Health Research Policy
and Systems

RESEARCH

Open Access

Data protection legislation in Africa and pathways for enhancing compliance in big data health research

Nchangwi Syntia Munung[1*], Ciara Staunton[2,3], Otshepeng Mazibuko[1], P. J. Wall[4] and Ambroise Wonkam[1,5*]

Abstract

Background The increasing availability of large volumes of personal data from diverse sources such as electronic health records, research programmes, commercial genetic testing, national health surveys and wearable devices presents significant opportunities for advancing public health, disease surveillance, personalized medicine and scientific research and innovation. However, this potential is hampered by a lack of clarity related to the processing and sharing of personal health data, particularly across varying national regulatory frameworks. This often leaves researcher stakeholders uncertain about how to navigate issues around secondary data use, repurposing data for different research objectives and cross-border data sharing.

Method We analysed 37 data protection legislation across Africa to identify key principles and requirements for processing and sharing of personal health and genetic data in scientific research. On the basis of this analysis, we propose strategies that data science research initiatives in Africa can implement to ensure compliance with data protection laws while effectively reusing and sharing personal data for health research and scientific innovation.

Results In many African countries, health and genetic data are categorized as sensitive and subject to stricter protection. Key principles guiding the processing of personal data include confidentiality, non-discrimination, transparency, storage limitation, legitimacy, purpose specification, integrity, fairness, non-excessiveness, accountability and data minimality. The rights of data subjects include the right to be informed, the right of access, the right to rectification, the right to erasure/deletion of data, the right to restrict processing, the right to data portability and the right to seek compensation. Consent and adequacy assessments were the most common legal grounds for cross-border data transfers. However, considerable variation exists in legal requirements for data transfer across countries, potentially creating barriers to collaborative health research across Africa.

Conclusions We propose several strategies that data science research initiatives can adopt to align with data protection laws. These include developing a standardized module for safe data flows, using trusted data environments to minimize cross-border transfers, implementing dynamic consent mechanisms to comply with consent specificity and data subject rights and establishing codes of conduct to govern the secondary use of personal data for health research and innovation.

Keywords Data laws, Data governance, Big data, Data sharing, Dynamic consent, Safe data flows

*Correspondence:
Nchangwi Syntia Munung
nchangwisyntia@yahoo.com; munung.nchangwi@uct.ac.za
Ambroise Wonkam
awonkam1@jhmi.edu
Full list of author information is available at the end of the article

Introduction

The vast amount of health-related data generated today and potentially available for biomedical research is astounding. These data come from diverse sources, including individuals participating in health research, electronic health care records, third-party service providers such as medical insurance companies, telehealth platforms and direct-to-consumer genetic testing companies [1]. Digital platforms and devices, including wearables, mobile phone apps and social media, also contribute substantially to the research data ecosystem [2–4]. These large volumes of data from diverse sources, also known as big data [5], can be leveraged to accelerate scientific research and innovation, validate research findings, improve disease surveillance, uncover trends in population health that might not be apparent in individual datasets [6–8], advance personalized medicine and inform the development of evidence-based public health policies [9]. However, alongside these opportunities are significant ethical, legal and governance considerations for the processing of big data for health research. This includes, for example privacy concerns/breaches, algorithmic bias, the potential for discrimination, upholding the rights of data subjects, national sovereignty over genetics and health data, and compliance with national requirements on secondary analysis and cross-border transfer of health and genetic data [10–12].

To give effect to the right to privacy and the right to data protection, many African countries have enacted legislation on the protection of personal data [13, 14]. In parallel, regional bodies such as the African Union (AU), the Southern African Development Community (SADC) and the Economic Community of West African States (ECOWAS), have introduced model data protection laws aimed at informing the sharing of personal data among their member states [15]. Ensuring compliance with data protection standards is essential for safeguarding the rights of individuals. However, there is a lack of clarity on their application to biomedical and data-driven health research especially in relation to secondary data analysis, cross border sharing of data and use of data for purposes different from that of which they were initially collected for [16].

Generally, data protection legislation serves as a broad legal framework and are not sector specific, meaning in most instances it will lack detailed and/or specific guidance on health research. While many of these laws include some exceptions for processing special categories of personal data, such as for scientific research, they can sometimes conflict with national and international research ethics regulations within the same jurisdiction [17], making data sharing in international collaborative research particularly challenging [18–20]. For example, uncertainty about the application of data protection laws in scientific research, along with fears of sanctions and penalties, may cause African scientists to hesitate in sharing data with other researchers and third parties [16], thus limiting opportunities for collaboration. This is even more pronounced with the sharing of health and genetic data [21], which are often afforded extra protections status and classified as sensitive data, with the effect being that data sharing and reuse may become increasingly restricted thereby stifling global health research efforts.

To advance data-driven health research in compliance with national data protection statues, it is critical to reflect on strategies that data science health research initiatives in Africa can adopt to remain compliant while re-using and sharing personal data for the benefit of science, medicine and innovation. To highlight and address the additional requirements brought about by data protection laws, we analysed 37 data protection laws in Africa to identify key requirements related to health research. On the basis of the analysis, we propose strategies that data science health research initiatives in Africa can implement to ensure compliance with national data protection laws while effectively re-using and sharing personal data for health research and scientific innovation.

Methods

We conducted a comparative analysis of data protection legislation in 34 African countries and 3 regional African economic/geographic blocks (Table 1), with the goal of identifying core bioethical elements that speak to the regulation of health research, particularly concerning data collection, storage, cross border sharing and reuse. Key areas of focus included: principles guiding data use and reuse; the rights of data subjects; informed consent requirements; regulation on cross-border sharing of data; and responsibilities of various stakeholders involved in data collection, management and use.

Full text of the data protection laws were sourced through personal contacts, official government websites, the United Nations Conference on Trade and Development (https://unctad.org/page/data-protection-and-privacy-legislation-worldwide), databases and general internet searches via Google. The documents were imported into QSR-NVivo 12, a qualitative data analysis software to facilitate the systematic extraction and organization of information.

