Title
Toward better governance of human genomic data

Journal
Nature Genetics, 53(1)

ISSN
1061-4036

Authors
O'Doherty, Kieran C
Shabani, Mahsa
Dove, Edward S
et al.

Publication Date
2021

DOI
10.1038/s41588-020-00742-6

Peer reviewed

HHS Public Access

Author manuscript

Nat Genet. Author manuscript; available in PMC 2021 September 19.

Published in final edited form as:

Nat Genet. 2021 January ; 53(1): 2–8. doi:10.1038/s41588-020-00742-6.

Towards better governance of human genomic data

Kieran C. O'Doherty[1], **Mahsa Shabani**[2], **Edward S. Dove**[3,*], **Heidi Beate Bentzen**[4], **Pascal Borry**[5], **Michael M. Burgess**[6], **Don Chalmers**[7], **Jantina De Vries**[8], **Lisa Eckstein**[7], **Stephanie M. Fullerton**[9], **Eric Juengst**[10], **Kazuto Kato**[11], **Jane Kaye**[12], **Bartha Maria Knoppers**[13], **Barbara A. Koenig**[14], **Spero M. Manson**[15], **Kimberlyn M. McGrail**[6], **Amy L. McGuire**[16], **Eric M. Meslin**[17], **Dianne Nicol**[7], **Barbara Prainsack**[18], **Sharon F. Terry**[19], **Adrian Thorogood**[13], **Wylie Burke**[20]

[1]Department of Psychology, University of Guelph, Canada.

[2]Metamedica, Faulty of Law and Criminology, Ghent University, Belgium.

[3]School of Law, University of Edinburgh, United Kingdom.

[4]Center for Medical Ethics, Faculty of Medicine, University of Oslo, Norway.

[5]Department of Public Health and Primary Care, KU Leuven, Belgium.

[6]School of Population and Public Health, University of British Columbia, Canada.

[7]Faculty of Law, University of Tasmania, Australia.

[8]Department of Medicine, University of Cape Town, South Africa.

[9]Department of Bioethics and Humanities, University of Washington, United States of America.

[10]Department of Social Medicine, University of North Carolina, United States of America.

[11]Graduate School of Medicine, Osaka University, Japan.

[12]Faculty of Law, University of Oxford, United Kingdom.

[13]Centre of Genomics and Policy, McGill University, Canada.

[14]University of California, San Francisco, United States of America.

[15]Colorado School of Public Health, University of Colorado, United States of America.

[16]Center for Medical Ethics and Health Policy, Baylor College of Medicine, United States of America.

[17]Council of Canadian Academies, Canada.

[18]Department of Political Science, University of Vienna, Austria; and Department of Global Health & Social Medicine, King's College London, United Kingdom.

[19]Genetic Alliance, United States of America.

[*]Corresponding author (edward.dove@ed.ac.uk).

Competing Interests:

Edward S. Dove is Co-Chair of the Regulatory and Ethics Work Stream of the Global Alliance for Genomics and Health (GA4GH). Jantina de Vries is a member of the H3Africa Steering Committee. The remaining authors declare no competing interests.

[20]Department of Bioethics and Humanities, University of Washington, United States of America.

Abstract

In this Commentary, we argue that in line with the dramatic increase in the collection, storage, and curation of human genomic data for biomedical research, genomic data repositories and consortia have adopted governance frameworks to address the dual objectives of enabling wide access while protecting against possible harms. However, there are ongoing debates in the scientific community about the merits and limitations of different governance frameworks in achieving these twin aims; and indeed, best practices and points for consideration are notably absent when it comes to devising a governance framework for genomic databases. Based on our collective experience of devising and assessing governance frameworks, our Commentary identifies five key functions of "good governance" (or what makes "better governance") and three areas where trade-offs should be considered when specifying policies within those functions. We apply these functions as a benchmark to describe, as an example, the governance frameworks of six large-scale international genomic projects.

