

Data sharing considerations and practice among health researchers in Africa: A scoping review

DIGITAL HEALTH
Volume 10: 1–34
© The Author(s) 2024
Article reuse guidelines:
sagepub.com/journals-permissions
DOI: 10.1177/20552076241290955
journals.sagepub.com/home/dhj



Oluchukwu Loveth Obiora^{1,2} , Dorothy Agnes Shead^{3,4}
and Benita Olivier^{2,5}

Abstract

Objective: To examine the way African health researchers share data. It summarized the types of data collected, the data sharing platforms, and how the geographical distribution of the African-based health researchers influenced data sharing practices. Ethical, legal, and social aspects were considered. Institutional and government matters such as research support and funding were identified.

Methods: PubMed, Web of Science, LILAC, African Journal Archive, and Scopus databases were searched. Full-text screening was conducted, and data was extracted using the data extraction tool published in an a priori Joanna Briggs Institute-published protocol. Discrepancies were resolved by consensus. Data were illustrated using a Preferred Reporting Items for Systematic Reviews and Meta-analyses flow diagram, figures, tables, and a narrative text.

Results: Of the 3869 studies that were identified, 32 studies were included in the final study. There was a spike in the number of published studies from 2015 to 2019 ($n = 24$, 75.0%), while a decline followed in the number of publications from 2020 to April 2023 ($n = 6$, 18.8%). Ten of the studies included were from South Africa, five were from Kenya, three each were from Nigeria and Tanzania, two were from Ghana and Sierra Leone respectively, while one each was from Malawi, Ethiopia, Cameroon, Mali, Gambia, Senegal, and Burkina Faso. Negative factors impacting data sharing practices of health researchers in Africa included barriers to individual research capacity, governmental bureaucracy and corruption, legal obstacles, technological problems, prohibitive costs of publication, lack of funding, institutional delays, and ethical issues.

Conclusion: This review identified how African health researchers undertook data sharing in their countries. It pinpointed how geographical location and the resultant challenges to data distribution both individually and institutionally influenced health researchers' ability to achieve data sharing and publication of their research. It was clear that many parts of Africa are still not participating in research due to the many factors that negatively impact health data sharing in Africa.

Keywords

Data sharing, information dissemination, open science, health, Africa

Submission date: 16 May 2024; Acceptance date: 25 September 2024

¹Nelda C. Stark College of Nursing, Texas Woman's University, Houston, Texas, USA

²Wits Cricket Research Hub for Science, Medicine and Rehabilitation, Department of Physiotherapy, School of Therapeutic Sciences, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

³Department of Anatomy, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

⁴The Witwatersrand Centre for Evidence-Based Practice: A Collaborating Centre of the Joanna Briggs Institute, Johannesburg, South Africa

⁵Department of Sport, Health Sciences and Social Work, Centre for Healthy Living Research, Oxford Institute of Allied Health Research, Oxford Brookes University, Oxford, UK

Corresponding author:

Oluchukwu Loveth Obiora, Nelda C. Stark College of Nursing, Texas Woman's University, 6700 Fannin Street, Houston, TX 77030, USA.
Email: oobiora1@twu.edu



Introduction

Globally, there has been an uptick in health research data being produced commensurate with increased technological communication.^{1,2} As data sharing has become more commonplace, factors have arisen regarding how it can and should be conducted in an efficient, ethical, technologically sound, well-governed, and equitable way.^{3,4} There is advocacy for responsibility in the sharing of research data proposed by the Yale University Open Data Access (YODA) project.⁵ Four foundational principles for Findable, Accessible, Interoperable and Reusable (FAIR) data sharing, namely, “findability, accessibility, interoperability, and reusability” have been described.⁶ A requirement for effective data sharing in lower- to middle-income countries (LMICs) is the need for the establishment and building of trust and confidence in research collaboration.⁷ Good governance of data sharing practices, together with the correct leadership and oversight, is seen as essential data sharing components.⁸ A cross-sectional survey of 160 sub-Saharan researchers and scientists from 43 countries reported that most respondents were open to sharing data within institutional data storage facilities and on data sharing platforms where governance and regulation were ensured and inclined to share data when data privacy was guaranteed.⁹

The global health research data picture is skewed as it mainly includes data generated in countries having advanced technological access, reliable electricity supply, and resource-rich settings.^{10,11} The last 10 years have seen increased scientific research being conducted in LMICs with an incremental increase in data generation.¹⁰ However, there has been a decrease in the sharing of such research globally.¹¹ In Africa, there are mainly LMICs with limited resources, poor technological infrastructure, and power outages that present a barrier to the management and sharing of data.¹⁰

On the African continent, out of a total of 54 independent countries, 33 fall into the low-income category, 12 into the lower middle-income category, and four into the upper middle-income category.^{12,13} The fact that 83% of African countries are in lower-income brackets dictates the amount of funding support that can be given locally to health researchers. In 2006 African Union member countries pledged to spend one percent of their Gross Domestic Product (GDP) on research and development but in 2019 the development funding only stood at 0.42% of GDP.¹⁴ In the last 10 years, only Kenya (0.98%), South Africa (0.82%), and Egypt (0.80%) are close to the 1% target.^{14,15} This contrasts poorly with the average global contribution of 1.7% of GDP.¹⁴ It follows that African researchers are dependent on outside financial assistance to get their research projects funded and published. Many research funders advocate that to receive such financial support researchers are obliged to share their data sets to make maximum use and value of such data.^{3,16,17}

The increasing need to share data from African LMICs in an affordable way has raised the issue of involving researchers from these countries in Open Data Access.¹⁸ In the global scenario, this links to the Open Science concept and the emergence of publicly funded research arising from government and research funder motivation.^{18,19} African researchers are increasingly being guided by the principles of Open Science.^{18,19} Within this context, Hulsen (2022)²⁰ reported on how academia is reluctant to share data before publication and suggests that using federated data might be a solution. In 2017, the Africa Open Science Platform (AOSP) was established as a federated system where individual or institutional researchers could share and reuse research data.²¹ Collaborative networking is essential in the data sharing process, and researchers who endeavor to pursue excellent research are encouraged to share their work on this platform.²¹

The establishment of any formative data sharing process requires ethical controls and careful governance.^{9,22} Globally, laws have been promulgated and instituted to offer protection to health research participants.^{23–27} Fernando, King, and Sumathipala (2019)⁷ recognized that “parachute research” where high-income countries (HICs) researchers’ utilized African data, without acknowledging the contributions of African participants or researchers, is now largely condemned. They also stated that for true partnerships between international colleagues to be established, recognition of contributions by African colleagues and participants had to be given credence, and the concept of beneficence no longer met this requirement. It is also hoped that going forward African researchers will be in research leadership roles and be involved in decision-making.²⁸

As part of the governance over how data is shared, institutional Research Ethics Councils (RECs) play an important part in ensuring that intellectual property rights and ethical sharing of research data are observed. The incorporation of Community Leaders into such RECs to give the local, rural, or cultural perspective on data collection has been suggested.²⁹ Individual African researchers are often worried that if they share raw data before using it for a degree or for publication, they might lose recognition for their work.^{29–31} Unfortunately, in some African countries, long delays in ethical approval for research projects can lead to research being delayed or even canceled, resulting in the loss of valuable data.²⁹

A scoping review approach was selected for this review as it was best suited to map the wide-ranging topic of how data sharing was considered and practiced in African countries by health researchers based in those countries. This scoping review aimed to examine how African health researchers share their data. It was informed by a Joanna Briggs Institute (JBI) a priori protocol,³² in which a search of JBI Evidence Synthesis, the Cochrane Database of Systematic Reviews, MEDLINE, Figshare, Open Science Framework, and PROSPERO showed that there was no relevant complete or in-progress reviews on this topic. The scoping review considered types of data generated/collected, and the data sharing

platforms used to disperse the data. The geographical location of the health researchers impacted the considerations that had to be factored in to enable data sharing and whether the researchers could practice data sharing individually or within institutional boundaries. Geographical context pinpointed the African health researcher's point of view. Ethical, legal, and social aspects were considered. Institutional and government matters such as research support and funding, and the publishing of data were reported. Credence was given to the perspective of the individual researcher. The review was instrumental in highlighting the diversity of data sharing practices in Africa, the different data sharing platforms available, and the lack of consensus between them. It identified gaps in the literature related to improving and enhancing health research data sharing on the continent.

Review questions

- What are the data sharing practices (including the relevance of geographical location, types of data, data management plans, and data sharing platforms used) of health researchers residing in Africa?
- What are the barriers or facilitators to data sharing for these health researchers?
- What ethical, legal, institutional, and funding aspects are being considered by the African-based health researchers who are sharing research data?
- What author-reported recommendations regarding ways to improve data sharing in Africa are given in the included literature?

Inclusion criteria

Participants

This review included studies that reported on how research data were shared by health researchers in Africa. The researcher's years of experience or clinical specialty was not a consideration for eligibility. Individual researchers and groups conducting health research through various research entities or at academic or tertiary institutions were included. Data generated by physiotherapists, doctors, nurses, and various other researchers working in the field of African health were included if the data were pertinent to the answering of the review research questions. Both primary and secondary research were included. Lab-based work and opinion pieces were considered for inclusion. Gender, age, and other sociodemographic factors were not used as exclusion factors.

Concept

The concept of interest was the health data sharing considerations and practices of researchers living on the African

continent. Data sharing enables researchers to avoid repeating existing research, distribute their research, and build upon research generated by others. It encourages better transparency, reproducibility, and ethical management in research (US Geological Survey, 2023 [Why share your data? usgs.gov]). Health research data sharing is important in developing governmental, institutional, or professional policies that dictate how people will be beneficially treated for health problems. In this scoping review, all the health data generated for research by individual health researchers, organizations, institutes, and health and medical facilities was termed health research data. Raw data and the cleaned data set, metadata, and summary-level data such as summary-level results posted on registries, lay summaries, and publications were able to be shared.³³

Research question two raised the concept of the barriers or facilitators to data sharing on the continent. Barriers are those factors that inhibit or prohibit data sharing and facilitators are those factors that assist, promote, and make easier the data sharing process. This opened the door to consideration of factors unique to the African diaspora that might impact the health data sharing process.

Therefore, the scoping review included studies on health data management plans, how health research data were shared including the geographical locations, data sharing platforms used, types of data collected, and legal, institutional, ethical and funding-related factors that influenced the sharing of that data. Cognizance was given to the fact that ethical considerations are based on human rights and morally accepted practices whereas legal considerations are based on written and passed laws. Included studies also described individual-level data sharing, clinical or patient-oriented data sharing (Table 1), and the challenges to data sharing among health researchers in Africa.

Context

The focus of this scoping review was on the sharing of research data generated by health researchers residing in this geographical region, whose interest was in the collection of data from local study participants. This is because in Africa, some context-related factors which are sometimes systemic, technological, and cultural, defined how researchers approach the challenges, barriers and unique circumstances that surround the collection and sharing of health data in both urban and rural communities.^{27,34} No language restrictions were applied for the articles to be included in the study. Furthermore, as stated in the peer-reviewed and published protocol for this study,³² data sharing is partly technology-driven^{34,35} and there has been a proliferation of research studies in this area in the last 10 years. This scoping review initially was limited to articles published from 1 January 2011 to 30 June 2021 but as due time was taken for the preparation of the manuscript a further search was undertaken up to 3 February 2024. This

Table 1. Relationship between study aim, geographical location of data collection, types of data collected and data sharing platform used.

S/ N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
1	Aidam and Sombié (2016)	The West African Health Organization's experience in improving the health research environment in the ECOWAS region	ECOWAS West African countries (Benin, Burkina Faso, Cape Verde, Cote d'Ivoire, The Gambia, Ghana, Guinea, Guinea Bissau, Liberia, Mali, Niger, Nigeria, Sierra Leone, Senegal, and Togo)	To document how the WAHO implemented a research development program in West Africa during 2009–2013 using the Knowledge for Better Health Research Capacity Development Framework, developed by Pang et al. (<i>Bull World Health Organ</i> 81(11):815–820, 2003)	Technical reports and financial reports of project activities, documented experiences of the author, staff of the various departments in WAHO, and other stakeholders in member countries	Epidemic Intelligence from Open Sources
2	Anane-Sarpong et al. (2018a)	You cannot collect data using your own resources and put it on OA: perspectives from Africa about public health data sharing	Ghana and Tanzania	To explore relevant experiences, contextual, and subjective meanings, and values that public health stakeholders in Africa attach to the scientific, socioprofessional, and ethical dynamics of data sharing	Population-level data to help inform public health activities	Databanks and data repositories such as the INDEPTH plus the HDSS
3	Anane-Sarpong et al. (2018b)	Application of ethical principles to research using public health data in the global south: perspectives from Africa	Ghana and Ethiopia	To explore the perspectives of professionals with a working knowledge of the HDSS to determine practical ways of appropriating the foundational principles of health research to advance the ever-growing opportunities in Research using Public Health Data	Real-time data collected from humans	HDSS
4	Anane-Sarpong et al. (2020)	Ethical principles for promoting health research data sharing with sub-Saharan Africa	Sub-Saharan Africa	This article proposes six principles for mandatory data sharing in relation to peculiar challenges faced by sub-Saharan African scientists to suggest ethical principles for rethinking and reframing solutions	Health research data sets which can be combined data sets and/or primary data sets from individual research	Databanks. Data repositories HDSS and INDEPTH. Data from journal publications

(continued)

Table 1. Continued.

S/ N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
5	Cole et al. (2016)	Implementing a national health research for development platform in a low-income country—a review of Malawi's Health Research Capacity Strengthening Initiative	Malawi	A case study describing one science granting council which operated as a platform for prioritizing, financing, and managing health research and development—the Malawi Health Research Capacity Strengthening Initiative	Health data from patients' health records	Biobanks, data silos, cyberbanks, and virtual research repositories
6	Conton (2017)	Build the Ebola database in Africa	Sierra Leone	To put forward the argument that an Ebola database should be established in West Africa and not in Oxford in England	N/A	National health research for development (R4D) platform
7	Denny et al. (2015)	Developing ethical practices for public health research data sharing in South Africa: the views and experiences from a diverse sample of research stakeholders	South Africa	To provide an analysis of stakeholders' views about data sharing in South Africa	Public health data, patient's health data	Infectious Diseases Data Registry
8	Dhai et al. (2015)	Biobanks and human health research: balancing progress and protections	South Africa	To document the reasons for encouraging the use of biobank assets while acknowledging the ethical, legal, and social issues associated with such usage	Human biological materials (specimens) and their associated data	Biobanks
9	Gorina et al. (2020)	Ensuring ethical data access: the SLED model	Sierra Leone		SLED - including distinct data systems for Ebola hot line calls, patient information collected by field surveillance officers, laboratory testing results, clinical information from Ebola treatment and isolation facilities, and burial team records	Journals, databases

(continued)

Table 1. Continued.