The data analysis focussed on specific provisions related to the following aspects: definitions of different types of data, specific requirements for scientific research, principles underpinning data protection, the responsibilities of data protection officers, the rights of data subjects and requirements for cross-border data transfer. A major limitation of the study is that the

Table 1 Overview of data protection legislation across Africa (grouped by language)

Country/regional blocks	Title of data protection legislation	Language
Egypt	Data Protection Law	Arabic and English
Chad	Loi portant protection des données à caractère personnel	Arabic and French
Tunisia	Loi portant sur la protection des données à caractère personnel	
Eswatini	Data Protection Act	English
The Gambia	Draft Data Protection and Privacy Policy and Strategy	
Ghana	Data Protection Act	
Kenya	The Data Protection Act	
Lesotho	Data Protection Act	
Malawi	Electronic Transactions and Cyber Security Act	
Mauritius	The Data Protection Act	
Seychelles	Data Protection Act	
South Africa	Protection of Personal Information Act	
Uganda	Data Protection and Privacy Act	
Zambia	Data Protection Act	
Zimbabwe	Data Protection Act	
Benin	Code du numérique en République du Bénin	French
Burkina Faso	Loi portant protection des données à caractère personnel	
Côte d'Ivoire	Loi relative à la lutte contre la cyber crilminalité	
Democratic Republic of Congo	Loi relative aux télécommunications et aux technologies de l'information et de la communication	
Gabon	Loi relative à la protection des données à caractère personnel	
Madagascar	Loi sur la protection des données à caractère personnel	
Mali	Loi portant protection des données à caractère personnel en République du Mali	
Mauritania	Loi sur la protection des données à caractère personnel	
Morrocco	Loi relative à la protection des personnes physiques à l'égard du traitement des données à caractère personnel	
Niger	Loi relative a la protection des données à caractère personnel	
Republic of Congo	Loi portant protection des données à caractère personnel	
Republic of Guinea	Loi relative à la lutte contre la cybercriminalité et la données à caractère personnel	
Senegal	Loi sur la protection des données à caractère personnel	
Rwanda	Law relating to the protection of personal data and privacy	Kinyarwanda, English and French
Tanzania	Muswada Wa Sheria Ya Ulinzi Wa Taarifa Binafsi Wa Mwaka	Kiswahili
Angola	Ante-Projecto de Lei da Protecção de Dados Pessoais	Portuguese
Carbo Verde	Lei de Proteção de Dados Pessoais	
São Tomé and Principe	Lei Protecção de Dados Pessoais	
Equatorial Guinea	Ley de Protección de Datos Personales	Spanish
Regional African Blocks		
The African Union	Convention on Cyber Security and Personal Data Protection (Malabo convention)	English
Economic Community of West African States (ECOWAS)	Personal Data Protection within ECOWAS	
Southern Africa Development Council (SADC)	SADC Data Protection Act (Model Law)	

language competency within our team restricted us to detailed analysis of legislation available in English and French. For laws written in other languages such as Kiswahili, Spanish and Portuguese, only basic information such as the name of the country, year and the title of the law was extracted.

Results

The complete text of 36 data protection statues and bills from across Africa were identified from the search (Fig. 1). This comprised 29 national data protection statues, one data protection and privacy bill, three cyber security acts, two model data protection laws from

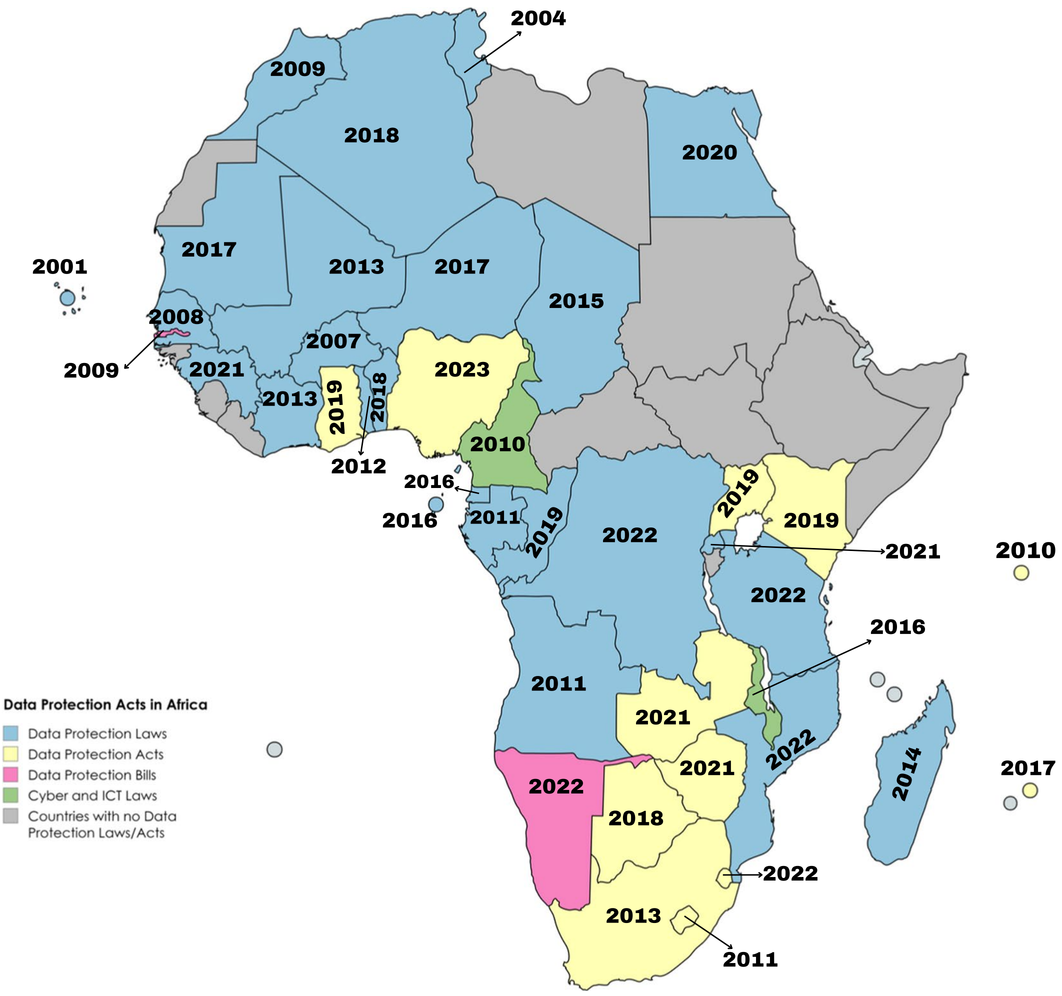

Fig. 1 Representation of African countries with data protection legislation/statutes and year enacted or drafted

