-type: none"> data-sharing, backed by enforceable governance mechanisms Accessing institutions should have their own policies supporting equitable data-sharing | <ul style="list-style-type: none"> Inclusion of LMIC stakeholders at all levels, including communities in developing data-sharing policies |
| <ul style="list-style-type: none"> Lack of enforceability of data-sharing agreements and policies | <ul style="list-style-type: none"> Establish frameworks and policies that support equitable data-sharing, backed by enforceable governance mechanisms | <ul style="list-style-type: none"> Enforceable data-sharing agreements between primary researchers and secondary researchers |
| <ul style="list-style-type: none"> Competitive scientific research environment, and rewards grounded in publication | <ul style="list-style-type: none"> Fair rewards for high-quality data collection, curation and sharing | <ul style="list-style-type: none"> Secondary user duties: offer collaboration with primary researchers where possible and appropriate; so that rewards and knowledge are shared with primary researchers Benefit-sharing and reciprocity: authorship opportunities or other forms of academic recognition on par with the burdens of data-sharing |

<p><i>Potential threats to researchers posed by unfair data-sharing</i></p>	<ul style="list-style-type: none"> • Worsening of unfair power relations with HIC researchers • Capacity building of LMIC research systems; provision of resources for sustainable and long-term data-sharing • Establish frameworks and policies that support equitable data-sharing, backed by enforceable governance mechanisms • Equitable partnerships with secondary researchers • Inclusion of LMIC stakeholders at all levels, including communities in developing data-sharing policies • Benefit-sharing and reciprocity: authorship opportunities or other forms of academic recognition on par with the burdens of data-sharing; secondary-analysis should have local translational benefits where possible; allow community input into research agendas • Burden-sharing: the cost of managing the data should be shared with secondary researchers • Secondary researcher duties: contribute to building local research capacity; offer collaboration with primary researchers
	<ul style="list-style-type: none"> • Opportunity and fiscal cost unfairly burden poorly resourced researchers and institutions; and the rewards are insufficient to offset the burdens • Capacity building of LMIC research systems; provision of resources for sustainable and long-term data-sharing • Funders to provide funding for sustainable data-sharing

- | | | |
|---|---|---|
| <ul style="list-style-type: none"> • Exclusion from participation in advancing science (LMIC researchers are stuck in a cycle of data production only) | <ul style="list-style-type: none"> • Accessing institutions should have their own policies supporting equitable data-sharing • Equitable partnerships with secondary researchers • Capacity building of LMIC research systems; provision of resources for sustainable and long-term data-sharing | <ul style="list-style-type: none"> • Funders to provide funding for sustainable data-sharing • Secondary user duties: offer collaboration with primary researchers |
| <ul style="list-style-type: none"> • Career and reputational damage leading to loss of funding opportunities | <ul style="list-style-type: none"> • Alignment of interests between primary and secondary researchers • Equitable partnerships with secondary researchers | <ul style="list-style-type: none"> • N/A |
| <ul style="list-style-type: none"> • Lack of recognition hampers career progression | <ul style="list-style-type: none"> • Equitable partnerships with secondary researchers | <ul style="list-style-type: none"> • Benefit-sharing and reciprocity: authorship opportunities or other forms of academic recognition on par with the burdens of data-sharing; secondary-analysis should have local translational benefits where possible; allow community input into research agendas |

<p><i>Potential threats to participants and communities posed by unfair data-sharing</i></p>	<ul style="list-style-type: none"> • Commercial exploitation of participants and communities 	<ul style="list-style-type: none"> • Establish frameworks and policies that support equitable data-sharing, backed by enforceable governance mechanisms 	<ul style="list-style-type: none"> • Restricted/managed access • Benefit-sharing and reciprocity secondary-analysis should have local translational benefits where possible; allow community input into research agendas
	<ul style="list-style-type: none"> • Stigmatisation 	<ul style="list-style-type: none"> • Equitable collaborations between primary and secondary researchers – reduces the risk of reporting results in a way that increases stigmatisation of participants and communities 	<ul style="list-style-type: none"> • Inclusion of LMIC stakeholders at all levels, including communities in developing data-sharing policies • Allow community input into research agendas
	<ul style="list-style-type: none"> • Benefits of research are not shared with originating communities 	<ul style="list-style-type: none"> • Establish frameworks and policies that support equitable data-sharing, backed by enforceable governance mechanisms 	<ul style="list-style-type: none"> • Secondary analysis should have local translational benefits where possible • Accountability: feedback of results of secondary analysis to community representatives

Potential threats to local health systems posed by unfair data-sharing

- Reduced opportunity to develop local research infrastructure and personnel
- Capacity building of LMIC research systems; provision of resources for sustainable and long-term data-sharing
- Secondary researcher duties: contribute to building local research capacity
- Funders to provide funding for sustainable data-sharing