S/ N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
10	Impouma et al. (2021)	Information management practices in the WHO African region to support response to the COVID-19 pandemic	Congo	To discuss difficulty of data collection, processing, analysis, and dissemination not only limited to the wide scope and large volumes of data required, but also the different structures and methods of data collection and dissemination adopted by each member state	The main types of data collected by AFRO were case-based or aggregate data on cases and deaths, data on laboratory tests performed, data on contact tracing, as well as signals for monitoring and verifying a unique evolution of the pandemic. Routine health system data, accessible through health management information system portals of the respective Ministries of Health, as well as open-source data on the public health restriction measures available from reputable institutions, were also collected to complement the epidemiological data on COVID-19 reported through official channels. Official COVID-19 data were sourced daily by a dedicated team of epidemiologists and data scientists from the emails sent to AFRO as well as the official websites or social media pages of the respective member states	Official websites or social media pages of the respective member states
11	Jao et al. (2015)	Research stakeholders' views on benefits and challenges for public health research data sharing in Kenya: the importance of trust and social relations	Kenya	To explore views on benefits and challenges in research data sharing	Health and demographic surveillance data from patients and communities	Clinical surveillance databases in hospitals

(continued)

Table 1. Continued.

S/ N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
12	Lötter and van Zyl (2015)	A reflection on a data curation journey	South Africa	To comment on a reflection on the experiences of data preservation and sharing practices in a South African research organization	Nationally representative, cross-sectional survey data dealing with attitudinal, behavioral, and health issues	District Health Information Software 2
13	Lwoga and Questier (2015)	Open access behaviors and perceptions of health sciences faculty and roles of information professionals	Tanzania	This study sought to investigate the faculty's awareness, attitudes and use of OA, and the role of information professionals in supporting OA scholarly communication in Tanzanian health sciences universities	Not specified	Journal websites for published research output and institutional research repository
14	Mbuthia et al. (2019)	Kenyan health stakeholder views on individual consent, general notification, and governance processes for the reuse of hospital inpatient data to support learning on health care systems	Kenya		Routine monthly audits for example, monthly aggregated data on patient diagnoses are routinely reported to the national government on a statutory basis across Kenyan public hospitals); an evaluation of a treatment guideline; and two types of CER, including a retrospective comparison of two antibiotics in routine use and a prospective comparison including randomization of patients to one of the two antibiotics under study	Not specified
15	Memiah et al. (2018)	Bridging the gap in implementation science: evaluating a capacity-building program in data management, analysis, utilization, and dissemination in low- and middle-income countries	Kenya and Tanzania	To evaluate health care providers' knowledge, competency, and skills related to data management, analysis, utilization, and dissemination before and after the training program	Demographic data Health survey data	Journal publication, conferences

(continued)

Table 1. Continued.

S/ N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
16	Moodley et al. (2022)	Ethics and governance challenges related to genomic data sharing in southern Africa: the case of SARS-CoV-2	South Africa	To report on the challenges facing African researchers with regard to the sharing of research data in the era of the Covid-19 pandemic. To show the implications of disclosing new variants in Africa	Data biospecimens, and code from human participants	Global Initiative on Sharing Asian Influenza Data
17	National Academies of Sciences, Engineering, and Medicine (2015)	Sharing research data to improve public health in Africa: a workshop summary	N/A (Report of international workshop in South Africa)	The workshop was to raise the profile of issues around the sharing of public health data in Africa, enable the Wellcome Trust and its international partners to highlight findings of previous sponsored research on data sharing in Africa, identify issues that mitigate against public health data sharing and pathways through research and policy venues to foster increased sharing, and, in general, serve as a way to bring more African voices and perspectives into the dialogue	Journal publications of raw health data used for research	Biobanks
18	Nature Medicine Editorial (2015)	Sharing data to save lives	Switzerland (data from Ebola epidemic in West Africa)	To address another concern, which a consensus felt was a formidable barrier to early data release: public release of data will have a negative effect on later scientific publication of the same data	Patient samples and scientific publications	Health Sciences Research Council (HRSC)

(continued)

Table 1. Continued.

S/N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
19	Nembaware et al. (2019)	A framework for tiered informed consent for health genomic research in Africa	South Africa and Cameroon	To propose a framework and template for the collection of informed ethical consent from study participants and to link those tools to relevant data-use domains	Health genomic research data	ICT (Platform was not specified)
20	Nnamuchi (2018)	H3Africa: An Africa exemplar? Exploring its framework on protecting human research participants	Nigeria (on whole of Africa)	The ultimate goal of this paper is to determine whether H3Africa's ethicolegal regime is consistent or at odds with international guidelines and best practices regarding protection given to research participants safe-guarding privacy, maintaining confidentiality of health information, and sharing of data/biospecimens	Data relating to the reports of the H3Africa consortium findings	Electronic health records
21	Nordling (2019a)	Africa's science academy leads push for ethical data use	Nairobi, Kenya	To showcase the African Academy of Sciences (AAS) goal to create the continent's first cross-disciplinary guidelines for sharing data	Participants' health information and biological specimens	Not specified
22	Nordling (2019b)	South African law may impede human health research	South Africa	To report on the new South African POPIA and how it may impede human health research	Not specified	Learning health care system
23	Obiora et al. (2022)	Perceptions of human movement researchers and clinicians on the barriers and facilitators to health research data sharing in Africa	All countries in Africa	To describe the perceptions of human movement researchers and clinicians on the barriers and facilitators to research data sharing in Africa	Primary/individual health data from research participants and patients, journal publication	Journal publications
24	Ruggeri (2014)	Disseminating health research in sub-Saharan Africa through journal partnerships	sub-Saharan Africa	To determine how junior sub-Saharan African researchers may best be helped to disseminate their research through journal partnerships	Not specified	Not specified

(continued)

Table 1. Continued.

S/ N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
25	Shabi et al. (2011)	Physicians' utilization of internet medical databases at the tertiary health institutions in Osun State, Southwest Nigeria	Nigeria	This study set out to answer the following questions: 1. What is the extent of utilization of the Internet-based medical databases by physicians in the two teaching hospitals in Osun State, Southwest Nigeria? 2. What are the purposes for seeking medical information on the Internet? 3. What factors determine the use of Internet as an information resource? 4. How skillful are the respondents in using the Internet? 5. What is the impact of the Internet-based resources on both the clinical and research output of the physicians?	Survey demographical data on internet use, access, and preferred information source	Scientific journal publications
26	Shaffer et al. (2018)	Development of a data collection and management system in West Africa: challenges and sustainability	Mali, Gambia, and Senegal	To describe the development of a data collection and monitoring system by the West African International Centers of Excellence for Malaria Research, and to encourage the linkage of DCMSs to international research and Ministry of Health data systems and repositories	Clinical laboratory results, epidemiological results/reports	H3Africa
27	Staunton and De Stadler (2019)	Ethical and practical issues to consider in the governance of genomic and human research data and data sharing in South Africa: a meeting report	South Africa	To identify and discuss challenges and opportunities in the governance of data sharing for genomic and health research data in South Africa	Genomic and human research data	H3Africa platform

(continued)

Table 1. Continued.

S/ N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
28	Staunton et al. (2019)	Protection of Personal Information Act No. 4 of 2013: Implications for biobanks	South Africa	To discuss the implications that the POPIA might have for biobanks	Personal information including individuals' health information and samples	Human Heredity and Health in Africa (H3Africa) and Bridging Biobanking and Biomedical Research across Europe and Africa (B3Africa)
29	Townsend et al. (2022)	The lawful sharing of health research data in South Africa and beyond	South Africa	The study time aimed to answer three broad questions: (a) What are the existing legal modalities governing and protecting data use and sharing in South Africa, and, more generally, in Africa? How can data be transferred into and out of South Africa? (b) What can be learned from recent international developments in data transfer? And, lastly, (c) where plausible how might data flows throughout Africa be facilitated in the interests of public health during times of pandemic?	Data acquired from digital health and medical technologies, such as wearable devices, digital health (or eHealth) applications, and medical devices and sensors; electronic health records and administrative hospital data; aggregated clinical trials; genetic and microbiomic sequencing of human biological materials; data analyses; health claims registries; and from social media, smartphones, and other electronic devices	Not clearly reported
30	Turcotte-Tremblay and Mc Sween-Cadieux (2018)	A reflection on the challenge of protecting confidentiality of participants while disseminating research results locally	West Africa-Burkina Faso	The objectives of this reflection are to (a) show how researchers conducting health systems research in LICs can experience, difficulty in protecting the confidentiality of participants locally and (b) present some potential strategies to minimize this difficulty	Result of health system research	Not specified

(continued)

Table 1. Continued.

S/N	Authors (year of publication)	Title	Study location/country	Aim of the study	Types of data	Data sharing platforms
31	Uzochukwu et al. (2016)	Health policy and systems research and analysis in Nigeria: examining health policymakers' and researchers' capacity assets, needs, and perspectives in southeast Nigeria	Nigeria	To report on the HPSR+A system in Nigeria	Health system and policy research data	Journal publications
32	van Hoving and Bysiewicz (2017)	African emergency care providers' attitudes and practices toward research.	In countries in Africa served by the African Federation of Emergency Medicine. Base is in Cape Town, SA	The study aimed to assess the perceptions and practices toward research among current emergency care providers in Africa	Demographic data of respondents	Databases

AFRO: WHO Regional Office for Africa; CER: comparative effectiveness research; COVID-19: Coronavirus disease 2019; DCMS: Data Collection and Management System; ECOWAS: Economic Community of West African States; HDSS: Health and demographic surveillance system; HPSR+A: Health Policy and Systems Research and Analysis; ICT: Information and Communication Technologies; INDEPTH: International Network for the Demographic Evaluation of Populations and their Health; LIC: lower-income country; N/A: not available; OA: open access; POPIA: Protection of Personal Information Act; SARS-CoV-2: severe acute respiratory syndrome coronavirus-2; SLED: Sierra Leone Ebola Database; WAHO: West African Health organization; WHO: World Health Organization.

timeframe reflects current developments in the data management and sharing space.

Types of sources

We considered all types of quantitative, qualitative, and mixed-methods studies that reported on the sharing of research data among health researchers in Africa. Unpublished articles, opinion pieces, policy documents, and dissertations that include information on health research data sharing among health researchers in Africa were also evaluated for admittance into this review. All primary, review, and grey literature that reported on the data sharing practices among health researchers in Africa was included in this review.

Methods

The proposed scoping review was done in conformity with the JBI scoping review methodological framework for systematic reviews.³⁶

Search strategy

A three-step search strategy was used. A preliminary restricted search of databases PubMed, Scopus, and LILAC was conducted. Then common wording found in both the title and abstract of extracted papers and of the index terms was collected to enable a description of the articles to be formulated. Key terms were decided through deliberations or exchange of views among the three authors (OLO, BO, and DAS).³⁷

A second search utilizing the keywords and index terms identified from the preliminary search was conducted in the PubMed, Web of Science, LILAC, African Journal Archive, and Scopus databases. Gray literature was sourced using Google scholar and consideration was given to relevant inclusions in the first 20 pages sourced. The search was repetitive as reviewers got accustomed to the supplementary keywords, sources, and search terms. The extra or supplementary keywords, sources, and search terms discovered to be important were included into the search strategy. The search involved the use of keywords or text words such as: data sharing, data access, open science, data management, open access, data ethics, data sets, data management activities, reposito*, health, medical, African countries, Africa, human. Also index terms like information dissemination and data systems were used for the search. The search strategies were designed categorically for each database using the relevant index and free text terms. A complete search strategy for each database is presented in Appendix 4 of the supplementary files.

In the final step, the reference lists of identified relevant articles were searched for possible additional references. The titles and abstracts of all articles found in the search were sorted, and the full-text versions of eligible articles

were obtained. The search for articles was carried out by OLO and DAS.

Source selection

No relevant conference proceedings were found in the above-mentioned databases. In this regard, Web of Science and Scopus databases publish reputable conference proceedings, and they are among the best databases to be searched. A deviation from the published study protocol was that no relevant articles were found in LILAC, so African Journal Archive was searched to ensure that a database that focuses on articles published in Africa was considered and searched. Any discrepancies related to source selection were resolved by consensus.

Study selection

Following the search, all results from the databases were exported into the (Covidence software) for systematic reviews.³⁸ Duplicates were removed before the articles were screened for eligibility. Study selection was done in two stages. First, titles and abstracts were screened against the inclusion criteria by two independent reviewers (OLO and DAS). Then, all potentially relevant full-text articles were retrieved and screened for inclusion in the final review. Hand searches of the reference lists of eligible studies was carried out to ensure that no relevant studies were missed. The basis for exclusion of full-text studies that did not qualify for the inclusion criteria were reported in the review, and discrepancies at this stage were resolved by the third reviewer (BO).

Data extraction

Two independent reviewers (OLO and DS) extracted data using the data extraction form created for this review.³² The third author (BO) then checked the relevance of the extracted data. Inconsistencies were sorted out through discussion until agreement was reached. The data extracted for this review encompassed details such as year of publication, country of origin, aim/purpose of the study, types of data, study population and sample size, perceptions, barriers and facilitators, ethical considerations, legal, institutional and funding-related aspects as well as the author-reported recommendations for mitigation of the challenges of data sharing in Africa. As endorsed in the JBI Manual for Evidence Synthesis,³⁶ the data extraction form was pilot tested on a small sample of studies to ensure that all relevant data were extracted. As this review was an iterative procedure, it was cumulatively filtered and regularly updated.

Data presentation

Following Preferred Reporting Items for Systematic reviews and Meta-analyses (PRISMA) guidance, the

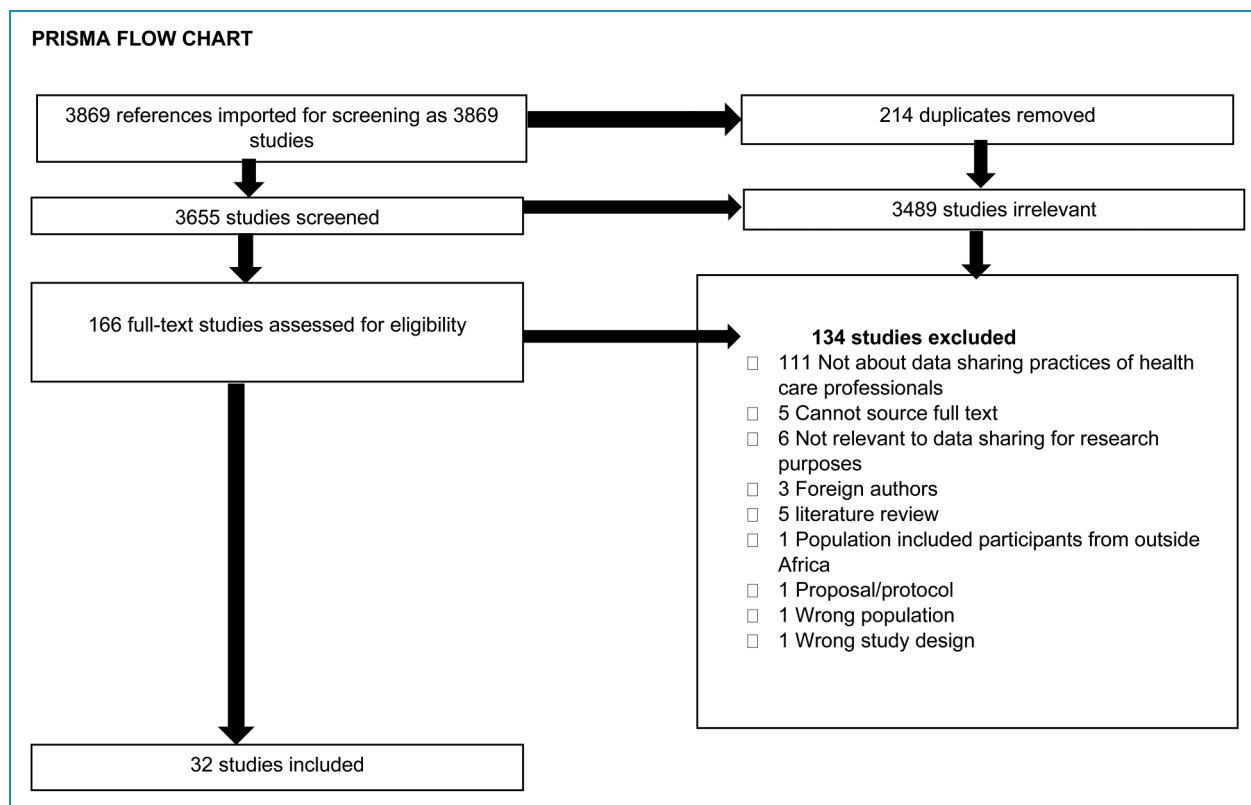


Figure 1. A Preferred Reporting Items for Systematic reviews and Meta-analyses (PRISMA) flow chart showing the search results and source selection and inclusion process.

results of the search were reported in full and presented in a PRISMA flow diagram (Figure 1 and Appendix 1 of the supplementary files)³⁹ The results of the scoping review are presented in tables and synthesized into relevant charts, to facilitate the readers' ability to understand and utilize the findings. The data sharing practices, country of origin, types of data aim/purpose of the study, study population, perceptions, barriers and facilitators, ethical considerations, legal, institutional, and funding-related aspects as well as the author-reported recommendations for mitigation of the challenges of data sharing in the articles were reported. The tabular summary (see Table 1) is detailed in a narrative summary, conveying the objectives of the scoping review.