African regional economic blocs and the African Union Convention on Cyber Security and Personal Data Protection. Out of these 36 documents, 31 were subjected to analysis, as they were available in either English or French, the two languages in which at least one member of the study team was proficient. The remaining documents were in Kiswahili, Portuguese or Spanish (Table 1). More than 50% of African countries have data protection legislation (Fig. 1).

Key concepts and definitions in data laws

Data protection laws defined different categories of data (Table 2) pertinent to health research, including sensitive data and biometric data. Health and genetic data fall

Table 2 Different categories and common definitions of data types

Data categories	General definitions from data protection legislations	Examples
Personal data	Any data relating to an identified natural person (data subject), or those identifiable, directly or indirectly, by reference to such data and to other	Name, voice, photograph, ID number, nationality, age, marital status, medical records, genetic data, race, ethnicity
Biometric data	Personal data resulting from specific technical processing based on physical, physiological or behavioural characterization	Blood typing, fingerprinting, DNA, earlobe geometry, retinal scanning, voice recognition
Sensitive personal data	Personal data which by its nature may be used to suppress the data subject's fundamental rights and freedoms	Genetic data, clinical records, biometric data, race, ethnic origins

within the category of sensitive data, warranting heightened levels of protection.

Processing of personal data for scientific research

The principles for the processing of personal data must be met for scientific research. In most instances, data protection laws typically accord exemptions or make special provisions on the processing of personal data for health or scientific research (Table 3). Tunisia, for example, introduces a specific provision for consent when processing data originally collected for a different purpose and subsequently needed for historical or scientific research. In such scenarios, data controllers are required to obtain the consent of the individuals involved or, in case of unavailability, their heirs or legal guardians. In Gabon, processing of personal data for research

Table 3 Country-specific provisions for the processing of sensitive data for scientific research

Country	Specific provisions for processing personal data for scientific research or processing for genetic data
Benin (Article 396)	Processing of personal data for scientific research is generally prohibited, unless specific conditions are met, for example, the research cannot reasonably be carried out without access to identifiable personal data, the information will not be used to contact individuals to participate in research and approval of the data controller
	The further processing of personal data personnel for historical, statistical or scientists carried out using anonymous data is admitted
	The processing of personal data for scientific research must adhere to the regulations and ethical standards governing the profession
Botswana (Section 24)	Processing must be compatible with specified, explicitly stated and for legitimate purpose
	Processing of the data must be approved by the data commissioner on the advice of a research and scientific ethics committee
Gabon (Chapter 5)	Processing of personal data for scientific research must be approved by the data commissioner on the advice of a research ethics committee and/or scientific committee composed of people competent in research in health, epidemiology, genetics and biostatistics
Niger (Article 7)	Authorization from the data protection authority (HAPDP) prior to processing genetic and health data for scientific research The processing of personal data for scientific research must adhere to the regulations and ethical standards governing the profession
Guinea	Processing of personal data (including genetic and medical data) for scientific research requires authorization from the data protection commission
Senegal (Chapter IV)	Processing of personal data (including genetic and medical data) for scientific research requires authorization from the *Commission des Données Personnelles*
	Request for data processing should include a research protocol specifying the objective of the research, the researchers involved, data analysis methods, the origin and nature of the personal data, justification and duration of use Scientific ad ethics review reports from relevant committees
	Where appropriate, scientific and technical justification for waiver of the requirement of access to anonymized data only and/or for storage of data beyond the required period
Tunisia (Section III)	The consent of the data subject or their heir is required for repurposing data for scientific research
	Doctors may communicate health data to persons/institutions for purposed of health research following a request and the authorization of the National Authority for the Protection of Personal Data (INPDP)
	The INPDP may, when issuing the authorization health data for research, set measures to ensure the protection of health data
	Personal data collected for scientific research may only be processed or used for scientific research
Zambia (Parts VII and IX)	Processing of sensitive data for scientific research by a person other than a public body will require the authorization of the Data Protection Commissioner
	When processing personal data for scientific research by a person other than a public body, that person shall ensure that the personal data are anonymized
	Where sensitive data are processed for scientific research, informing the data subject may be postponed if it would significantly prejudice the research; there is no evident risk of infringement of the rights of the data; and the data were collected initially on the basis of consent
ECOWAS (Article 12)	The processing of personal data relating to genetic data and health research is subject to authorization by the data protection authority
SADC (Part VII)	The data protection authority shall establish appropriate safeguards for personal data retained longer than permitted scientific research purposes
	Where sensitive data are processed for scientific research purposes, and there is no discernible risk of violating the data subject's rights, notification to the data subject may be deferred until the conclusion of the research, provided that informing the data subject earlier would significantly compromise the research. Under these circumstances, the data subject must have previously provided written consent for the processing of their personal data for scientific research, including agreeing to postpone notification until the appropriate time

requires an opinion from a research ethics committee. In the ECOWAS region, the use of health and genetic data for research purposes mandates permission from a data protection authority. Meanwhile, within the SADC region, the model data protection law stipulates that in cases where sensitive personal data are processed for scientific research and there is no apparent risk of privacy infringement or decision-making based on individual data, notification to the data subject may be postponed until the conclusion of the research. However, this delay is permissible only if informing the data subject would significantly prejudice the research. In such instances, the data subject must have previously provided written consent to the processing of their personal data for scientific research purposes, including postponement of notification for this reason.