Introduction

Recent years have seen a dramatic increase in the collection, storage, and curation of human genomic data for biomedical research. To optimize the knowledge and benefits deriving from genomic data, managers of data repositories and funding organizations have increasingly sought to enable wide access to these resources. However, expanding access to human genomic data also intensifies a number of well-articulated ethical, legal and social concerns about the potential risks of these data collection efforts, such as privacy violations, misuse of data, and unauthorized access to data[1,2].

Genomic data repositories and consortia adopt governance procedures to address the dual objectives of enabling wide access while protecting against possible harms. There are ongoing debates in the scientific community about the merits and limitations of different governance approaches to achieve these twin aims, such as the adequacy of broad consent and the degree to which different stakeholders (including the public) have an opportunity to participate in governance[3]. What is currently missing is a comprehensive assessment of the ethically salient issues to be addressed. Part of the challenge is that different kinds of repositories and consortia may require different forms of governance. The purpose of this article, therefore, is to identify the functions that governance of genomic data should fulfil, as the basis for the design, implementation, and evaluation of governance frameworks for particular cases. We do not advocate for or against particular governance frameworks. Instead, we identify five key functions of "good governance" and examine three areas where tensions may arise between achieving competing functions and where trade-offs need to be considered when specifying policies. We illustrate these issues with the governance frameworks of six large-scale international genomic projects.

Key functions of good governance

One key function of good governance is enabling data access. Making genomic data widely available supports research efficiency and scope and is the underlying justification for data repositories and biobanks. There are several challenges, however, to wide data access, including (i) legal and technical barriers that may hinder the ability to share data across jurisdictions (e.g., real or perceived regulatory constraints, lack of interoperability), (ii) the ongoing sustainability of a data repository, including the willingness and ability of researchers to contribute high-quality data[4], (iii) lack of transparency regarding the governance arrangements of the repository, including such issues as data access processes and licensing, and (iv) arrangements that allow private sector collections to limit public access to their data, even when they build upon publicly funded research. To address these challenges, a good governance framework should provide appropriate incentives for researchers to contribute and make data available, address logistical and jurisdictional barriers, and adopt transparent policies and procedures for equitable data access. Across the six genomic projects analyzed (Table 1), the majority (with the notable exception of the Personal Genome Project) aim to make aggregate data available to vetted researchers, who can in turn (subject to governance approvals) contact participants for access to individual data. The ease by which researchers across different regions of the world can access these data, however, remains subject to wide variation. Efforts by different organizations across the globe remain ongoing to develop governance solutions to reduce legal and technical barriers to making data available, and to develop tools to incentivize researchers to make data more widely available[5,6].

A second key function of good governance is compliance with applicable national laws and international agreements. Rules adopted by a data repository must adhere to relevant laws governing matters such as data protection, human subjects research and genomic data sovereignty[7]. However, regulations in these areas are often complex, vague on the specifics of sharing genomic data, and vary considerably internationally. There can be multiple domestic, international, and professional standards that may apply, and international regulations (such as the General Data Protection Regulation (GDPR)[8], a European Union law that protects personal data) might be interpreted differently by various institutions or countries. Given the variability in interpretation of the same regulations, and differences across jurisdictions in law/regulation, good governance should specify what regulations apply and ensure that the framework is compliant with them[9].

A third key function of good governance is supporting appropriate data use and mitigating potential harms. Widening access to genomic data could lead to a variety of uses with potential for informational, financial, material, and psychosocial harms. In many jurisdictions, safeguards exist to prevent harms, but use of genomic data could result in unintentional harms or objectionable research (as perceived by different groups of stakeholders) even without breaking laws[10]. For instance, personal genomic sequencing data generated by direct-to-consumer companies or other businesses (to which regulations restricting the use of healthcare data, depending on the jurisdiction, might not apply) could be used to conduct warrantless surveillance, deny or limit access to health or other social resources, to deny entry into a country, or to undermine the reputation of

particular population groups[11]. There may be overriding collective concerns about data use that could be reasonably foreseen to cause harm to groups, such as stigmatizing particular ethnic groups, even when individuals have given consent for the use of their data in research[12]. Participants may also find some uses of data to be objectionable on moral, religious, or cultural grounds, with such uses influencing their willingness to donate samples to biobanks or repositories[13]. Notably, these concerns cannot be addressed by laws and regulations established for protecting personal identifiable data. Therefore, a good governance framework should specify the scope of research for which data may be used, including any restrictions based either on the original consent or on guidelines generated for the repository, and specify measures it will use to mitigate or prevent unintended harms and misuses, including through transparent decision-making and oversight processes.