Results

Source inclusion

Through database searches conducted in April 2023, 3971 studies were identified, and 214 duplicates were removed. An additional two studies were identified by hand search. A total of 3759 studies were screened by title and abstract for inclusion. Of those, 3592 were excluded because they were irrelevant to the study aims. The remaining 167 studies were assessed for inclusion based on full-text

review, and 135 full-text studies were excluded (see Figure 1 for reasons for exclusion). The final data set consisted of 32 studies for data extraction. The search results are summarized in a PRISMA flow diagram (Figure 1).³⁹

Characteristics of included studies. Review findings

Review question #1: What are the data sharing practices (including geographical locations, types of data, data sharing platforms used) of health researchers in Africa?

Year of publication and geographical locations. There was a spike in the number of published studies from 2015 to 2019 ($n=24$, 75.0%), while a decline followed in the number of publications from 2020 to April 2023 ($n=6$, 18.8%). One study was published in each of 2011 and 2014 ($n=2$, 6.2%). Only one relevant study was included between the beginning of 2022 and 20 April 2023.

Not all studies included in this scoping review focused on data management and sharing in a specific country. The more generalized study regions were named as: Africa (12.9%, $n=4$), sub-Saharan Africa (6%, $n=2$), and West Africa (6%, $n=2$). The remainder of the studies referred to specific countries in Africa (see Appendix 2 of the supplementary files). These countries are indicated on the map in Figure 2. Of these countries 10 were done in

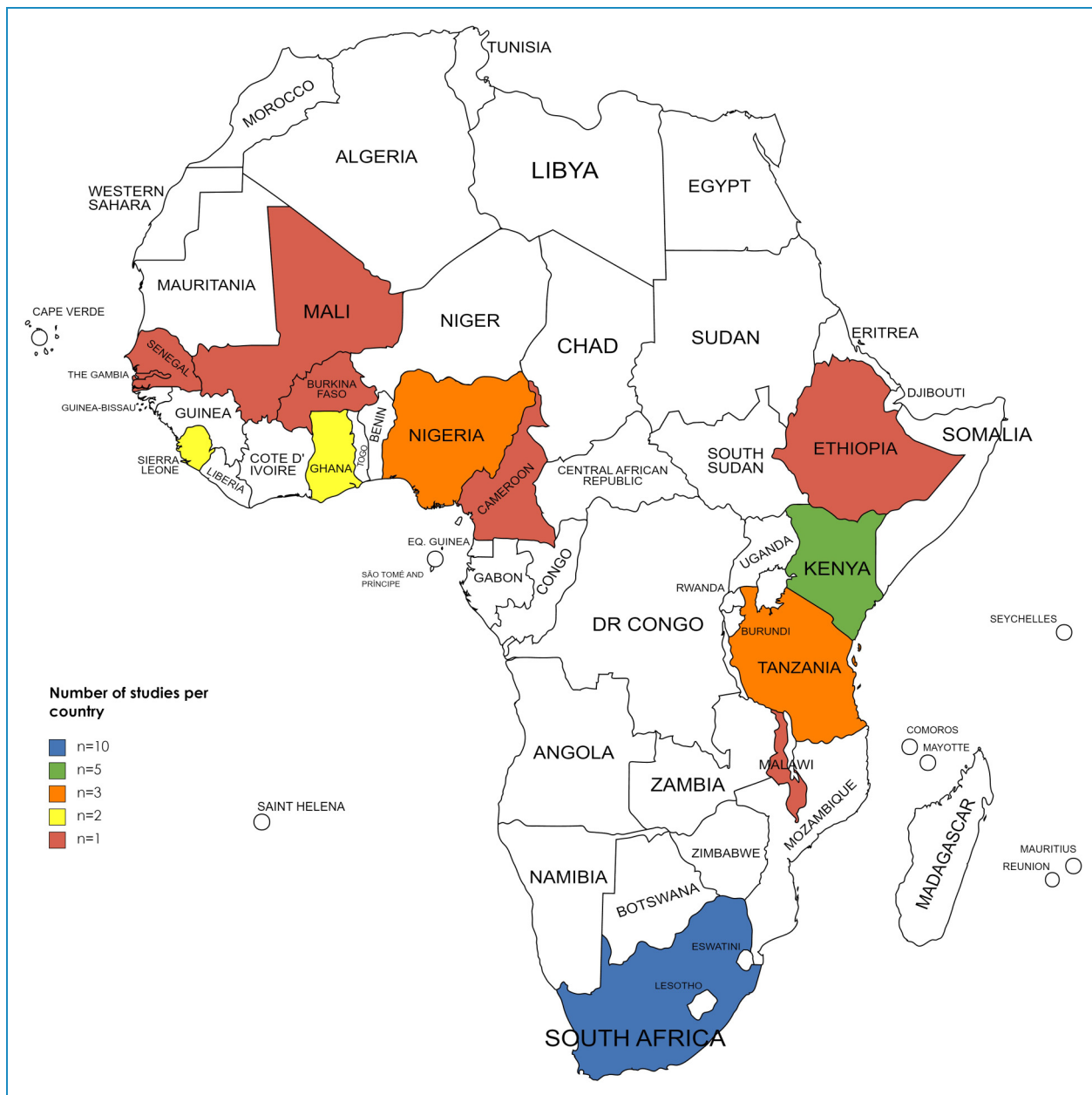


Figure 2. A mind map of factors impacting data sharing between health researchers in Africa.

South Africa, five were from Kenya, three each were from Nigeria and Tanzania, two were from Ghana and Sierra Leone, respectively, while one each were from Malawi, Ethiopia, Cameroon, Mali, Gambia, Senegal, and Burkina Faso. A number of these studies involve more than one country per study.

Types of data. The types of data that were reported to be shared by health care researchers in Africa were real time health and demographic data collected from human patients.^{30,40–53} Population-level health data,^{30,34,54} health data from patient’s health records,^{40,41,43,47,51} health

system and policy research data,^{43,55–58} epidemiological results and reports,^{43,44,59} secondary health data from primary research articles,^{29,34} monthly aggregated data on patients’ diagnosis,⁴⁶ human health data from databases and repositories,^{30,34,40,41,44,45,47,52–54,59,60} human body specimens and their associated data,^{40,42,50–53,59} data acquired from digital health and medical technologies, such as wearable devices, digital health (or eHealth) applications, and medical devices and sensors, routine health system data, data on contact tracing,^{44,61} electronic health records (EHR) and administrative hospital data,^{30,34,40,41,45,54,57,58,62} data from laboratory tests

performed,⁴³ aggregated clinical trials⁴⁶; genetic and microbiome sequencing of human biological materials,^{52,53,61,63} health and demographic surveillance data from patients and communities,^{30,34,45,54,63} national representative survey data⁵⁷ dealing with attitudinal, behavioral, and health issues,^{30,63} data analyses,⁴⁷ health claims registries, and from social media,⁴⁴ technical reports of health project interventions, smartphones, and other electronic devices,⁴⁴ ethical and confidentiality and security matters related to data sharing,^{29,30,34,42,49,53,61–65} and journal matters^{43,47,48,66–68}

Cognizance was taken regarding the possibility of some overlap of similar types of extracted data and this was addressed by sorting and categorizing the data by means of tabulation using color coding under specific headings.

Data sharing platforms. For the purposes of this review the term databank was used as a generic term referring to any form of data collection. The technical term database was used when referring to a collection of data managed by a Database Management System (DBMS). A data repository described a central location where data was stored and managed. In order to clarify these types of data sources a priori plan was in place to sort and group such data.

The platforms that were used to share these data were relevant databases, databanks and data repositories like the Human Heredity and Health in Africa (H3Africa),^{52,53,59} and Epidemic Intelligence from Open Sources (EIOS),⁵⁶ and International Network for the Demographic Evaluation of Populations and their Health (INDEPTH)^{30,34}; institutional repositories,^{30,34} journal websites,^{43,47,58,66,67} biobanks^{40,42,48} like Bridging Biobanking and Biomedical research across Europe and Africa (B3Africa),⁵³ smartphone software applications,⁶² EHR,^{30,34,40,41,45,54,57,58,62} clinical surveillance databases in hospitals,^{30,34,45,54,63} district health information software,^{30,34,54} official websites or social media pages of respected member states,⁵⁹ infectious disease data registry,^{41,52,53,56,59} national health research for development (R4D) platform (IDDO),⁴¹ data silos,⁴⁰ cyberbanks,⁴⁰ and virtual research repositories.⁴⁰ A summary of the characteristics of the included studies is presented in Table 1.

Review question #2: What are the perceptions, barriers, facilitators, and author-reported recommendations of African health researchers on data sharing in Africa?

The perceptions, barriers, and facilitators to health data sharing among health researchers in Africa were summarized into individual research capacity concerns, issues related to journal publications, governmental issues, social issues, and technology issues.

Individual research capacity concerns. Several studies^{30,41,45,46,48} reported that even though there is willingness to share data at management levels, individual

scientist/researchers are reluctant, unconvinced, and hesitant to share data. The reason for this low motivation among individual health researchers was due to distributive justice concerns (fairness), trust, sensitivity of the data, reciprocity, and inclusiveness.^{30,45,46,48,53} According to Anane-Sarpong et al.,^{30(p399)}

“Data is used out of the country without the original collectors only to later hear of a new publication. It’s not fair”; and “it’s all been taken for granted... If somebody at the country level does not raise eyebrows, [data] just goes.”

Other causes of reluctance among individual researchers were due to inequities in the environment, research skills, privileges, burdens, incentives, opportunities, reluctance to change established work habits, concerns about the additional burden of detailed recording of metadata, insufficient funding, and rewards.^{29,34,41,45,48,55–58,68} Some authors^{29,34,45,57,68} reported that although individual scientists are aware of the benefits of data sharing in health research, the inability of the researchers to fully analyze and utilize their data (due to poor research statistical skills, lack of access to software for research analysis, lack of infrastructure and technology to enable adequate analysis of data) before publishing it makes them reluctant to share data. A participant in the study by Anane-Sarpong et al.^{34(p90)} stated that:

data which is under-utilized at first publication due to manual processes or use of sub-standard analytical tools holds the greatest potential for further analysis and new manuscript preparation.

Under-resourced African scientists are most likely to collect such data. They fear that data sharing will expose them to stiff competition for the use of their data with better resourced competitors as they as data producers do not have many rights and privileges.^{34,48} This fear of ideas and data being overtaken by more resourced researchers in the global professional race for impactful research publications leads to African researchers, holding back their data.⁴⁵ Also, even among researchers in Africa, there were reports of researchers stealing data from other researchers for their personal research studies.⁴⁵

Human Heredity and Health in Africa (H3Africa) reported that data at H3Africa was held by them for a 23 month period of time to allow for the data contributors to have the right of exclusive use of that data.⁶² This longer exclusivity period was designed to assist African researchers from the resource challenged continent to exploit, analyze, and publish their data, thereby preventing other researchers from publishing such data first.⁴⁸

Other reasons that make the African researchers not to be able to fully utilize their data within a reasonable time period include: time constraints due to the researchers

working multiple roles and working under extreme pressures due to severe shortage of staff and technology,^{29,34,48,55} lack of competence to preserve data in ways that make for effective aggregation and third party analysis,^{29,34,57,68} lack of data sharing precedence,⁴⁸ lack of specialized infrastructures such as data management and curation,⁴⁸ fear that data would lose their value once placed in the public domain,⁴¹ the potential for shared data to be misused and misunderstood so that it produces false conclusions,^{41,46,48} the problem of data collectors going unrecognized while someone else will analyze the data and get all the recognition and publications,^{41,48} worry about losing control on the future use of shared data,⁴⁰ low compliance with informing the data-curation centers of all outputs derived from shared data; this is because despite that the access to online data collections is usually governed by a time-dependent expiration of the end-users' subscription, the downloaded content may remain in the end users' possession indefinitely.⁴¹

Issues related to journals. African researchers reported that although several journals require data sharing, the guidelines around it are “very loose.”³⁰ Data sharing practices relating to journal publication that were reported included: holding on to data with plans of fuller use in manuscript writing, delaying the submission of ready manuscripts until the highest number of manuscripts is ready for concurrent submission, in order to retain control of the data in those manuscripts for as long as one is engaged in other roles, and avoiding the publication of novel complex new ideas requiring the release of copious data and perhaps metadata.³⁴ A noted facilitator to the dissemination of research was that some journal publishers and the World Health Organization (WHO) were trying to increase the accessibility and affordability of scientific publications to researchers in Africa.³⁴

Identified incentives that were shown to propel African health researchers toward sharing data were the traditional outputs and metrics of research around the requirements for the career progression of academic faculties.⁴⁸ Many research stakeholders expressed strong support for journal policies on sharing the data sets that underpin the data analyses, while ensuring that the primary communities involved in the data collection should be recognized in all future publications that would emanate from the shared data.⁴⁵

A study by Lwoga et al.⁶⁷ on the open access behavior and perception of health science faculties revealed that only one institution in Tanzania hosted a local journal and that faculty had published not more than 38% of their articles, and that they have self-archived not more than 26.8%. According to this same study,^{67(p44)}

Most researchers are not aware that a growing number of publishers allow archiving of pre- or post-print articles into repositories prior to their publication. Authors are not familiar

with the Sherpa/RoMEO service that provides researchers with information regarding publishers' self-archiving policies and the permissions they grant to authors to disseminate different versions of a published article.