Principles guiding the processing of personal information

All the data protection laws are built upon a set of principles that govern the lawful collection, storage and use of data (Table 4). The processing of personal data for scientific research must follow these principles. There are, however, in most regulations, certain exceptions to some of these principles if the processing is for research.

The rights of data subjects

All data protection regulations afford certain rights to data subjects (Table 5) including the prerogative to request organizations or data controllers to delete their personal data or opt out from the processing of their personal data, provided such objections are grounded in legitimate and justifiable reasons.

Cross border sharing: storage and sharing of scientific data

All countries that have data protection regulations in place do not permit the trans-border sharing of data unless the transfer falls within one of the grounds for the trans-border sharing of data specified in the regulation. The exact grounds vary according to jurisdiction and the precise definition of the ground differs, but they generally include some or a combination of the following:

- Sharing of data with a country that has an adequate level of protection (adequacy);
- Standard contractual clauses that provide a similar level of protection;
- Binding corporate agreements that provide a similar level of protection;
- The transfer is necessary for the performance of a contract between the data subject and the controller or measures prior to the conclusion of such a contract;
- Data subject consents to the transfer;
- The transfer is necessary to safeguard the vital interests of the data subject;
- The transfer is necessary or made legally binding for the protection of an important public interest, or for the establishment, exercise or defence of legal claims.

In the research context, the transfer mechanisms that are likely most appropriate are: adequacy, standard contractual clauses, binding corporate agreements or consent. As can be seen in Table 6, Madagascar, Mali and South Africa are the only countries surveyed that explicitly state binding corporate rules as a ground for transfer if the binding corporate rules would provide an adequate level of protection. Madagascar, Mali, South Africa and Zambia explicitly provide for standard contractual clauses as a ground for transfer. Thus, in the context of international collaborative research within Africa, adequacy and consent are most likely the grounds to be used in the transborder sharing of data. With the exception of Togo, Mali, Egypt and the Republic of Congo, consent is a ground under which personal data can be shared across borders. The consent would need to be specific to the transfer and specifically state the country that it is going to.

Responsibilities of individuals under data protection law

The data protection laws outline the roles and obligations of key data protection stakeholders (Table 7). For the purposes of scientific research, the data protection laws in Gabon, Senegal and Lesotho mention an advisory or scientific committee as a critical stakeholder for the processing of personal data for scientific research. By contrast, Botswana, Mauritania, Zimbabwe and SADC data protection laws stipulate that health-related data may only be processed under the responsibility of a healthcare professional.

Navigating data protection laws: proposed strategies for ensuring compliance in big data health research initiatives

Data protection laws introduce strict requirements on the processing and sharing of personal data. For instance, while informed consent stands as an ethical imperative in all research endeavours, under data protection regulations, it constitutes merely one potential lawful basis for processing personal data, subject to specific conditions and exceptions [22]. Consent may also be the lawful basis on which to transfer data internationally, or under adequacy, if the receiving country has an adequate level of protection [23]. Data science research initiatives in Africa need to develop mechanisms for navigating the complexities of processing personal data for health research. On the basis of our analysis, we recommend several

Table 4 Key principles outlined in data laws and their definitions/descriptions

Principles	Descriptions of principles	Examples of countries that list the principles
Confidentiality	Personal data must be processed securely to retain confidentiality and integrity in consistency, accuracy and trustworthiness over its entire life cycle	Benin, Chad, DRC, Gabon, Guinea, Mauritania, Nigeria, Morocco, Senegal, Zambia ECOWAS
Non-discrimination	Personal data should not be used to the disadvantage of individuals or groups	Benin, Tunisia
Transparency	Data subjects should be made aware of how their data are being processed, the recipient of the data, and the purpose of processing Data controller and processors should use clear, simple and plain language to communicate with the data subject Data subject should have easy access to the information about processing of their data	Benin, Congo, Gabon, Gambia, Guinea, Mauritius, Morocco, Niger, Nigeria, Rwanda, Senegal, Uganda, Zambia, ECOWAS
	Data processor should keep a register (available to the public) of all automatic processing operations of data	Zimbabwe
Storage limitation	Personal data should be kept in a form which identifies the data subject for no longer than is necessary	Algeria, Benin, Chad, DRC, Eswatini, Egypt, Eswatini, Gabon, Gambia, Ghana, Guinea, Ivory Coast, Kenya, Lesotho, Malawi, Mali, Mauritius, Morocco, Madagascar, Niger, Nigeria, Republic of Congo, Rwanda, Senegal, Seychelles, South Africa Uganda, Zambia, Zimbabwe,
Legitimacy	Data must be processed for explicit, specified and legitimate purposes Data processing should not violate public moral norms	Algeria, Benin, Burkina Faso, Chad, Egypt, Gabon, Gambia, Guinea, Senegal, Seychelles, Togo, Tunisia, Uganda
Purpose specification	Data must be collected for a specific purpose and cannot be processed later in a manner incompatible with the purpose for which it is collected	Algeria, Burkina Faso, Chad, Egypt, Gabon, Gambia, Ghana, Guinea, Madagascar, Mauritania, Morocco, Niger, Nigeria, Rwanda, Senegal, Seychelles, Togo, Tunisia, Uganda, Zambia
	Data obtained for one purpose should not be used for another purpose without the consent of the data subject	
Integrity	Collection, processing, storage and recording of data should be lawful, fair and non-fraudulent	Botswana, Kenya, Gambia, Madagascar, Mali, Mauritius, Niger, Togo, Tunisia, Zambia
Fairness	Give data subjects the highest degree of autonomy over control of their data Personal data should be obtained directly from the data subject	Botswana, Burkina Faso, Kenya, Gambia, Madagascar, Mali, Mauritania, Niger, Seychelles
Non-excessiveness and data minimality	Collect and process only for a specific purpose Collect only data that are necessary for the defined purpose Limit the possibility of repurposing personal data	Algeria, Burkina Faso, Chad, Egypt, Gambia, Mauritania, Mauritius, Morocco, Seychelles, South Africa, Uganda, Zambia
Openness	A data user must make personal data policies and practices known to the public	Ghana, Zimbabwe
Accountability	Provide sufficient guarantees of compliance with security measures as defined by the law Have internal mechanisms for demonstrating compliance to data laws (both to the data subject and to the data commission)	Benin, Gambia, Ghana, Kenya, Mauritania, Mauritius, Morocco, Nigeria, South Africa, Uganda, Zimbabwe