A fourth key function of good governance is promoting equity in access, use, and analysis of genomic data. Potential equity barriers to exchanging genomic data occur as a result of unequal opportunities for researchers to access, use, or analyze data as a function of local capacity, specifically limitations in human capital, fiscal resources, and technological sophistication[14]. Inequities in research capacity are most evident between resource rich and resource poor nations, though they also arise within nations of both types. In particular, there may be limited capacity for the interpretation of genomic data among groups who are instrumental in providing those data[15], as well as differences in the capacity to benefit from generating genomic data. This is evident in the fact that people of European descent still account for 88 percent of the genomes in GWAS, which form a key source of information for genetic reference databases[16]. Finally, there is the potential for genomic data to be used in ways that exacerbate, rather than reduce, health care disparities across or within societies, especially if there are inequities in the underlying data collection and analysis processes[17]. For example, genomic research projects investigating the prevalence of obesity and type 2 diabetes, which disproportionately affect minority populations in the United States, might in fact exacerbate health disparities among a wide segment of a society if genomic explanations are emphasized rather than integrated into broader social models of disease and interdisciplinary research methods[18]. A good governance framework should identify measures to alleviate inequities in access, use, and analysis. Here, we note, as one example, the effort of the Human Heredity and Health in Africa (H3Africa) Initiative (Table 1) to boost capacity building for Africa-based scientific efforts and to encourage genomic research that benefits African populations across the continent.

A fifth key function of good governance is using genomic data for public benefit. Genomic databases may require significant public resources and their use can affect whole populations and societies. This implies an obligation to act for the public good. However, what constitutes the public good is not always self-evident and what is considered "good" for some may be detrimental or irrelevant to others. When management of health data has been viewed as objectionable, this has led to a breakdown in relationships of trust and loss of important data and associated research benefits[19]. Preconditions for trust vary over time and are contingent on the histories of particular communities, including their experiences of marginalization, exploitation, and past relationships with researchers and governments. A good governance framework should clarify how its operations enhance public trustworthiness and the public good. These might include mechanisms for meaningful

patient and public engagement in which publics are involved in formulating what constitutes public benefit for uses of genomic data and how particular data may be used[20], either as one-time public deliberation processes for particularly contentious issues[21], or formation of bodies such as committees or community advisory boards to provide ongoing public input to, and oversight of, a repository or consortium's management[3].

Tensions and trade-offs

As is clear from the discussion above, governance frameworks must consider how different governance functions may be in tension with one another. In such cases, governance must consider how to balance competing values, to which degree one might be prioritized over another in particular contexts, and who should be responsible for making and reviewing these decisions. Here, we consider three key trade-offs.

A tension and trade-off involves data access control. The fundamental trade-off for secondary use of genomic data relates to providing unrestricted access to data versus introducing oversight and restrictions to ensure appropriate data uses. *Open access*, which is endorsed by the Personal Genome Project (Table 1), in principle offers more immediate availability of data to any researcher, thus promoting (more) equitable access and more opportunities to investigate research questions, as well as opportunities to expand participation in the research process by non-professionals, such as through citizen science. This approach supports wide data access but provides no means to address potential objectionable uses, ensure equitable outcomes, or protect individuals and/or communities from informational and other harms. In contrast, *controlled access* offers the ability to vet appropriate research use of the data, and to assess whether data users are qualified and trusted to comply with data use requirements (e.g., the commitment not to re-identify individuals). A fair number of genomic projects operate a controlled access model, as reflected in the examples from Table 1. Intermediate approaches, such as *registered access*[22], allow data access to individuals who have been vetted, affording them more immediate availability, but like controlled access may delay access to the data. These different access models are the subject of live debate and exploration by different organizations. As one example of recent initiatives to address aspects of data access control, the Global Alliance for Global Health (GA4GH) has advocated the benefits of registered access model as a means to advance responsible and harmonized genomic data access and sharing, via its "GA4GH Passports and the Authorization and Authentication Infrastructure"[23].