Aidam and Sombié⁵⁶ reported that the funding and efforts provided the West African Health Organization (WAHO) resulted in the publication of four research articles in international peer reviewed journals, as well as the organization of two regional scientific congresses. All these were aimed at promoting research and data sharing among the health researchers in West Africa.

A study on disseminating health research in sub-Saharan Africa through journal partnerships⁶⁷ reported that because large studies submitted to top journals for possible publication can take a long time to have peer review feedback, resubmission, or rejection simply because of numbers of competing submissions. This long period of waiting is problematic due to the limited available funds to support write-up after project completion.

Government related aspects. One of the government-related barriers to the sharing of health data in Africa is that African countries do not have the capacity to monitor nor enforce the compliance of local and foreign researchers to the regulations surrounding the data transfer agreements.³⁰ An article⁴⁴ which discussed the information management practices in the WHO African Region to support response to the Coronavirus disease-19 (COVID-19) pandemic reported that some political barriers affected the efforts of the WHO African region to monitor and disseminate information on the transmission dynamics of the COVID-19 in 2020 in the member states. An example was that the United Republic of Tanzania did not share the data collected on the COVID-19 situation in their country, while the other member States who established political structures that coordinated COVID-19 data flows, did not adhere to the International Health Regulation (IHR)-2015 on data sharing requirements.⁴⁴

Aidam and Sombié⁵⁶ who reported on the experience of the WAHO in improving the health research environment in the Economic Community of the West African States (ECOWA) region, said that there was no budget allocation for research in many of the member States. Also, some of the countries contract out their research programs instead of investing the funds for the contract to develop and maximize the capacity of their researchers.⁵⁶ Many of those countries were reported to be among those that consistently rank high on the global lists of corruption, and had extensive bureaucracies that further exacerbated their challenges.⁵⁸ Despite these challenges, the WAHO has continued to support Economic Community of West African States (ECOWAS) through providing research funding and capacity-building program in data management, analysis, utilization, and dissemination.^{47,56}

An editorial by Nature Medicine⁴⁹ on “sharing data to save lives” mentioned that based on their experience of data sharing during the Ebola epidemic in West Africa, there were government restrictions on the export of patient samples, and only scientists who had the cooperation of governments could ensure that data were released as quickly as possible. A study⁵⁸ on the health policy and system’s research and analysis in Nigeria revealed that the poor commitment of the government to follow through with plans and policies for health systems research hampered the materialization of any good will.

Social aspects. Health research and data sharing in Africa were reported to have some social peculiarities. It was reported that seeking the permission of elders, not just institutional permission from the government nor political custodians, before any important activities, including collecting and sharing data from specific communities was very important in Africa.^{54,65} Thus, the participants of the study by Anane-Sarpong et al.⁵⁴ on the application of ethical principles to research using public health data mentioned that it was important to get the community leaders actively involved in all research using their data. Also, Nordling^{51(p284)} stated:

Some people say the communities wouldn’t understand the research, Dandara says. But if they don’t understand, why are we researching them? Maybe it’s us who don’t understand.

Furthermore, it was reported that there are underlying systemic factors, embedded in historically and politically rooted structural issues that negatively impact data sharing in under-resourced African communities.³⁴ Globally, researchers receive different levels of support in the research environment they work in, and until such support becomes comparable, researchers in Africa will remain unequal as far as data sharing governance is concerned.³⁴

Also, conceptual differences across communities, cultures, and countries creates the need for harmonization, standardization of values,⁴⁴ and developing definitions for terms used in the data, which requires the “involvement of scientists who are working with the data, familiar with the context for the data, and have a stake in how the data are used.”⁴⁸ A South African study on public health data sharing^{41(p294)} revealed that some researchers in Africa felt anxious that data sharing may be similar to the:

neo-colonialist behavior... where the raw materials are taken out of the country and the beneficiation happens outside, and South Africa is the poorer for it.

Also, researchers from Kenya and South Africa reported the possibility of harm created by imprecise or stereotyped reporting of the data collected from African

communities^{41,48} Some senior researchers in this same South African study reported that their primary concern was ensuring the validity of secondary research. According to them,

it was imperative for the end user (secondary data user) to be able to exhibit on request that they have engaged with what they want to use the data for on a conceptual level.⁴¹⁽²⁹⁶⁾

This implies that the secondary data user must have critically reviewed the metadata catalogues and other documentation on the data nuances, so as to reduce chances of data misuse and misinterpretation.⁴⁶

Regarding sharing community health data with foreign researchers, it was reported that additional regulations to protect the interest of communities from where the data was generated should be applied.^{45,53,62} to ensure that they acknowledge the contributions, good partnership, and working together with the community.^{45,62} Therefore, data sharing practice should not be detrimental to the well-being of the community^{46,48,55} One of the participants in a South African study mentioned the need for a benefit sharing strategy so that the investigator will not just have patents generating billions of money while the community that provided the data are left to continue in poverty.⁴¹

Other participants were reported to have said that making data available for reuse (data sharing) demonstrates respect for the respondents, in that the researchers care about the opinion (data) of the respondents, and not just discard it after analyzing and using it.^{41,62} The need to disseminate the research output of shared data to the participant communities was also emphasized.⁴¹ However, giving feedback to participant communities can be challenging due to: the difference in openness and receptiveness of different communities and to feedbacks, data deidentification challenges, and lack of funds to provide feedback.^{41,48}

Of great importance was the report that prior individual awareness and consent of the participants from whom data was collected, were seen as important for data sharing to be done.^{29,45,48} Broad consent, where individual consent for their data to be shared was not feasible, was said to be a compromise and never an ideal, and only if linked to fair decision-making when data requests were made.^{48,51} Staunton et al.⁵³ reported that although broad consent is currently adopted for many genomic studies across Africa, its use is only proper if subject to appropriate oversight and governance procedures that foster trustworthiness. However, he wrote that with some exceptions, a general prohibition on the processing of “special information” that includes genetic data is imposed by Section 26 of the Protection of Personal Information Act (POPIA).⁵³ Furthermore, a study on a template for tiered informed consent for genomic research in Africa⁵⁰ reported many participants in Africa may be vulnerable, including those

with little access to health care and socioeconomic resources, disenfranchised women, persecuted ancestral groups, etc.⁵¹ Therefore, the researchers recommended that specialist advice be sought to ensure an appropriate informed consent process for the use of their data.⁵⁰

A study that explored the stakeholders views on the benefits and challenges to research data sharing in Kenya reported that the main potential harm data sharing can create for primary communities was stigmatization, where individuals, groups, or communities were identified and associated with potentially sensitive data.⁴⁵ This is because despite the efforts of researchers to make the data anonymous, so that individual participants cannot be linked with their data, the use of geopositioning data such as village names, name of the tribe, or ethnic group, etc., may sometimes bring stigma to the communities.^{45,48} Generally, it might be difficult to determine which data are sensitive or not. However, the researchers who participated in the study done in Kenya by Jao et al.⁴⁵ mentioned some data which could be seen as sensitive; clinical information, genomics data, sickle cell disease status, information about gender violence, sexual orientation, and sexual exploitation.⁴⁵

In a synopsis of some of the challenges related to the sharing of genomic data Moodley et al.⁶³ reported on the unexpected outcome of the sharing of data which was realized by South African researchers when they alerted the world about the discovery of the Covid-19 omicron variant. The resultant travel ban affected the country's economy and had an inhibiting effect on the country's research progress. The travel ban reaction led to the call for a collectivist data sharing approach in times of pandemics and warnings of the possibility of more dangerous pathogens emerging if a nationalistic viewpoint was allowed to continue.⁶³

Technological aspects. The National Academies of Sciences, Engineering and Medicine,⁴⁸ and Impouma et al.⁴⁴ reported that there are resource inequalities for data sharing in Africa and limited interoperability of the different data capture systems within the member states. The institutions lack the needed technical capacity to manage and share data,^{48,53} as well as the methodological capacity needed to analyze complex data, putting them at a disadvantage.⁴⁸ The lack of adequate resources was said to have impacted the compliance of research institutions and even individual researchers to data sharing, provision of training to student researchers on research ethics and data protection compliance.⁵³ Also, researchers struggle with limited research data management skills and maintaining data quality for longitudinal individual-level surveillance is a challenge in Africa, because of the highly mobile population with no unique individual identifiers.^{44,48}

However, unlike most African countries, South Africa has a linkable database which the Department of Health is trying to operationalize by setting up a preapproved

database and procedures for using it. This will increase researchers' access to the linked data.⁴⁸ Initiatives like Human, Heredity and Health Africa (H3Africa), INDEPTH, the pan-African bioinformatics network called H3ABioNet, and the Sierra Leone Ebola Database (SLED) focuses on promoting data management, storage, and analysis in Africa.^{34,43,48,62} According to the National Academies of Science, Engineering and Medicine,^{48(p25)} H3Africa operates on the following principles:

Maximizing the availability of research data in a timely and responsible manner, protecting the rights and privacy of human subjects who participated in research studies, recognizing the scientific contribution of researchers who generated the data, considering the nature and ethical aspects of proposed research while ensuring the timely release of data, and promoting deposition of genomic data in existing community data repositories whenever possible.⁴⁸

In addition to the limited manpower, equipment, lack of access to advanced technology, paucity of databases, and funding experienced by African health researchers,^{29,45,66} several research facilities in countries like Sudan, Zambia, and several West African countries have difficulties with internet access, software and equipment maintenance for laptops, scanners, etc.^{29,49,56,59} Research takes place in communities with unstable or no electricity and water supplies, poor communication, transportation and housing, and this affect all aspects of the research process.⁵⁶

Review question #3: What ethical, legal, institutional, and funding aspects are being considered by health researchers who are sharing health research data in Africa?

Ethical aspects. Anane-Sarpong et al.⁵⁴ raised the question of the rightful ownership of data as an integral baseline hurdle to overcome when ensuring ethical data sharing. This conundrum was seen from the perspective of the participant:

"If [the re-use of data] is anything besides what I initially consented to then I need to know and [be] informed. It's my right to know" and from the aspect of a data manager:

"We do not know what the value of data will be for different purposes and...from an administrative point of view to guarantee the use of data for specific purposes not practically possible."^{41 (p294-295)} It was also reported that a researcher queried the ethical ownership protections for work cited in a systematic review:

"...where primary researchers wonder if their research is being used, and if so, in a proper, relevant and pertinent way."

while another researcher called for the “foot soldier” to be treated fairly and acknowledged for the collection of data and given credit for work done.^{48(p13)}

The matter of informed consent was debated in the review literature. Manager-participants were unable to reach consensus on specific ways to gain consent for reuse of data and leaned toward the use of broad consent.⁴¹ However, one of the manager-participants in this study noted that:

...one of the drawbacks [to broad consent] is that the patient or participant is now less in the know...they really do not know what will happen to their data, so they are less informed to make a truly informed decision.^{41(p295)}

A study in Kenya by Mbuthia et al.^{46(p8)} showed that a few health care providers, hospital managers, and researchers inferred that if a patient came in to seek care this meant that they had implicitly granted consent for their data to be collected and used. In this regard, a researcher said:

Sometimes even you as an individual your rights sort of ends where there is a bigger purpose at hand.

A health care manager said:

By them coming in (or) choosing to come to the laboratory for services...in one way or the other it's like assumed you are offering this data, to be aggregated later.^{46(p7)}

In this study a health care manager also expressed, from a patient's perspective, why they might feel obligated to allow their data to be used:

“So, when I came here, I needed to be cured and you're telling me that I'm going to treat you but be aware that one of these days your data may be used for an evaluation. Of course, I'm forced to say yes, because it's like if I don't say yes, I won't get the treatment.” The manager also added “So, for me still ethically, it's not right.”^{46(p8)}

Similarly, it was shown that vulnerable health research participants might undergo a similar “bait and switch” situation whereby data collected for the health research were unethically used for ancestry related genome-research on a secondary basis.^{50(p1569)} These authors also suggested that specialist advice on informed consent procedure should be sought where research involved participants from vulnerable communities or groups in society. They raised the important issue of how to use informed consent in times of crisis such as when there is an outbreak of an infectious disease. The reuse of data may be critical to disease detection and treatment and the issue of what is most in the public's best interest may come into play. These authors recommended that ways to waiver informed

consent might have to be considered in such circumstances where individual rights may have to be overridden such as in cases where patients were unable to consent due to the severity of their illness. The authors proposed a tiered-consent model with reduced levels of consent might be considered in such instances.⁵⁰

One of the issues identified with the reuse of data was that the participant trusted the initial researcher with their data and believed that it would not be used in a harmful way.⁴⁸ It was suggested that trust may not carry forward in the case of the person reusing the data, so it is essential to respect the original consent terms when data is recycled, and this is especially pertinent in the case of historical data.⁴⁸ In Africa the issue of consent can be complicated and take time. Poor literacy plays into the understanding of what is meant by consent by possible participants and there is the need for a witness to be present to authenticate the participant's “mark.”⁶² Dhai et al.⁴² reported that biobanks were of specific beneficial use for the treatment of many diseases. They emphasized the need for the public to trust and be confident in such institutions especially with regard to ethical, social, and legal matters. They pointed out that South African law was not necessarily in sync with improvements in science and technology but RECs in some institutions in the country made sure that ethical oversight was carried out properly.

These authors⁴² also reported that the advent of biobanks has raised new problems regarding the issue of individual and anonymous informed consent such as: the nature of this type of research means that there are numerous participants; data are needed over an indefinite period of time, perhaps intergenerational periods of at least 25 years; cell lines may be generated from samples leading to duplication and exchange widely through networks; the conventional ideas of individual informed consent, particularly for ongoing reuse of data, cannot be applied. These authors^{42(p56)} suggested a different framework of ethical consent must be applied in these circumstances based on the “utilitarian common good” of using such data. Identification of origin of the samples poses the main risk to personal participants privacy especially with sensitive information. In Africa, this matter may present more difficulties as the definition of sensitive data varies widely throughout the continent.⁶¹

It was noted by the National Academies of Sciences, Engineering and Medicine⁴⁸ that recognition of the part played by a researcher in the collection, preparation, and sharing of data is linked to the intellectual ownership of that data. Anane-Sarpong et al.^{30(p399)} reported that:

“Issues of intellectual property, patenting and ownership” were critical: “A researcher who is tapping into the data of another should give credit where credit is due” and “Transactions should be mutual for everybody to be happy.”

In this regard, African researchers were seen to be often associated with the production of vast amounts of data but were not involved with data dissemination via publications or acknowledgement of participation in research projects.³⁴

To address shortcomings pertaining to African health researchers not getting adequate recognition or scientific exposure for their research efforts these researchers^{34(p91)} put forward six principles to enable and encourage data sharing within sub-Saharan Africa. The principles are:

Justice, respect for scientists whose data are shared, minimizing risks, maximizing benefits, collaborative partnership, and transparency.

In the interest of fairness, they also recommended that original data producers should be informed of other researcher's intentions to reuse their data. The same authors³⁴ noted that the lack of reward for data production is associated with poor data sharing in the south and some scientist may justify giving mandatory/obligatory data sharing another look based on the professional and systemic inequalities they experience. These authors warned against such activity believing it would lead to problems in research and collaborative relationships.³⁴ In this regard, it was suggested by researchers attending the workshop of the National Academies of Science, Engineering, and Medicine⁴⁸ that principles incorporating fairness, respect, and capacity building would assist in developing worthwhile collaborations.