Table 5 Rights of data subjects as defined in national data protection laws

Rights of data subjects	Description of rights	Examples of countries that have these rights
Right to be informed	Data subjects have the right to be informed about how their personal information will be used	Algeria, Benin, Burkina Faso, Chad, Egypt, Eswatini, Kenya, Gabon Ghana, Guinea, Lesotho, Mali, Mauritania, Mauritius, Morocco, Niger, Nigeria, Republic of Congo, Rwanda, Senegal, Seychelles, South Africa, Tunisia, Uganda, Zambia, Zimbabwe
Right of access	Data subjects can request a copy of their personal data held by data controllers, processors or third parties	Algeria, Benin, Burkina Faso, Chad, Egypt, Eswatini, Kenya, Gabon, Ghana, Gambia, Guinea, Lesotho, Madagascar, Mali, Mauritania, Mauritius, Morocco, Niger, Nigeria, Republic of Congo, Rwanda, Senegal, Seychelles, South Africa, Tunisia, Uganda, Zambia, Zimbabwe
Right to rectification	Data subjects have the right to have inaccurate personal data corrected and incomplete data completed	Algeria, Benin, Burkina Faso, Chad, Gabon, Egypt, Eswatini, Ghana, Guinea, Kenya, Lesotho, Madagascar, Mali, Mauritania, Mauritius, Morocco, Niger, Nigeria, Republic of Congo, Rwanda, Senegal, Seychelles, South Africa, Tunisia, Uganda, Zambia, Zimbabwe
Right to erasure/deletion of data	Data subjects can request the deletion or removal of their data	Benin, Chad, Ghana, Guinea, Egypt, Eswatini, Kenya, Lesotho, Madagascar, Mali, Mauritania, Mauritius, Morocco, Niger, Nigeria, Republic of Congo, Rwanda, Senegal, Seychelles, South Africa, Tunisia, Uganda Zambia, Zimbabwe
Right to restrict processing	Data subjects have the right to limit data processing by the data controller for a specific duration	Algeria, Benin, Burkina Faso, Chad, Kenya, Egypt, Gabon, Gambia, Ghana, Guinea, Lesotho, Madagascar, Mauritania, Mauritius, Morocco, Niger, Nigeria, Republic of Congo, Rwanda, Senegal, South Africa, Tunisia, Uganda, Zambia, Zimbabwe
Right to data portability	Data subjects can request their personal data to be transferred to another data controller	Benin, Kenya, Mali, Niger, Nigeria, Rwanda, Zambia, Zimbabwe
Right to seek compensation	If data subjects suffer serious damage or loss due to violations of data protection laws, they may seek compensation	Ghana, Rwanda, Seychelles

Table 6 Relevant grounds for the transborder transfer of personal data

	Adequacy (supervisory authority)	Adequacy (data controller)	Binding corporate rules	Standard contractual clauses	Consent
Algeria	X				X
Benin	X				X
Burkina Faso		X			
Egypt		X			
Eswatini		X			X
Gabon	X				
Gambia		X			X
Guinea	X				
Kenya		X			X
Madagascar	X		X	X	X
Mali	X		X	X	
Mauritius		X			X
Nigeria	X				X
Republic of Congo	X				
South Africa		X	X	X	X
Togo	X				
Uganda				X	X
Zambia	X			X	X
Zimbabwe	X				X

approaches to address the complexities of re-use and cross-border sharing of personal data for health research while ensuring compliance with data laws. This includes the use of trusted research environments, establishing a module for safe data flows in Africa, adopting dynamic consent, developing codes of conduct to complement data laws and engaging the public on big data for health research.

Establishing a module for safe data flow for health research in Africa

For scientific research, the grounds for what can be shared between jurisdictions is based on one of the following: adequacy, standard contractual clauses, binding corporate agreements or consent (Table 6). These mechanisms ensure that health and genomic data can flow securely across borders while adhering to the diverse national and regional legal standards that protect personal data. To meet these demands it is necessary for African data science research consortia to establish a safe data module that provides a structured framework for lawful and ethical management and transfer of personal data for health research and public health purposes. The module should focus on informed consent, adequacy assessments, exploring alternative grounds for data transfers, training in data protection principles and processes and monitoring and compliance. Drawing on the analysis of data protection legislation in African countries and our experience in data-driven health research, we propose a set of practical recommendations for creating a robust, compliant and effective module for safe data flow (Table 8).

Adopting technical approaches to data analysis that limit cross border data transfer

The implementation of trusted research environments (TREs), designed to offer remote and pre-approved access to health data [24], may prove necessary, perhaps indispensable, within the current data protection landscape in Africa. TREs effectively restrict researchers from directly copying individual-level data while allowing other researchers to access and analyse data using techniques such as federated data sharing [25] and data visiting [26]. However, the implementation of these techniques in Africa would require the development of harmonized codes of conduct for data access, significant investment in data infrastructure, trained workforce in cloud computing and use within TREs. To ensure compliance to data protection laws, it would be essential to anchor the codes of conduct on principles outlined in data protection laws (Table 4), as well as those identified as key to fostering equity in research partnerships in Africa [27, 28]. Initiatives in the United Kingdom have also proposed the five safes framework as a code of conduct that is central to the use of TREs [29], and its application to big-data-driven research in the United

Table 7 Responsibilities of different stakeholders as listed in different data laws

Stakeholder	Description of stakeholder	Examples of responsibilities/rights of the stakeholder	Examples of stakeholders in biomedical research
Data subject	An individual who is the subject of personal data	Rights are set out in Table 4	Research participants Individuals enrolled into a health registry Biobank donors
Data controllers	Individual or public/private company, agency or association which, alone or jointly with others, takes the decision to collect and process personal data and determines the purposes thereof	Comply with data protection requirements Provide the data commissioner with proof of assessment of appropriate safeguards for cross border transfer	Universities Research institutions Research groups Data coordinating centres Biorepositories Pharmaceutical companies
Data processors	An individual, private entity, public authority or agency or any other body who or which processes personal data on behalf of, or at the direction, of a data controller	Process personal data as specified by the data controller Ensure adequate level of security of the personal data	Consultant researchers Consultant data analyst
Data commissions or data protection authorities	Independent administrative and public authority responsible for ensuring compliance to data protection laws	Promote educational activities relating to protection of personal data Monitor and adopt authorizations for transborder flow of data and facilitate international cooperation Maintain a register of data controllers, if applicable Prepare and disseminate code of practice for data controllers, if applicable Establish a list of countries with similar or higher level of protection of personal data Authorize cross border transfer, if applicable	Nigeria Data Protection Commission The Office of the Data Protection Commissioner The National Commission for the supervision of Personal Data Protection
Research and scientific ethics committees	A board/committee that provides scientific and/or ethical review and opinion on a research project	Advise the data commissioner when there is a request to process sensitive data for ethics committee, where applicable	Research ethics committee Scientific advisory council
Health professionals	A health professional registered with a recognized health professional council	Responsible for the processing of personal health-related data	Medical doctors Nurses

Table 8 Recommendations for a safe data module for data sharing

Component	Description	Examples of Implementation in data-driven health research consortia in Africa
Informed consent	Ensure that data transfer across borders is based on explicit consent from data subjects, specific to each transfer and names the destination country	Funders: Support the development of standardized data sharing consent forms tailored to national research ethics guidelines of participating countries Data coordinating centres: Map consent form details to the requirements of participating countries and implement a digital consent management system Ethics committees: Ensure that consent for international collaborative projects meets both ethical and national legal requirements
Adequacy assessment framework	Evaluate whether the receiving country provides adequate protection for personal data through pre-approved lists, assessments by regulatory bodies or self-assessments by responsible parties/data controllers	Researchers: Verify adequacy through local data protection authorities or pre-approved lists Research consortium: Establish and maintain an updated list of countries with adequate protection; create tools and checklists for adequacy assessments Funders: Encourage the use of recognized adequacy assessments for international collaborations Policymakers and researchers: Collaborate with African regulatory bodies and regional bodies (AU, SADC, ECOWAS) to develop and harmonize a pre-approved list of adequacy requirements
Alternative grounds for data transfer	When adequacy is not achievable, explore alternative grounds for data transfer, such as explicit consent	Research consortium: Establish guidelines for determining and documenting alternative grounds for data transfer Funders: Support initiatives that provide templates, tools and resources for assessing and documenting data sharing on the basis of alternative grounds
Training in data protection legislation and processes	Provide training for researchers and data managers and data protection officers o, adequacy assessments	Funders: Finance training programs and resources for researchers on personal data protection Researchers: Engage in workshops and certification courses on data protection Research institutions: Support comprehensive training initiatives on data protection
Monitoring and compliance	Regularly review and update the data transfer module to adapt to changes in data protection laws and research needs	Researchers: Implement regular reviews to monitor compliance to data protection legislation Funders: Provide funding for monitoring tools and compliance audits

Kingdom has proven to very beneficial [30–32]. The five safes framework (safe projects, safe people, safe data, safe settings, safe outputs) could serve as a valuable tool for thinking through codes of conduct for data access and use in TREs in Africa. However, empirical studies on the feasibility and preferences of TREs and remote data access and analysis methods (e.g. data visiting, data federation) by scientists in Africa would be required to inform their rapid adoption and use in big data health research in Africa.

Dynamic consent: a solution to consent specificity and rights to restrict processing

Data protection laws place emphasis on the specificity of consent for the processing of personal data or the transborder flow of data. Where consent is not the lawful basis for the processing of personal data, data subjects have certain rights, which can include the right to object to the processing of their personal data (Table 4). Tunisia, for example, introduces a specific provision for consent when processing data originally collected for a different purpose and subsequently needed for historical or scientific research. In such scenarios, data controllers are required to obtain the consent of the individuals involved, or in case of unavailability, their heirs or legal guardians. In such cases, dynamic consent [33] offers a promising digital solution for managing the complexities of consent specificity and data subjects' rights.

Dynamic consent employs digital platforms to foster continuous communication and engagement between data custodians and research participants [33, 34] by providing updates on data use and research progress, aligning with principles of autonomy, legitimacy, purpose limitation and fairness. Another significant benefit of dynamic consent is that it empowers research participants to exercise their rights as prescribed in data protection laws, such as the right to object to the processing of personal data. Furthermore, emerging data suggest that research participants would like to be re-contacted for future use of their data and samples for health research [35]. This further strengthens the argument for dynamic consent, as it provides a flexible and participant-centred approach to managing consent over time. A couple of initiatives have already proposed dynamic consent platforms tailored for use in big data health research [36–38]. However, the feasibility and acceptability of dynamic consent in Africa would need to be explored.

Data governance: approaching data privacy through a socio-cultural lens

The data protection legislation in all the countries is heavily informed by the rights of natural persons to data privacy. However, the effectiveness and adequacy of data protection laws as it applies to health research in Africa would be contingent upon socio-cultural factors that shape perceptions of privacy, trust and data sharing practices in health research. Generally, culture exerts a profound influence on people's perceptions of privacy, data protection and willingness to share personal information [39, 40]. In communal cultures, prevalent across Africa, where solidarity is prioritized, there may be a greater willingness to share personal information for the greater good of the community [41]. Empirical studies conducted across Africa have shown that research participants often express a willingness to share their data for research purposes, particularly if it is to be used for, the public good [42, 43]. Additionally, data from some of our public engagement activities on genomics and big data research data leans towards support for the concept of data solidarity, with participants stating that they will favour minimal restrictions to data sharing if benefits accrue to their communities and they were informed of how their data are contributing to the public good. However, it would be essential to further explore whether communities view data sharing for research purposes as encroaching upon their privacy and autonomy and if that requires stringent rules for data sharing within and across borders. Such insights can inform the development of codes of conduct or harmonized data protection frameworks for research, focussing on the benefits and risks associated with different data uses, rather than solely emphasizing stringent rules around personal data.