Another pertinent issue identified by Nordling⁵¹ was that related to the use of African data by foreign, particularly northern hemisphere researchers especially during times of deadly disease outbreaks such as the 2014–2016 Ebola outbreak where many specimens were removed to foreign countries possibly without donor consent. To avert such situations, Conton⁵⁵ suggested that before a data repository is established and data contributors solicited certain rules of the game must be laid out regarding collaboration strategy to decide the governance of intellectual property, who will pay for analysis of data and who has the right to pass on results for commercial gain and who reaps such financial gain. The formation of databases and health networks in Africa have offered ethical challenges of their own especially in the case of anonymization of participants data. In this regard, it was found that the International Network for the Demographic Evaluation of Populations and their Health (INDEPTH) network was shown to be conforming with data published in a way that protects personal identity and anonymization of microdata.⁴⁸

In the study by Anane-Sarpong et al.,⁵⁴ an investigation was held into how the use of a Health and Demographic Surveillance System (HDSS) could inform Research using Public Health Data (RUPD). They found that, the fact that anonymization processes were introduced into the (INDEPTH) HDSS; many practitioners viewed the

HDSS and RUPD as already compliant in this regard and therefore thought it not necessary to check on the anonymization of data in these systems believing it is already the norm. The authors suggest that a specific ethical framework should be instituted to address such a problem. Nnamuchi^{62(p160)} reported on the ethical remit of H3Africa on the African continent and discussed that H3Africa, which deals with the management of biospecimens, preferred a broad based fully informed consent defined as:

Consent that allows the use of samples of genomic and phenotypic data for future research with ethics approval and the possibility to withdraw.

This type of consent still delineates timelines for sample storage and use. However, they also raised the issue of major concern in the way that anonymization of data collected in studies by H3Africa was carried out to ensure participant's confidentiality and privacy.⁶²

Impouma et al.⁴⁴ reported on problems with Information management practices in the WHO African region during the Covid-19 pandemic raising problems such as: the changing of reporting templates as pandemic grew; nonadherence to email nomenclature rules; poor email data security and confidentiality; wide use of poorly secured data spreadsheets with subsequent data loss; lack of uniform data collection standards; and easy access to personally identifiable information were identified and led to member countries being reluctant to share their data. The Covid-19 pandemic also further opened the debate on the ethical and legal frameworks around public health research and the restricted access versus unrestricted access to databases. In this regard, African researchers preferred databases where there were protections in place for those that generated and shared the data.⁶³

For ethical research to be carried out oversight by RECs is essential. Lötter and van Zyl⁵⁷ reported that due to the lack of or insufficient training of personnel, limited resources and a paucity of modern research management tools being available to allow good functioning and easy interaction with researchers many RECs in Africa were underdeveloped and that resulted in delays in or lack of research approval. These authors noted that to address this deficit, the WAHO had worked for five years in a strategic technical capacity with the Council on Health Research for Development (COHRED) to provide ethics training and thereby improve research capabilities. This was followed in South Africa by the REC of the Human Sciences Research Council (HSRC) ensuring that from 2011 planned data sharing and preservation were part of research protocols in that organization. Dhai et al.⁴² pointed out that South African REC guidelines make use of the wording of repository and biobank on an interchangeable basis, and this means that oversight by an REC is not a

mandatory procedure. Obiora et al.²⁹ found that bureaucracy and institutional red tape were identified as stumbling blocks to an African researcher's ability to share data. The secure holding of research for a period of only five years by institutions prohibited long-term usage of stored data.²⁹ Turcotte-Tremblay and Mc Sween-Cadieux⁶⁴ also showed that bureaucratic stakeholders may make the keeping of participants' confidentiality difficult during the dissemination of results due to a small pool of population participants being available, hierarchical concerns, a very limited number of research sites and their own vested interests.

Legal aspects. Staunton and De Stadler⁵² and Townsend et al.⁶¹ reported that privacy laws both across the African continent and globally are often disparate leading to the level and extent of data protection varying considerably. They noted that in Europe, the strict European Union regional Data Protection Regulation (GDPR), is in effect to regulate the use of data. Townsend et al.^{61(p22)} also reported that the African Union Convention on cyber security and personal data protection held in 2014 endeavored to lift Africa's "protection of personal privacy" to an international status but the recommendation included that appropriate cultural and social matters would be incorporated appropriate to the African setting. The same authors⁶⁰ said that currently 25 African countries have enacted privacy laws that regulate how personal data is collected and used. Of the 25 countries 12 have recently put such laws in place or have amended existing laws. They include: Burkino Faso, Cape Verde, Cote d'Ivoire, Gabon, Ghana, Lesotho, Mali, Mauritius, Morocco, Seychelles, South Africa, and Tunisia.

Nordling⁶⁵ discussed whether the South African POPIA was prohibitive to the free collection and sharing of research data. The South African POPIA was based on the personal data protection recommendations in an early draft of the GDPR. Problems with freely conducting research arising from the earlier GDPR were addressed in Europe but POPIA was not adjusted. The law was created in 2013 to protect South Africans' personal information and was enacted when a year-long grace period for compliance ended on 30 June 2021. The POPIA raises the issue of whether broad consent can be obtained legally from those participating in a study and suggests that legal advice should be sought to avoid transgressing local legislation.

Dhai et al.⁴² reported on legal aspects relating to the establishment of biobanks in South Africa noting that the South African National Health Act (NHA) and its regulations dictated the legalities regarding the use of human tissue and research in that regard. In addition, four studies looked at the way in which POPIA might limit genomics and biobank research.^{51-53,65} Three of these studies recommended that POPIA provisions regarding research might have to be adjusted to avoid the upending of research

projects and creation of problems with international collaborations.^{51,53,65} Nordling,⁶⁵ in particular, raised the issue of obtaining broad consent by study participants so researchers could store and analyze samples for unspecified usage stating that local scientists were lobbying to get a POPIA exemption for this practice from the government. Their argument was that POPIA as it stands could hamper valuable research on such diseases such as tuberculosis and HIV.

Staunton and De Stadler⁵² stated that POPIA was focused on the protection of individual data in an administrative sense and that a code of conduct should be considered for researchers regarding obligations and protections necessary in research. This was particularly applicable in relation to not only protecting the individual's data but also that of the community as open science became more common. A suggestion to address this problem was that, under section 55 of POPIA, information officers could be appointed to assist institutions to adapt to such challenges. Staunton et al.⁵³ reported that in the case of POPIA and biobanks Section 8 of the Act would suggest that institutions had to assume overall responsibility for handling compliance with the act and would be subject to paying fines resulting from breaches of compliance. Townsend et al.⁶¹ recommended that consideration should be given to the setting up of data trusts and data governance regulations by means of new legal data management entities.

According to Uzochukwu et al.,⁵⁸ in 2001 the National Health Research policy was formulated in Nigeria in an endeavor to increase the uptake of Health Policy and Systems Research and Analysis (HPSR+A). This document has not been implemented and remains as a draft. The authors note that even though collaboration between the Federal Government of Nigeria and a Canadian entity resulted in the National Strategic Health Development plan in 2010 there is slow implementation of this due to poor follow through on decisions. This lack of operational planning and strategy implementation leaves the government's stewardship of research development and policy strategy awry. This impacts the management of knowledge and research and slows the generation of HPSR +A and its use.

Cole et al.⁴⁰ drew attention to the fact that none of their participants considered that data should be shared with the private sector as they felt that it should not be shared for profit, and it should rather be kept within the realm of academia or offered for the good of the local citizens. The participants in this study also believed that the sharing of data nationally had to be well governed to avoid data being exploited. The issue of how sponsors or funders were contractually obligated to guard privileged data that they had access to was also raised.

The Human Sciences Research Council Act 17 of 2008 was instituted in South Africa to lay out the objectives and

purposes of the HSRC.⁵⁷ The institution was tasked with directing the way forward for research in the country. In one of the clauses, it was said that the HSRC was required:

To develop and make publicly available new data sets to underpin research, policy development and public discussion of the key issues of development and to develop new and improved methodologies for use in their development.^{57(p339)}

Institutional aspects. Cole et al.⁴⁰ reported that such institutional platforms as the Health Research Capacity Strengthening Initiative (HRCSI) situated in Malawi developed research capacity, increased the number of highly rated researchers in the country, encouraged young scientists to be interested in research, and provided support for the conduction of research. It also fostered the sharing of knowledge by helping delegates attend both institutional and national meetings or conferences. The platform was actively marketed resulting in calls for research being well supported. As a manager said:

“Non academics, are asking when the next HRCSI calls are coming out which shows the demand for research” and “Funds were made available to institutions and people that would not have been able to access them.”^{40(p8)}

Anane-Sarpong et al.³⁴ put forward the idea that where people from different backgrounds collaborated in the research space there was strengthened rigor, quality, and quantity of research data produced. This sharing of data within teams was effective in creating new knowledge sources and processes, encouraged the sharing of data to negate data ownership conflicts, promoted networking, and increased productivity, resulted in communities getting improved feedback and evidence-based health care quicker and resulted in easier policy decisions. The same authors also pointed out that in undersourced settings there is a decreased adherence to data sharing.

In a study by Denny et al.,^{41(p293)} participants described data sharing as “ad hoc” decisions of an informal nature or formalized institutional interactions with agreements struck between the data sharing body, the primary researcher, and any funding entity. Data curation activities, which included the reuse of “donor-funded research data” and planned data sharing was encouraged. Participants in the same study raised the issue of obstacles to institutionalized data sharing. A senior manager noted that:

People didn’t really think about data sharing [as] I think they were less understanding of the benefits...

This participant added that an open data policy presented:

“Challenges, in terms of [changing] people’s attitudes and [growing] a sharing culture” within the institution.^{41(p293)}

It was recommended that participants should receive feedback regarding any sharing of their data and that institutional RECs should make sure that this practice was put in place and adhered to.

Lötter and van Zyl⁵⁷ recommended that to encourage research production and dissemination institutional repositories should be established, internet infrastructure should be improved with alternate power being supplied during outages and computer access being available, the use of open access should be enhanced by means of mandated policies and open access information should be supplied in many ways by librarians and they should also instruct on copyright matters so researchers could safely archive their research.

In this regard, Lwoga and Questier⁶⁷ looked at the matter of open access in institutions in Tanzania. The study findings indicated that open access was poorly adopted because few librarians were associated with open access activities, over the previous five years few materials for faculty research were made available on open access platforms, there was poor technical infrastructure, poor knowledge of open access issues, skills were lacking for publishing online, and researchers were not well-versed in the author-pay model of submitting for publication and scared or had misconceptions regarding journal copyright and plagiarism instructions.

A participant in the National Academies of Science, Engineering and Medicine Conference⁴⁸ raised the point that in the Global North, institutions were quick to implement data sharing as they had financial and other resources to put such practices, platforms, metrics, and policies in place. In Africa, where often resources were wanting, it was harder to get such structures up and running and maintained going forward. Another participant argued for the need for capacity that would lead to the generation of quality data, increased ability to process such data and the capacity to utilize it to encourage and facilitate data sharing.⁴⁸ She suggested that to enable data sharing institutions should develop correct research guidelines, policies, and research contract procedures as per those used by the COHRED. This participant also recommended that in place of numerous institutions archiving data consideration should be given to expanding regional institutional archiving such as that found in the African Population and Health Research Center (APHRC) and the focus should be on the generation of data and its curation analysis and management. The idea put forward was that this approach would ultimately lead to data sharing as an outcome. Townsend et al.⁶¹ also recommended that the sterling work by the African Union and other African regional African bodies should be carried forward to develop greater collaboration and integration regionally on the African Continent.

A study by Lwoga and Questier⁶⁷ reported on a recommendation to WAHO that there should be the formation of a network of institutions regionally resulted in the West African Health Research Network (WAHRNET). This network, established in 2010, consisted of 22 medical schools and 30 research institutions and was recognized in 2012 by the Assembly of Health Ministers affiliated ECOWAS. Lötter and van Zyl⁵⁷ found that in 2006 the HSRC in South Africa began a process to manage data better, preserve data and make it available for future use. In 2008, the initial research team moved to developing a long-term plan for data curation, preservation, and dissemination. By the time of their study in 2015, 87 data sets had been curated by the HSRC since 2007. Data are available to the public but under strict ethical protections and conditions. These authors further recommended that correct training should be given to postgraduate students and supervisors so that they were aware of the ethical implications of using data again in future research and that all researchers should know how to plan for reuse of data.⁵⁷

Gorina et al.⁴³ looked at the SLED in relationship to the protection of sensitive information. This anonymity not only extended to the individual patients whose data were in the database but also protected the institutions that were supplying data and using the SLED database for statistical purposes. There was also protection from countries being identified due to unique cultural information. The database used behavioral and technical methods to ensure confidentiality of data.

Jaο et al.⁴⁵ reported that the issue of difficulty in accessing the technical expertise for data analysis was not merely linked to LMICs. In fact, it was rather related to poorly resourced and funded research institutions. A Kenyan participant in this study said:

If it's open access, someone else who has the skills that you may not necessarily have at the moment could very quickly do an analysis that you'd want to do... someone in Harvard could walk very quickly to a next-door neighbor and get that analysis done overnight and published.^{45(p9)}

Poor opportunities related to data analysis was not the only problem facing researchers. Lwoga and Questier⁶⁷ reported that some research institutions could not carry out research projects due to insufficient capacity to manage such projects. This was both in the areas of managerial skill and technical skill with poor data retrieval methods and archiving abilities.

Uzochukwu et al.⁵⁸ reported on the teaching of HPSR+A in four institutions in Nigeria to enhance the capacity, needs, assets, and perspectives of researchers and health policy makers. Problems which were experienced with infrastructure included poor electrical supply and difficulty getting research tools such as e-journals and computer software. These authors reported that to implement the use of HPSR+A there was a

need for institutions to first identify researcher groups and then encourage networking and research collaboration among the groups. It was suggested that research activities needed a framework, a research group database needed to be set up, a managed online library should be set up, and calls for research protocols should be announced.

Staunton and De Stadler⁵² pointed out that institutions such as universities experienced difficulties with standardizing and centralizing research compliance and ethics with data sharing. They noted that in the South African situation researchers were inadequately trained in these issues and that government guidelines regarding regulations on managing health related data were poor. Universities often fail to comply with data protection law and in this regard, lump research data handling into the same category as administrative data.

Mbuthia et al.⁴⁶ argued that in Kenyan public hospitals it was preferable that patients were openly told about possible reuse of their data and what it might be used for as this might instill confidence in the health care system and encourage them to give more pertinent information when describing their symptoms. A health provider said^{45.(p6)}

If they are asked and they give ...all the information they have... they'll feel happy that...somebody else has gained from it.^{45(p6)}

The reuse of clinical data was hampered by certain factors. In this regard, a health care manager said:

What we need to appreciate is that most of the public hospitals, they are congested...and the health workers are few. So, if we say we are going to explain to our patients what we are going to do with their clinical data, it might take long and at the end of the day very few patients will be served.^{46(p7)}

Another factor raised was that costs could be significant putting in place those personnel that could communicate and explain complex issues such as deidentification and what the data might be used for in future research. The possibility that patients might have low levels of education with limited formal education was also noted.

Funding aspects. Anane Sarpong et al.³⁴ reported that resource constraints in sub-Saharan Africa led to the limitation of large-scale collection of data pertaining to health research. Research funding needed to be supported where sponsorship was limited so the burden did not fall on the shoulders of the individual researcher. Obiora et al.²⁹ listed reasons as barriers to data sharing in Africa such as: poorly resourced institutions; researchers being poorly paid; costly research not being carried out in poor settings; costly data and a paucity of funding opportunities.

Shabi et al.⁶⁶ in Nigeria did not find financial restrictions to the use of the internet or purchasing of data. This gave researchers access to online digital platforms such as HINARI. A participant in the National Academies of Science, Engineering and Medicine⁴⁸ conference who was an African researcher in a nongovernmental institution said:

“It is important to think of data sharing in the context of funding cycles.” She also said that at the start of a year her institution had: “no committed funds for anything” but “Toward the end of the year, the focus of the board is how to continue to cover the existing staff the following year. Having conversations around data sharing is not a priority. In addition, funding for projects versus funding for core activities is a huge issue.”

As data sharing is rated as a “core support function” it is regarded as a “luxury” and that prohibits data sharing even though people want to do so.^{48(p21)} In this regard, Denny et al.⁴¹ recommended that research institutions should be granted funding opportunities or that costs should be subsidized to promote data sharing practices and afford data curation to take place. They also recommended that the cost of data curation should be covered in the standard budget of all research grants and that RECs should require it for ethical approval. The Nature Medicine Editorial⁴⁹ suggested that once guidelines were in place so data producers and users came together funders would create policies to ensure that the guidelines were followed enabling data producers to be encouraged to share their data.

Cole et al.⁴⁰ noted that Malawi’s HRCSI funding was slow as often funding applicants in that country had to wait for over a year in some cases for awards to be made and monies granted. A manager said:

HRCSI lost credibility with stakeholders.^{40(p5)}

The methodology was improved in 2011 with an HRCSI board being established but its decision still had to be approved by the original funders which led a producer of research to say:

HRCSI money comes from the Wellcome Trust [partially] and takes much time.^{40(p6)}

In 2013, the HRCSI lost its external funding. Aidam and Sombié⁵⁶ reported on WAHO’s contribution to a research development program that ran in West Africa from 2009–2013. The funding was in the region of US\$857 000 for 24 critical research projects including nine in the field of HIV and tuberculosis, three in both health financing and malaria and one in each of the fields of noncommunicable disease, typhoid fever, maternal health, sickle cell disease, dengue, and medicinal plants.

The storing of samples and curation of data in a database scenario is expensive. In Anane-Sarpong et al.,^{30(p399)} a participant said:

because data is maintained at a cost, there should be a fee for use. You have to contribute to make sure we keep it going.

The need for funding support for an Ebola database in Sierra Leone led Conton⁵⁴ to suggest that similar to the way in which biobanks in the developed world charged fees to remain sustainable, consideration should be given to charging fees to support such a database. It was also suggested by a participant in the National Academies of Science, Engineering and Medicine⁴⁸ conference that HDSS data would increase by the sharing of data costs as opposed to charging for data for longitudinal studies especially in these open access times.

Nordling⁵¹ looked at the fact that data sharing agreements between African researchers and those elsewhere internationally often have a skewed balance of power. This particularly in the case of younger African scientists. Powerful donors or research partners may lead to African researchers not questioning their role in research for fear of losing funding opportunities.

Discussion

This scoping review sought to delineate factors that influence how African health researchers share their data and discuss the relevance of certain barriers or facilitators to the conduction of data sharing among this research population. It was apparent that many of the factors under consideration impacted more than one aspect of this review’s findings with an overlap of influence on the findings inevitable. One such example was the fact that ethical considerations and legal decisions often had social impact and that institutions had to have reputable RECs to ensure that ethical research was conducted regarding both the collection of research data and the reuse of such data in the process of data sharing. Hence, this discussion will endeavor to contrast and compare findings while considering the interwoven nature of the complexities of data sharing by health researchers in Africa.

The 32 included studies were very varied in the focus of their reporting although their outcomes were relevant to the overall topic of the scoping review. The review findings were reported in line with the answering of the review research questions to present all available information in a clear and concise manner, to reach conclusions that would possibly lead to new insight into what factors are barriers to data sharing on the continent and to identify facilitatory factors to increase data sharing going forward. Most studies were carried out in 2015 (28.9%),^{41,42,45,48,49,57,67} followed by 18.8% in 2018^{30,47,54,59,62,64} and 2019,^{46,50–53,65} 9.4%

in 2016^{38,56,58} and 2022,^{29,61,63} 6.3% in 2017^{55,60} and 2020^{34,43} and 3.1% in 2011,⁶⁶ 2014,⁶⁸ and 2021.⁴⁴ Hence, a downward trend in the production of information on the topic of data sharing among African health researchers was noticed after 2019. The results from 2020 to 2022 may reflect the fact that worldwide the focus of research turned to reporting on the Covid-19 pandemic. However, one might speculate that interest in data sharing in Africa may have met an impasse due to the many technological, financial, and bureaucratic barriers that African health researchers face in trying to gather data, analyze data, publish data, and share data in other ways.

Geographical locations and data sharing practices

The African continent is vast and includes 54 independent countries.¹³ However, the production and sharing of research data were found to be limited or nonexistent in certain regions. In this regard, the search for information on data sharing among health researchers in central Africa found one article⁴⁴ and only one article was sourced from Malawi.⁴⁰ No countries in North Africa contributed to the literature available on the topic under review. Similarly, Obiora, Shead, and Olivier,²⁹ sought information on data sharing by African health researchers by means of qualitative interviews, but they were unable to get participants from central Africa and only managed to interview one Malawian participant and one Sudanese (North African) participant. These same authors found that most of their interviewees were from South Africa and Fonkou et al.¹¹ also noted that most intra and extra-African collaboration and production of research literature occurred in South Africa. Likewise, in this review it was found that the production and distribution of research was most prolific from researchers based in South Africa.^{29,41,42,48,50,52,53,57,60,61,63,65}

Types of data and data sharing platforms

There were widely differing data shared by health care researchers. These can be grouped as data collected in the public health arena and those collected on a more individual personal health-related basis. These data included results from surveys, epidemiological data, demographic data, data from human databases and repositories, EHR, and data collected by means of digital devices or by using medical sensors among others.

Data sharing platforms, databases, and repositories were to be found on the African continent and yet Obiora et al.^{29(p5)} had participants who had not come across such databases and said:

“I’m not sure we have a platform like that,” and “There is currently nothing.”

This study uncovered the fact that existing databases or data repositories containing health-related data varied widely in their regional location and in the types of data that they were storing. In this regard, many databases/repositories had been set up in relation to epidemic disease challenges such as the SLED,^{43,55} the National Health Research for Development (R4D) platform,⁵⁴ the Infectious Diseases Data Registry (IDDO),⁴¹ and the Epidemic Intelligence from Open Sources (EIOS).⁵⁶ Conton,⁵⁵ a pioneer in the effort to establish an Ebola database in Africa, argued that research on the disease should be carried out in Africa for the benefit of African communities afflicted by it instead of specimens being removed to other countries like the United Kingdom with the added concern that sometimes this was not done in an ethical manner. Other platforms, such as the South African Research Council (SAMRC), Malawi HRC SI, and the HDSS, are for the gathering of health-related population data. Data from the HDSS can be used for Research using Public Health Data (RUPD)⁵⁴ or entered in data repositories such as the International Network for the Demographic Evaluation of Populations and their Health (INDEPTH).⁵⁴ With the advent of increased biobanking in Africa there has been a necessity for the formation of the Human, Heredity and Health in Africa platform (H3Africa) to facilitate the handling of genomic human research data.^{52,53} This current review found that West African countries and South Africa were found to have the greatest number of data sharing platforms, databases, or data repositories.

Individual research capacity concerns

Despite efforts to build research capacity over decades more researchers are needed in Africa.¹⁵ Measurement of health research capacity is currently dependent on metrics such as: resources available for health research; number of publications, clinical trials, research institutions, and research personnel policies and regulations in place.⁶⁹ The calculation of research personnel has no standard measure and may include those with PhDs or others with designations such as laboratory technicians.⁶⁹ Availability of funding is another factor that influences the number of active researchers in a society. It is suggested that the United Nations Educational, Scientific and Cultural Organization (UNESCO) is a fairly comprehensive source of information on researchers per capita of the population of an African country.⁶⁹ In 2015, UNESCO⁷⁰ reported that per million people globally there was an average of 1478 researchers. This included those doing research in many scientific fields. At that time only Tunisia in Africa exceeded that statistic with 2000 researchers. Other African countries’ totals of researchers per a million people were listed in descending order as Morocco (1100), Egypt (680), Senegal (550), and South Africa (494). But the majority of Sub-Saharan African countries had an average of less than 50 researchers

per a million people.^{15,70} In 2021, Wenham et al.^{71(p7)} tabulated African health research data based on findings formulated from information on the WHO's Global Observatory on Health Research and Development (GERD) as a proportion of Gross Domestic Product (GDP), number of institutional facilities and the degree of official medical research and grass roots health assistance proportional to gross national income. This showed in descending order that the 12 African countries falling into the top tercile of the range with regard to number of health researchers per one million people were: Tunisia (1965), Morocco (1069), Egypt (680), Senegal (549), and South Africa (473). Eleven countries fell into the middle tercile and 12 into the low tercile. Rwanda (12), Niger (seven), and the Democratic Republic of the Congo (seven) were shown to have the least numbers of researchers per one million people.⁷⁰ In their table Wenham et al.^{71(p7)} indicated that overall 39 sub-Saharan African countries (72% of the total African countries) showed a researcher per million of the population quota of under 50 with 20 of these countries not even registering any data based on this metric.

Individual researchers in Africa face enormous challenges due to inadequate facilities, limited access to technology and a paucity of funding and resources. These factors were shown to sometimes present insurmountable barriers to the conduction of research particularly in lower-income countries like Zambia, Sudan and a number of West African countries.^{29,49,56,59,66} A study in Kenya by Jao et al.⁴⁵ found that most researchers and community helpers wanted to share research. However, other studies reported that the eagerness to share data was expressed more by management structures and not by individual researchers who were hesitant, reluctant, and unsure about sharing their data.^{30,34,41,46,55} A reason for the reticence of African researchers to share their data was that when they collected data there was either poor or lacking resources for analysis of said data.^{29,34,45,57,68}

The fact that this data is underutilized when initially produced can lead to the fear by under resourced African researchers that if they share their data it will be poached by better resourced researchers who will be able to analyze it and publish it without proper acknowledgement of the original researchers input.^{34,48} Furthermore, instances whereby data were collected by African researchers but then samples were hastily transported to first world countries for interpretation with no acknowledgement of the African input also led to the continents researchers being ignored and being unable to publish their own findings. Such "helicopter research" practices led to the suggestion that an Ebola database should be built in Africa instead of England to avoid unethical practices related to the speedy collection of patient's data and samples and the noninclusion of African researchers in publications.⁵⁵

The National Academies of Sciences, Engineering and Medicine⁴⁸ on data sharing in Africa, reported that it is

more likely that somebody from Harvard will want to analyze data from an African country than somebody from Africa requesting data from America. This may be because although data from the United States are available in well managed data sharing platforms, it is more likely to get data that has not been analyzed or is poorly exploited from African countries. On the other hand, Jao et al.⁴⁵ mentioned that researchers from sub-Saharan Africa felt that sharing data mainly among themselves (coresearchers in sub-Saharan Africa) will help build local scientific capacity, and by so doing increase public acceptability of data sharing.

African researchers expressed the need to use their data for publication prior to the end of their degree process and felt that having the protection of a DOI number would show they owned the data and allay some of those fears.²⁹ Globally, a similar view was seen as there was some reticence among researchers to share their data citing intellectual property and privacy concerns.⁷⁰ Some review studies' authors felt that trust issues, worries about fairness of data distribution, lack of reciprocal acknowledgement, and inclusion problems were reasons for researchers being reluctant to share their data. In that regard, Conrad, Delahunty, and Ding⁷² put forward that researchers, universities, publishers, and funders when entering into data sharing should adopt the "Findability, Accessibility, Interoperability and Reusable" principles proposed by Wilkinson et al.,⁵ and Bull et al.⁶ addressed the need for trust and confidence to be built between collaborators in LMICs in order to facilitate data sharing.

Matters related to research institutions

An institution that is well resourced can foster an environment where researchers from differing backgrounds can collaborate with each other and share data.³⁴ Such a space can encourage networking and increase productivity, negate conflicts over intellectual property rights, improve feedback to communities, increase knowledge on topics of interest, and lead to better health care using evidence-based practice.^{29,34} The cross-pollination of research ideas and outcomes can lead to increased rigor in data collection and increase the quantity and quality of research data collected. However, participants in the study by Denny et al.⁴¹ raised obstacles that might arise regarding data sharing in an institutionalized setting such as people's attitudes not being pro data sharing and the lack of a data sharing culture.

The National Academies of Science, Engineering, and Medicine⁴⁸ attendees suggested that in the poorly resourced setting in Africa it was difficult to implement data sharing as the institutions lacked the practices, platforms, and policy structures that were easily put in place in the global north institutions. An attendee at this conference also recommended that COHRED research contract

procedures and policies might assist in the establishment of better data sharing opportunities. She further suggested that the way forward to facilitate data sharing would be to expand regional institutional archiving such as that found in APHRC. In this regard, Townsend et al.⁶¹ were also in favor of the generation of regional cooperation and collaboration under the banner of the African Union or similar African bodies.

While lack of access to adequate technology was seen as a barrier to institutional research capability it was noted that this arose from the lack of resources and inability to source technical expertise in poorly funded institutions.⁴⁵ These authors also found that information on or the finding of personnel to assist with statistical analysis of data was difficult to access resulting in delays in finalizing outcomes and being able to present and possibly publish findings. In Tanzania, the matter of poor management of projects with resultant inadequate data retrieval and archiving of data was attributed to lack of appropriate technical and management skills.⁶⁷

Uzochukwu et al.⁵⁸ tried to implement the use of HPSR +A in four Nigerian institutions but realized that before such a system could be utilized infrastructure had to be up to the task of supporting such an implementation. He realized that problems such as poor electrical supply, lack of computer software, and access to e-journals had to be addressed prior to the roll out of the program. He also found that a research framework had to be designed, researcher groups needed to be established and encouraged to collaborate, a researcher database had to be set up, an online library created and research protocols had to be called for.