Public engagement and education on data laws in health research

Public engagement activities aimed at raising awareness about data protection laws can empower individuals to make informed decisions about their privacy rights and secondary uses of their data for research and innovation. It should involve educating the public about the transformative potential of data-driven scientific advancements and empowering the public to appreciate the possibilities that that the use of their personal data can bring to advances in health research and medicine. Equally important is addressing the ethical and social concerns that may arise when sensitive data are repurposed and used for secondary research or commercial purposes.

Conclusions

While data protection laws are not primarily designed for scientific research purposes, they will significantly influence the way African researchers approach data sharing. Through a comparative analysis of data laws across Africa, we propose that to harness the full potential of big data for health research and innovation while adhering to data protection legislation, initiatives in data science for health should consider adopting the following strategies: (1) the use of data access and analysis methods that allow for data localization; (2) Implementation of dynamic consent to meet requirements of specificity of consent; (3) public engagement and education on sharing of personal data for health research as prescribed in data protection laws; and (4) development of codes of conduct for the responsible sharing, reuse and repurposing of personal data for scientific research and innovation. The development of codes of conducts should take into consideration societal perceptions of privacy. Finally, the formulation of the recommended guidance, policies and codes of conduct would greatly benefit from input and support from African regional and international agencies such as the African Union Development Agency-New Partnership for Africa's Development (AUDA-NEPAD), the Africa Centres for Disease Control and Prevention, the WHO and the World Economic Forum, that have a mandate to promote science policy and diplomacy in Africa and/or have a vested interest in fostering the responsible use of big data for global health research.

Author contributions

N.S.M. and A.W. conceptualized and designed the study. Primary data extraction was done by N.S.M. Secondary checks for data extraction was done by O.M. Analysis and interpretation were carried out by N.S.M., C.S, P.J. and A.W. The first draft was written by N.S.M. and C.S. All authors contributed to the revision of the manuscript and approved the final version for publication.

Funding

This study is funded by the NIH-National Institutes of Mental Health (award no. U01MH127692) for the project "Public Understanding of Big data in Genomics Medicine in Africa (PUBGEM-Africa)" under the Harnessing Data Science for Health Discovery and Innovation in Africa (DS-I Africa) program". A.W. and N.S.M. receive research funding from the NHLBI (U24HL135600) as part of the Sickle Africa Data Coordinating Center (SADaCC). C.S. is funded through the U.S. National Institute of Mental Health and the U.S. National Institutes of Health (award no. U01MH127690) as part of the DSI-Law project. The content of this article is solely the author's responsibility and does not necessarily represent the official views of the U.S. National Institutes of Health.

Data availability

No datasets were generated or analysed during the current study.

Declarations

Competing interests

The authors declare no competing interests.

Author details

[1]Division of Human Genetics, University of Cape Town, Cape Town, South Africa. [2]Institute for Biomedicine, Eurac Research, Bolzano, Italy. [3]School of Law, University of Kwazulu-Natal, Durban, South Africa. [4]ADAPT Centre Trinity College, Dublin, Ireland. [5]McKusick-Nathans Institute and Department of Genetic Medicine, John Hopkins University School of Medicine, Baltimore, MD, United States of America.

Received: 25 July 2024 Accepted: 20 September 2024
Published online: 15 October 2024

References

1. Alonso SG, de la Torre DI, Rodrigues JJPC, Hamrioui S, López-Coronado M. A systematic review of techniques and sources of big data in the healthcare sector. J Med Syst. 2017;41(11):183.
2. Hussain MS, Tewari D. Social media applications in biomedical research. Explor Digit Health Technol. 2024;2(4):167–82.
3. Munos B, Baker PC, Bot BM, Crouthamel M, de Vries G, Ferguson I, et al. Mobile health: the power of wearables, sensors, and apps to transform clinical trials. Ann N Y Acad Sci. 2016;1375(1):3–18.
4. Bietz MJ, Bloss CS, Calvert S, Godino JG, Gregory J, Claffey MP, et al. Opportunities and challenges in the use of personal health data for health research. J Am Med Inform Assoc. 2016;23(e1):e42–8.
5. Kitchin R, McArdle G. What makes big data, big data? Exploring the ontological characteristics of 26 datasets. Big Data Soc. 2016;3(1):2053951716631130.
6. Saberwal G. The many uses of data in public clinical trial registries. Curr Sci. 2021;120(11):1686–91.
7. Król ZJ, Dobosz P, Ślubowska A, Mroczek M. WGS data collections: how do genomic databases transform medicine? Int J Mol Sci. 2023;24(3):3031.
8. Bourgeron T, Chapeau UK. Biobank! A revolution for integrated research on humans and large-scale data sharing. CR Biol. 2022;345(1):7–10.
9. Vickers AJ. Whose data set is it anyway? Sharing raw data from randomized trials. Trials. 2006;7:15.
10. Lalova-Spinks T, De Sutter E, Valcke P, Kindt E, Lejeune S, Negrouk A, et al. Challenges related to data protection in clinical research before and during the COVID-19 pandemic: an exploratory study. Front Med. 2022. https://doi.org/10.3389/fmed.2022.995689.
11. Staunton C, Adams R, Anderson D, Croxton T, Kamuya D, Munene M, et al. Protection of personal information act 2013 and data protection for health research in South Africa. Int Data Privacy Law. 2020;10(2):160–79.
12. Boscarino N, Cartwright RA, Fox K, Tsosie KS. Federated learning and Indigenous genomic data sovereignty. Nat Mach Intell. 2022;4(11):909–11.
13. Makulilo AB. African data privacy laws. Berlin: Springer; 2016.
14. Daigle B. Data protection laws in Africa: A pan-African survey and noted trends. J Int'l Com & Econ. 2021:1.
15. Greenleaf G, Cottier B. International and regional commitments in African data privacy laws: a comparative analysis. Comput Law Secur Rev. 2022;44:105638.
16. Peloquin D, DiMaio M, Bierer B, Barnes M. Disruptive and avoidable: GDPR challenges to secondary research uses of data. Eur J Hum Genet. 2020;28(6):697–705.
17. Gefenas E, Lekstutiene J, Lukaseviciene V, Hartlev M, Mourby M, Cathaoir KÓ. Controversies between regulations of research ethics and protection of personal data: informed consent at a cross-road. Med Health Care Philos. 2022;25(1):23–30.
18. Fleming N. Proposed EU data laws leave researchers out in the cold. Nature. 2023. https://doi.org/10.1038/d41586-023-01572-2.
19. Nordling L. A new law was supposed to protect South Africans' privacy. It may block important research instead. Science. 2019. https://doi.org/10.1126/science.aax0768.
20. Lewis D. China's souped-up data privacy laws deter researchers. Nature. 2023. https://doi.org/10.1038/d41586-023-01638-1.
21. Townsend B. The lawful sharing of health research data in South Africa and beyond. Inform Commun Technol Law. 2022;31(1):17–34.
22. Staunton C, Adams R, Horn L, Labuschaigne M. A framework to govern the use of health data for research in Africa: a South African perspective. In: Zima T, Weisstub DN, editors. Medical research ethics: challenges in the 21st century. Cham: Springer International Publishing; 2023. p. 485–99.
23. Scheibner J, Ienca M, Kechagia S, Troncoso-Pastoriza JR, Raisaro JL, Hubaux JP, et al. Data protection and ethics requirements for multisite research with health data: a comparative examination of legislative governance frameworks and the role of data protection technologies. J Law Biosci. 2020;7(1):010.
24. Varma S, Hubbard T, Seymour D, Brassington N, Madden S. Building trusted research environments-principles and best practices. London: Towards TRE ecosystems; 2021.
25. Kairouz P, McMahan HB, Avent B, Bellet A, Bennis M, Nitin Bhagoji A, et al. Advances and open problems in federated learning. Found Trends Mach Learn. 2021;14(1–2):1–210.