Technological concerns

The technological landscape of Africa differs widely with low-income countries like Sudan having interrupted power supply, that impacts the use of computers which in turn are in poor supply.²⁹ The different health data capture systems are not uniform across the continent or even within certain countries, so sharing of data is difficult or impossible and differing resources make the managing of the sharing of data a challenge.^{44,48,53} In this regard, Anane Sarpong et al.^{30,54} described the HDSS system that has been set up in Ghana, Tanzania and Ethiopia, Lötter and van Zyl⁵⁷ talked to the use of District Health Information Software 2 in South Africa, Jao et al.⁴⁵ referred to clinical surveillance databases in Kenya and Nnamuchi⁶² highlighted the use of EHR in Nigeria. In 2021, The National Academies of Science, Engineering, and Medicine⁴⁸ reported that South Africa had a linked health database and the Department of Health was attempting to ensure procedures to increase researchers access for the use of such linked data. However, in 2021 Tsegaye and Flowergay⁷³

reported on challenges to interoperability in South Africa, proposed a new architectural framework for addressing the problems and indicated a way to ensure a national interoperable EHR system. But it is not just the material inadequacies that delay or prohibit the sharing of data the human element also plays a role where researchers do not always have data management skills.^{44,48}

Funding considerations

In order for health data to be shared it first has to be collected. The fact that 85% of the countries in Africa fall into the low income and low middle income categories explains why money availability for the funding of research is often very limited or unattainable.¹² Lack of funding for research not only impacts the individual's propensity for gathering research data but it dictates how institutions expand their knowledge base and improve the teaching of students. Obiora et al.²⁹ also reported that their participants had raised the issues of how inadequate funding led to poorly paid researchers, lack of resources in institutions and difficulty in carrying out research in impoverished settings. This can be a real impediment to early career researchers developing and furthering their careers.⁷⁴

Anane-Sarpong et al.³⁴ reported that the large-scale collection of health data, so important for public health policies and for the formulation of treatment criteria, was sorely impeded by resource constraints caused by poor funding. Omungo⁷⁵ reported that a very low percentage (2%) of African researchers from only a few countries received significant grants of over a million dollars to aid with their research. This author also noted that most African researchers did not get any research funding, leading to self-funded research occurring in the majority of cases. However, he did comment that health sciences research stood a higher than average chance to be given grants by European and American funding institutions.

International funding entities such as the National Institutes of Health in the United States,⁷⁶ UK Research and Innovation⁷⁷ and Wellcome Trust⁷⁸ issue calls for grant applications but these offerings are usually oversubscribed by researchers seeking funding. An applicant may also have to do the research in the country where the grant is offered and if funding for accommodation, travel, and other basic requirements is not supplied the applicant may not be able to take up the grant opportunity anyway. In South Africa, the National Research Fund⁷⁹ offers researchers funding opportunities in an ongoing program. In the past in Malawi the HRCIS offered funding supported by the Wellcome Trust but unfortunately, in 2013 this funding source was discontinued.⁴⁰ This scheme was criticized for money only being supplied to researchers after a protracted length of time. Similarly, funding supplied by WAHO to aid research development in West Africa ended in the same year.⁵⁶

No other major funding schemes were uncovered by this review pointing to the fact that international support for research in Africa might be waning and this does not bode well for research on the continent. This fact may also indicate that for Africa to step up to the challenge of producing world class research output and building research capacity, it will be reliant on more funding from its own entities and governments.

Issues related to journals and journal editors

Conrad, Delahunty and Ding⁷² in their capacity as professional publishers suggested that there are many benefits to the reuse of original data and that the open science principle can no longer be ignored by their peers. They draw attention to how the recent Covid-19 pandemic demanded that data sharing and reuse became the norm rather than the exception globally. The need for research to move quickly in addressing such health crises has already been mirrored by the hastened use and reuse of data and its entry into SLED such as took place during Ebola outbreaks in Africa.^{43,55} The workshop summary that came from the meeting of members of the National Academies of Sciences, Engineering and Medicine⁴⁸ stated that researchers opined that data that was released prior to peer review would be deemed to be nonnovel. This would impact the release of data such as that collected in a health crisis like an Ebola outbreak and prevent dissemination of data for the public health good. Representatives from leading biomedical publications met and decided that this would be detrimental to public health and that the fast-track release of data would not prejudice journal acceptance and publication.

It is tempting for African researchers to be lured into the Open Access world and be enticed by offers of publication from predatory journals as publication in reputable journals is often prohibitively, costly. However, there are efforts from reputable journals and the WHO to try and get easier access and cheaper rates for African researchers to publish.³⁴ Similarly, WAHO assisted with payment of the publication of four international journal articles in West Africa.⁵⁶ The open access era has seen publication of research in journals and the payment by author model being adopted. In Africa, the submission for publication in such databases or journals is difficult due to prohibitive costs for the author and ambiguous descriptions of submission procedures.⁶⁸ On the other hand, the sourcing from open access data bases and journals has facilitated research in Africa as researchers can source such information for free. One problem that was identified by Lwoga and Questier⁶⁷ in Tanzania was that few librarians were familiar with the open access format, faculty open access research materials were in short supply and poor technological infrastructure meant that access to open access online material was limited or absent.

The fact that a Tanzanian institution hosted their own local journal was reported by Lwoga and Questier⁶⁷ and these authors added information on the practices of journals where publishers unbeknown to authors allowed archiving of pre- and postprint articles. The fact that authors were unaware of the construct of the permissions granted to them related to dissemination of different versions of an article once it had been published was also raised by Lwoga and Questier.⁶⁷

Anane-Sarpong et al.³⁰ reported that several journals now require data to be shared as a proviso for publication, and the guidelines applied to such a request do not have many boundaries. Researchers are trying to counteract the possible loss of their data by withholding novel data including new ideas that will involve the disclosure of much data and perhaps even metadata, keeping data for future projects and not sending it for publishing and archiving manuscripts on similar research topics and submitting them together to ensure that credit is given to the original authors for the research outcomes. However, contrarily researchers in Kenya were supportive of the sharing of data sets that underpin data analysis subject to the primary researchers who collected the original data being recognized in any future publications arising from the use of their data.⁴⁵ In a possible solution to such a conundrum, Shaffer et al.⁵⁹ reporting on the H3Africa group, noted that H3Africa applies an approximate 23-month publication embargo period where the researchers that produced the data had an exclusive right to analyze that data. This enables under-resourced African researchers time to utilize, analyze, and publish their data to avoid other researchers in Africa or elsewhere in the world from “scooping” such data for their own publications.

Ethical, legal, governmental, and social aspects of data sharing by African researchers

These aspects to be considered in relation to data sharing by African researchers interlink with each other. However, they also dictate how data collection is influenced both from the viewpoint of the individual African researcher and the research institution. In current times collection of much of research data is associated with technological devices and data storage facilities. The digital age has presented the keepers of data security and moral boundaries with new challenges. This change has necessitated a shift in how ethical values and social protections are handled.

Ethics in relation to the collection of data is guided by certain principles. In the case of the initial collection of health data the process of obtaining informed consent is nonnegotiable and must be carried out following specific guidelines.^{80(p204)} However, when data is to be reused or distributed so other researchers can use it to carry out further research problems arise. These problems are not

unique to the African continent but do impact how data is shared on the continent.

The basic question of “who owns data” put forward in Anane-Sarpong et al.⁵⁴ is integral to the understanding of how data could be shared. Researchers were concerned as to whether a participant’s initial consent covered the further use of such data for another purpose and how the specific reuse of data could be predicted. The question was raised regarding what the participant’s right was regarding being informed of the intention to reuse their data, in what way it would be used and for what purpose this would be carried out.⁴¹ These authors raised the concept of broad consent where permission to redistribute data was covered by a blanket-like consent at the time of original consent being obtained. A participant in their study argued that this concept left a participant blind to the many possibilities for future reuse of their data.

Rather concerningly, Mbithia et al.⁴⁶ reported on how in Kenya data collected in hospitals or laboratories might be considered as the property of the institution merely because the patient had come in for treatment or testing. It was a quid pro quo situation where the patient received the treatment in return for giving up their data. This meant that the individual patient had very little control over the use of their data. This blatant disregard for the ethical rights of the patient overlapped with the social aspects of data use where the individual in a community had the right to decide how their own data would be used in the future, how the need to prevent harm to such an individual or community needed to be addressed and how trust of participants had to be guaranteed and maintained.⁴⁸ This type of “bait and switch” situation was highlighted further by Nembaware⁵⁰ in relation to ancestral-related genome research in the community or societal groups. Dhai et al.⁴² also wrote about challenges related to collection of data for placement in biobanks. The large number of participants, and the prolonged use of their data, maybe up to 25 years, across various networks raised peculiar problems with obtaining informed consent for data reuse.

Research Ethics Committees in institutions are not only bound by the institutional ethical rules and regulations but also by the law of the land as it impacts ethical research practices.

Data protection and privacy laws globally and across Africa vary considerably in the level and extent of data protection.^{52,61} In the United States, the Health Insurance Portability and Accountability Act of 1996,⁸¹ GDPR⁸² protects patients’ health data and privacy. In Europe the GDPR effects the ethical use of health data. This review found that currently only 25 of the 54 independent African countries had some type of data protection in place to ensure anonymity and privacy of participants. However, only 12 (Tunisia, South Africa, Seychelles, Morocco, Mauritius, Mali, Lesotho, Ghana, Gabon, Cote d’Ivoire Cape Verde, and Burkina Faso) of these countries have enacted laws or

amended previously passed laws in this regard to conform to international standards of data protection. The most notable contribution to African data protection laws is the South African POPIA that was created in 2013 but only enacted on 30 June 2021.^{53,65} This Act was based on an early draft of the GDPR. However, the GDPR was amended in line with points raised regarding problems experienced by researchers in Europe related to ongoing data sharing, but this was not the case with POPIA.

The issue of biobank research further complicated the POPIA implementation as there was some malalignment between POPIA dictates and the Human Tissue Act of 2004, which regulates legalities relating to the use of human tissue in research.⁴² Four studies looked at how POPIA might limit biobank and genomics research.^{51–53,65} This has resulted in certain South African researchers suggesting that POPIA may have to be adjusted to accommodate matters raised by them in relation to the halting of research projects or international research collaborations that might transgress the law.^{51,53,65} The three mentioned studies also elaborated on how POPIA deals with the matter of informed consent. In that regard, POPIA questions whether informed consent can be obtained from study participants and suggests that legal advice should be sought to avoid local laws being broken.

In 2020, before the POPI Act was due to be enforced health researcher members of the Academy of Science of South Africa (ASSAf) looked into the development of a Code of Conduct for such research that would assist with the application of the POPIA. The POPI Act does provide for such scientific bodies to make recommendations regarding the ethical conduction of research. Matters to be addressed by the Code of Conduct include which consent models would be permitted under POPIA; inherited characteristic processing issues that arise in relation to genetic research; the use of Innovations in Mobile Privacy and Security by researchers and how personal information sourced from social media platforms can be used for research purposes. Intellectual property rights including patents but not exclusively related to them and the commercial use of research data were also raised for consideration. The purpose of the Code is primarily to guide the use of personal scientific research information in a lawful and responsible way.⁸³ The Code of Conduct should enable a smoother transition into the conduction of research under the POPI Act.

Ethical and legal factors in the conduction of open science research impact not only individual participants but also the communities that they are affiliated with. It was reported that there would be reduced, or no benefits for the community from which the data came, after secondary analysis and the geographical detachment between the data source and the end user, thereby increasing their risk for potential harm.^{41,45,48,53} In the historical past of Africa, collection of research data was often associated

with exploitation of communities and harmful malpractices.⁴¹ Accordingly, it was suggested that research using data collected from a community should be carried out by the people in that community.⁴¹ One of the participants from a research study expressed the opinion that research was done with a community not on a community.^{41,48} Research data collection usually entails getting specific permissions from institutions and government departments, but it was suggested that community elders should also be asked for permission to conduct research in their community as it was important for researchers to respect cultural and religious norms.^{54,65} Inequality in research resources found between the global north and Africa was found to still lead to problems in data sharing between these geographical regions.³⁴ Therefore, it might be necessary to ask secondary data users to consider ensuring potential benefits to the original participants when requesting data.^{41,46,48,53}

Furthermore, even within the African region the widely differing data that can be collected across communities, countries, and cultures has led to the reasoning that standardization and harmonization should be applied to data definitions and terminology to ease data sharing.⁴⁴ It was suggested that secondary data users should be able to conceptually describe what they would use the data for as well.⁴¹ The issue of the interpretation and the understanding of the data by the secondary user was also raised in connection with secondary data being misused or misinterpreted.⁴⁶

The matter of exploitation of African communities for the collection and distribution of data that would be used for acquiring patents to benefit large companies or industries while the communities that the data were collected from wallowed in poverty was also raised.⁴¹ These same authors suggested that a benefit-sharing strategy should be created to help the community from where the data was being procured. In general, it was reported in three articles that more regulations should be devised to protect vulnerable communities.^{45,53,62} The POPIA in South Africa goes some way in implementing the sentiments expressed by these authors.

A social consideration for ethical collection of data both initially and in its secondary use is protection of participants anonymity. Sometimes in Africa data is collected from a small geographical area, perhaps a village or a small community. Research collecting sensitive data must be particularly carefully undertaken to avoid any chance of participants being identified or communities stigmatized.^{45,48} This type of data might include personal clinical information, genomics data or perhaps information relating to sexual abuse, exploitation, or orientation.⁴⁵

The monitoring and enforcement of compliance with rules and regulations pertaining to the ethical carrying out of research by local and foreign researchers are not always in place or enforced by African governments.³⁰ The Covid-19 pandemic highlighted this problem within the WHO African Region where noncompliance with data

sharing was observed in Tanzania and elsewhere in the region Covid-19 data flows were not adherent to IHR-2015 sharing regulations.⁴⁴ In the ECOWAS region, lack of funding for research or the contracting out of research programs has led to poor support for researchers and lack of capacity building.⁵⁶ Throughout Africa many countries are bureaucratic, and corruption is rife leading to research programs underfunded or not funded at all.⁵⁸ These authors also pointed out that in Nigeria the government makes promises regarding improving health research but often these promises never materialized leading to ill will against the state.

In the international sense, Seastedt et al.⁸⁴ put forward the idea that for global health care fairness to be realized more data should be shared not less data. However, they do not mention Africa in their assessment. They elaborate on how the increase in data sharing will require more governmental oversight and laws will have to be enacted such as the 116th Congress (2019–2020): National Artificial Intelligence Initiative Act of 2020 that was passed in the United States.⁸⁵

Limitations and recommendations

The natural limitations of the did not allow for an in-depth analysis of the studies included in this review. The study content was diverse, and the databases or repositories used to store and share data were numerous and not clearly defined in their usage parameters or restrictions that were in place to govern their use. The many different legal measures controlling the sharing of data across the African continent give rise to a confusing picture of how data sharing is controlled.

This scoping review mapped out the data sharing practices, perceptions and considerations among health researchers in Africa, and it identified the opportunity for a qualitative systematic review exploring the perceptions, barriers, and facilitators to data sharing among health researchers in Africa in greater depth to be conducted. A table showing the critical contributions of this scoping review in comparison with other reviews on health research data in Africa is in Appendix 3 of the supplementary files. Moreover, research studies that compare the legal and policy frameworks for data sharing across African countries are needed. This could help identify best practices and gaps in current regulations.

It is evident that African countries, especially those in the Northern and Central regions of Africa, need to invest in reliable and accessible technological solutions to support data storage, sharing, and analysis. Also, the importance of providing training and resources to health researchers to improve their skills and knowledge in data management and sharing cannot be overemphasized. There is also a need for harmonized legal frameworks that would facilitate data sharing while protecting privacy and intellectual property rights in Africa.

Conclusions

This review has identified that Africa is making progress toward realizing a network of health databases with the possibility of some standardization of data. There are also moves toward protecting data in repositories and biobanks, increasing protection of intellectual property, and protection of the ethical rights of those participating in research whether individuals or communities. However, many parts of Africa are still not participating in research due to poor technological infrastructure, lack of electricity, and very poor resources for the conduction of research brought about by poverty and governmental bureaucracy and corruption. Furthermore, the need for clarity regarding the timeous sharing of data during global health emergencies such as the COVID-19 pandemic has raised new concerns regarding the possible economic and scientific detrimental outcomes resulting from the disclosure of discoveries of new pathogens or variants of those already in existence. This matter needs to be addressed or it could impair the collection and dissemination of valuable data that might make future diagnoses and treatments of disease a reality and prevent such discoveries being shared with the world.

Acknowledgments: The Scoping review contributed toward the Hillel Friedland Foundation's Postdoctoral project on data sharing practices in the Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa.


Author contributorship: BO and OLO researched literature and conceived the study. OLO, DAS, and BO were involved in protocol development, data collection, and analysis. OLO wrote the first draft of the manuscript. All authors reviewed and edited the manuscript and approved the final version of the manuscript.

Declaration of conflicting interests: The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Ethical approval: Not applicable.

Funding: The authors received no financial support for the research, authorship, and/or publication of this article.

Guarantor: OLO, DAS, BO.

ORCID iD: Oluchukwu Loveth Obiora  <https://orcid.org/0000-0002-2372-3645>

Supplemental material: Supplemental material for this article is available online.

References

- Schwalbe N, Wahl B, Song J, et al. Data sharing and global public health: defining what we mean by data. *Front Digit Health* 2020; 2: 612339.
- Lang T. Advancing global health research through digital technology and sharing data. *Sci* 2011; 331: 714–717.
- Terry RF, Littler K and Olliaro PL. Sharing health research data – the role of funders in improving the impact. *F1000Res* 2018; 7: 1641.
- Wellcome Trust. Sharing research data to improve public health: full joint statement by funders of health research, <https://wellcome.org/what-we-do/our-work/sharing-research-data-improve-public-health-fulljoint-statement-funders-health#notcover> (2022, cited 27 December 2022).
- The Yale University Open Data Access (YODA) Project. Forging a unified scientific community, <https://yoda.yale.edu> (2022, cited 22 December 2022).
- Wilkinson MD, Dumontier M, Aalbersberg IJ, et al. The FAIR guiding principles for scientific data management and stewardship. *Sci Data* 2016; 3: 160018.
- 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 Human Res Ethics* 2015; 10: 302–313.
- Fernando B, King M and Sumathipala A. Advancing good governance in data sharing and biobanking - international aspects. *Wellcome Open Res* 2019; 4: 184.
- Kabanda SM, Cengiz N, Rajaratnam K, et al. Data sharing and data governance in sub-Saharan Africa: perspectives from researchers and scientists engaged in data-intensive research. *S Afr J Sci* 2023; 119: 1–12.
- Bezuidenhout L and Chakauyac E. Hidden concerns of sharing research data by low/middle income country scientists. *Global Bioethics* 2018; 29: 39–54.
- Fonkou MDM, Bragazzi ML, Tsinda EK, et al. COVID-19 Pandemic related research in Africa: bibliometric analysis of scholarly output, collaborations and scientific leadership. *Int J Environ Res Pub Health* 2021; 18: 7273.
- World Bank. <https://blogs.worldbank.org/en/opendata/world-bank-country-classifications-by-income-level-for-2024-2025> (2024, retrieved 10 October 2024).
- Worldometer. Worldometer-number of countries in Africa, <https://www.worldometers.info/geography/how-many-countries-in-africa/> (2022, cited 28 December 2022).
- Midega J, et al. *Countries should up support for research and development*. *University World News*, 2021. <https://www.universityworldnews.com/post.php?story=20210519073727854>.
- Jackson J, Viney C, de-Graft Aikins A, et al. Pathways to research leadership for early career researchers in Africa: a potential role for African and global funders. *S Afr J High Educ* 2022; 36: 38.
- National Institute of Health (NIH). Final NIH Policy 1474 for data management and sharing, <https://sharing.nih.gov/data-management-and-sharing-policy/about-data-management-and-sharing-policies/data-management-and-sharing-policy-overview> (2023, cited 25 May 2023).
- World Health Organisation (WHO). New WHO policy requires sharing of all research data, <https://www.who.int/news/item/16-09-2022-new-who-policy-requires-sharing-of-all-research-data> (2022, cited 28 December 2022).

18. Skelly L and Chiware ERT. African researchers do not think differently about open data. *Front Res Metr Anal* 2022; 7: 950212.
19. Academy of Science South Africa (ASSAF). African open science platform part 1: landscape study, <https://research.assaf.org.za/handle/20.500.11911/177> (2019, cited 2 January 2023).
20. Hulsen T. Sharing is caring-data sharing initiatives in health-care. *Int J Environ Res Pub Health* 2020; 17: 3046.
21. African Open Science Platform (AOSP). The future of Science and Science for the future, <https://aosp.org.za/about-us/> (2022, cited 2 January 2023).
22. Ross MW, Iguchi MY and Panicker S. Ethical aspects of data sharing and research participant protections. *Am Psych* 2018; 73: 138–145.
23. United Kingdom Government. Data ethics framework, <https://www.gov.uk/government/publications/data-ethics-framework/data-ethics-framework-2020> (2020, cited 3 January 2023).
24. European Union. General Data Protection Regulation (GDPR), <https://gdpr.eu/what-is-gdpr/> (2018, cited 3 January 2023).
25. Centers for Disease Control and Prevention (CDC). Health Insurance Portability and Accountability Act of 1996 (HIPAA), <https://www.cdc.gov/phlp/publications/topic/hipaa.html> (2023, cited 15 April 2023).
26. Brand D, Singh JA, Nienhaber McKay AG, et al. Data sharing governance in sub-Saharan Africa during public health emergencies: gaps and guidance. *S A J Sci* 2022; 118: 1–6.
27. Nordling L. South African Law may impede human health research. Data protection measure may block reuse of data in biobanks. *Science* 2019; 363: 802.
28. Munung NS, Mayosi BM and de Vries J. Equity in international health research collaborations in Africa: perceptions and expectations of African researchers. *PLoS ONE* 2017; 10: e0186237.
29. Obiora OL, Shead DA and Olivier B. Perceptions of human movement researchers and clinicians on the barriers and facilitators to health research data sharing in Africa. *Physiother Theory Prac* 2022; 40: 516–527.
30. 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. *Develop World Bioeth* 2018a; 18: 394–405.
31. Bangani S and Moyo M. Data sharing practices among researchers at South African Universities. *Data Sci J* 2019; 18(28): 1–14.
32. Obiora OL, Olivier B, Shead DA, et al. Data sharing practices of health researchers in Africa: a scoping review protocol. *JBIM Evid Syn* 2022; 20: 681–688.
33. Lo B. Sharing clinical trial data: maximizing benefits, minimizing risk. *JAMA* 2015; 313: 793–794.
34. Anane-Sarpong E, Wangmo T and Tanner M. Ethical principles for promoting health research data sharing with sub-Saharan Africa. *Dev World Bioeth* 2020; 20: 86–95.
35. Karczewski KJ, Tatonetti NP, Manrai AK, et al. Methods to ensure the reproducibility of biomedical research. *Pac Symp Biocomput* 2017; 22: 117–119.
36. Aromataris E and Munn Z. *JBIM manual for evidence synthesis (formerly JBI reviewer's manual)*. Adelaide: JBI, 2017.
37. McGowan J, Sampson M, Salzwedel DM, et al. PRESS peer review of electronic search strategies: 2015 guideline statement. *J Clin Epidemiol* 2016; 75: 40–46.
38. Covidence. Better systematic review management, <https://www.covidence.org> (2023, cited 10 April 2023).
39. Page MJ, McKenzie JE, Bossuyt PM, et al. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. *Br Med J* 2021; 372: 71.
40. Cole DC, Nyirenda LJ, Fazal N, et al. Implementing a national health research for development platform in a low-income country – a review of Malawi's health research capacity strengthening initiative. *Health Res Pol Syst* 2016; 14: 24.
41. 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 Ethic* 2015; 10: 290–301.
42. Dhali A, Mahomed S and Sanne I. Biobanks and human health research: balancing progress and protections. *S Afr J BL* 2015; 9: 55–59.
43. 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.
44. Impouma B, Mlanda T, Bukhari A, et al. Information management practices in the WHO African region to support response to the COVID-19 pandemic. *Epid Infec* 2021; 149: 5.
45. Jao I, Kombe F, Mwalokure 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.
46. Mbuthia D, Molyneux S, Njue M, et al. Kenyan health stakeholder views on individual consent, general notification and governance processes for the re-use of hospital inpatient data to support learning on healthcare systems. *BMC Med Ethic* 2019; 20: 40.
47. Memiah P, Mu TA, Penner J, et al. Bridging the gap in implementation science: evaluating a capacity-building program in data management, analysis, utilization, and dissemination in low- and middle-income countries. *Pop Health Man* 2018; 21: 1–7.
48. National Academies of Sciences, Engineering and Medicine. Chapter 1: Preface and acknowledgements. In: O'Connell ME and Plewes TJ (eds) *Sharing research data to improve public health in Africa: a workshop summary contributors*. Washington, DC: The National Academies Press, 2015, vii–xi.
49. Nature Medicine Editorial. Sharing data to save lives. *Nat Med* 2015; 21: 1235.
50. Nembaware V, Johnston K, Diallo AA, et al. A framework for tiered informed consent for health genomic research in Africa. *Nat Genet* 2019; 51: 1566–1571.
51. Nordling L. African academy leads push for ethical data use. *Nature* 2019a; 570: 284–285.
52. Staunton C and de Stadler E. Protection of personal information act No. 4 of 2013: implications for biobanks. *S Afr Med J* 2019; 109: 232–234.
53. Staunton C, Adams R, Dove ES, et al. Ethical and practical issues to consider in the governance of genomic and human research data and data sharing in South Africa: a meeting report. *AAS Open Res* 2019; 2: 15.
54. Anane-Sarpong E, Wangmo T, Sankoh O, et al. Application of ethical principles to research using public health data in the global south: perspectives from Africa. *Dev World Bioethic* 2018b; 18: 98–108.

55. Conton B. Build the Ebola database in Africa, <https://www.nature.com/articles/d41586-017-05676-4> (2017, cited 3 January 2023).
56. Aidam J and Sombié I. The West African health organization's experience in improving the health research environment in the ECOWAS region. *Health Res Pol Syst* 2016; 14: 30.
57. Lötter L and van Zyl C. A reflection on a data curation journey. *J Emp Res Hum Res Ethic* 2015; 10: 338–343.
58. Uzochukwu B, Mbachu C, Onwujekwe O, et al. Health policy and systems research and analysis in Nigeria: examining health policymakers' and researchers' capacity assets, needs and perspectives in south-east Nigeria. *Health Res Pol Syst* 2016; 14: 1–13.
59. Shaffer JG, Doumbia SO, Ndiaye D, et al. Development of a data collection and management system in West Africa: challenges and sustainability. *Infect Dis Pov* 2018; 7: 125.
60. Van Hoving DJ and Brysiewicz P. African emergency care providers' attitudes and practices towards research. *Afr J Em Med* 2017; 7: 9–14.
61. Townsend B. The lawful sharing of health research data in South Africa and beyond. *Inf Com Tech Law* 2022; 31: 17–34.
62. Nnamuchi O. H3Africa: an Africa exemplar? Exploring its framework on protecting human research participants. *Dev World Bioeth* 2018; 18: 156–164.
63. Moodley K, Cengiz N, Domingo A, et al. Ethics and governance challenges related to genomic data sharing in Southern Africa: the case of SARS1593 CoV-2. *Lancet Glob Health* 2022; 10: e1855–e1859.
64. Turcotte-Tremblay A-M and Mc Sween-Cadieux E. A reflection 1594 on the challenge of protecting confidentiality of participants while disseminating research results locally. *BMC Med Ethics* 2018; 19: 45.
65. Nordling L. South African Law may impede human health research. Data protection measure may block reuse of data in biobanks. *Science* 2019b; 363: 802.
66. Shabi IN, Shabit OM, Akawukereke MA, et al. Physicians utilisation of internet medical databases at the tertiary health institutions in Osun State, South West Nigeria. *Health Info Lib J* 2011; 28: 313–320.
67. Lwoga ET and Questier F. Open access behaviours and perceptions of health sciences faculty and roles of information professionals. *Health Info Lib J* 2015; 32: 37.
68. Ruggeri K. Disseminating health research in sub-Saharan Africa through journal partnerships, www.thelancet.com/lancetgh (2022, cited 01 November 2022).
69. Mijumbi-Deve R, Parkhurst J, Jones C, et al. Beyond the metrics of health research performance in African countries. *BMJ Global Health* 2021; 6: e006019.
70. United Nations (UNESCO). *UNESCO science report: towards 2030*. Paris, France: UNESCO Publishing, 2015.
71. Wenham C, Wouters O, Jones C, et al. Measuring health science research and development in Africa: mapping the available data. *Health Res Pol Syst* 2021; 19: 1–13.
72. Conrad LY, Delahunty R and Ding W. The promise and the future of research data sharing. *Learn Publ* 2022; 35: 4–6.
73. Tsegaye T and Flowergay S. A system architecture for ensuring interoperability in a South African national electronic health record system. *SACJ* 2018; 33.
74. Stuart D, Baynes G, Hrynaszkiwicz I, et al. White paper. Practical challenges for researchers in data sharing, https://figshare.com/articles/journal_contribution/Whitepaper_Practical_challenges_for_researchers_in_data_sharing/5975011 (2018, cited 11 February 2023).
75. Omungo M. Africa Science 'Millionaires': survey spotlights top-funded researchers, <https://www.nature.com/articles/d41586-018-07418-6> (2018, cited 11 February 2023).
76. National Health Institutes (NIH). Grants and funding, <https://www.nih.gov> (2023, cited 14 February 2023).
77. UK Research Institute (UKRI). Find funding, <https://www.ukri.org> (2023, cited 14 February 2023).
78. Wellcome Trust. Grant funding, <https://wellcome.org/grant-funding> (2023, cited 14 February 2023).
79. National Research Fund (NRF). Funding, <https://www.nrf.ac.za> (2023, cited 14 February 2023).
80. Manti S and Licari A. How to obtain informed consent for research. *Breathe* 2018; 14: 145–152.
81. Centers for Disease and Control (CDC). <https://www.cdc.gov/phlp/publications/topic/hipaa.html> (2023, cited 14 February 2023).
82. General Data Protection Regulation (GDPR). <https://gdpr.eu/what-is-gdpr/> (2023, cited 20 April 2023).
83. Adams R, Adeleke F, Anderson D, et al. POPIA code of conduct for research. *S Afr J Sci* 2021; 117: 1–12.
84. Seastedt KP, Schwab P, O'Brien Z, et al. Global healthcare fairness: we should be sharing more, not less, data. *PLoS Dig Health* 2022; 1: e0000102.
85. Johnson E. Text—H.R.6216 - 116th Congress (2019–2020): National Artificial Intelligence Initiative, <https://www.congress.gov/bill/116th-congress/house-bill/6216> (2020, cited 1 May 2023).