26. Weise M, Kovacevic F, Popper N, Rauber A. OSSDIP: open source secure data infrastructure and processes supporting data visiting. Data Sci J. 2022. https://doi.org/10.5334/dsj-2022-004.
27. Munung NS, de Vries J, Pratt B. Genomics governance: advancing justice, fairness and equity through the lens of the African communitarian ethic of Ubuntu. Med Health Care Philos. 2021;24(3):377–88.
28. Yakubu A, Tindana P, Matimba A, Littler K, Munung NS, Madden E, et al. Model framework for governance of genomic research and biobanking in Africa—a content description. AAS Open Res. 2018. https://doi.org/10.1268/aasopenres.12844.2.
29. Desai T, Ritchie F, Welpton R. Five safes: designing data access for research. Economics. 2016;1601:28.
30. Brophy R, Bellavia E, Bluemink MG, Evans K, Hashimi M, Macaulay Y, et al. Towards a standardised cross-sectoral data access agreement template for research: a core set of principles for data access within trusted research environments. Int J Popul Data Sci. 2023;8(4):2169.
31. Kerasidou CX, Malone M, Daly A, Tava F. Machine learning models, trusted research environments and UK health data: ensuring a safe and beneficial future for AI development in healthcare. J Med Ethics. 2023;49(12):838–43.
32. Ritchie F, Tilbrook A, Cole C, Jefferson E, Krueger S, Mansouri-Benssassi E, et al. Machine learning models in trusted research environments—understanding operational risks. Int J Popul Data Sci. 2023;8(1):2165.
33. Kaye J, Whitley EA, Lund D, Morrison M, Teare H, Melham K. Dynamic consent: a patient interface for twenty-first century research networks. Eur J Hum Genet. 2015;23(2):141–6.
34. Kaye J, Curren L, Anderson N, Edwards K, Fullerton SM, Kanellopoulou N, et al. From patients to partners: participant-centric initiatives in biomedical research. Nat Rev Genet. 2012;13(5):371–6.
35. Moodley K, Sibanda N, February K, Rossouw T. "It's my blood": ethical complexities in the use, storage and export of biological samples: perspectives from South African research participants. BMC Med Ethics. 2014;15(1):4.
36. Haas MA, Teare H, Prictor M, Ceregra G, Vidgen ME, Bunker D, et al. 'CTRL': an online, dynamic consent and participant engagement platform working towards solving the complexities of consent in genomic research. Eur J Hum Genet. 2021;29(4):687–98.
37. Miranda P, Kaur J, Morita PP. Designing dynamic informed consent for public health research. Eur J Public Health. 2023;33(2):60.
38. Wang Z, Stell A, Sinnott RO. A GDPR-compliant dynamic consent mobile application for the Australasian type-1 diabetes data network. Healthcare. 2023;11(4):496.
39. Reviglio U, Alunge R. "I am datafied because we are datafied": an ubuntu perspective on (relational) privacy. Philos Technol. 2020;33(4):595–612.
40. Basu S. Privacy protection: a tale of two cultures. Masaryk Univ J Law Technol. 2012;6(1):1–34.
41. Kim J, Kwan M-P. An examination of people's privacy concerns, perceptions of social benefits, and acceptance of COVID-19 mitigation measures that harness location information: a comparative study of the U.S. and South Korea. ISPRS Int J Geo-Inform. 2021;10(1):25.
42. Mohammed-Ali AI, Gebremeskel EI, Yenshu E, Nji T, Ntabe AC, Wanji S, et al. Informed consent in a tuberculosis genetic study in Cameroon: information overload, situational vulnerability and diagnostic misconception. Res Ethics. 2022. https://doi.org/10.1177/17470161221106674.
43. Nansumba H, Flaviano M, Patrick S, Isaac S, Wassenaar D. Health care users' acceptance of broad consent for storage of biological materials and associated data for research purposes in Uganda. Wellcome Open Res. 2022;7:73.

Publisher's Note