A second tension and trade-off involves data de-identification. Data are typically de-identified by removing information such as name and other information that could easily identify an individual. This offers substantial, but not complete, privacy protection. A second option for de-identifying data is anonymization, which means full, irreversible destruction of the link between identifiers and individual level data. However, given the nature of genomic data, which includes uniquely identifying information about the participants, genomic data cannot be considered anonymous, even when de-identified. Another option is pseudonymization (also known as key-coding), whereby the key-code is retained but kept separately. Pseudonymization may achieve a better balance in genomic and health-related research whereby data can still be linked and participants can be re-contacted as needed, but

privacy-protecting measures are also enhanced. Pseudonymized data and anonymized data may also be treated differently in particular regulatory contexts[24].

While anonymization may have initial appeal, retaining individual identifiers may enhance the value of the data by allowing: linkage to other data sources such as electronic medical records; longitudinal data collection from participants; consent from participants for new, future uses of data not envisioned in the original consent form; reports to participants about research findings, either as a routine practice or under specific circumstances (e.g., research identifies a medical finding that triggers a duty of care); and participants to withdraw or access data.

We note that genomic data may be also made available in an aggregate-level form, via the publication of summary statistics (e.g., "privacy-preserving" statistics for GWAS studies and genome "Beacon" queries[25]). In practice, there are several techniques that are implemented in collaborative research efforts to mitigate the privacy risks associated with the sharing of genomic data[26-28]. We also note that important advances in computational science mean that new forms of data protection may become available in the future, such as running analyses on encrypted data and running analyses in distributed formats[29].

A third tension and trade-off involves designing or navigating different models of consent. Researchers may have compelling reasons to use data for purposes not described in the original consent form. One option is to provide participants the opportunity to re-consent specifically – or to opt in or out – of additional research studies. Empirical research indicates that some participants value this opportunity[30]. This approach requires an interface between the researcher and the research participant so that requests for participation can be made, with a link between the participant and the individual-level data. Key trade-offs here are between preservation of voluntary participation in research versus: (1) reduced availability of data for research; (2) time and resources required for the re-consent or opt out process; (3) potential for loss of representativeness of sample; and (4) privacy risks associated with maintaining a system to re-contact participants. An alternative approach, and practiced by a number of genomic research projects (as seen in Table 1), is *broad consent*: a consent approach that informs the participant about broad categories of future secondary uses, sometimes within certain boundaries (e.g., "cancer-related research", "no commercial use"), which is generally subject to ongoing governance oversight by a research ethics or data access committee. In this approach, participants are not informed about the specificities of data use; in essence, they are asked to consent to specified governance of their data and participation[31, 32]. We also note another approach for consent is consent for broad sharing and future research use, which the US National Institutes of Health (NIH) has issued guidance on in the context of genomic studies[33]. Finally, dynamic and meta-consent models enable people to select different consent preferences using digital resources to record individual consents. While meta-consent has set preferences for how and when to be asked for consent, dynamic consent, dynamic consent enables a range of different kinds of consents to be offered to individuals tailored to changing research needs over time and enables longitudinal bi-directional communication[34].

An illustration using governance frameworks in six projects

Table 1 describes the governance frameworks of six large-scale international genomic research projects: the Human Heredity and Health in Africa (or H3Africa) Initiative; the All of US Research Program; the Personal Genome Project; the Taiwan Biobank initiative; the Program for Engaging Everyone Responsibly (PEER); and the 100,000 Genomes Project. The projects are used to illustrate governance choices, as well as their approaches to important trade-offs and how those are reflected in their governance functions, given contextual factors. These six projects were selected for diversity of setting and approaches, not to necessarily exemplify best practices.

Table 1 draws on publicly available information about the projects as well the knowledge of co-authors who have worked on some of the projects. The information presented in the table is necessarily abbreviated and is intended primarily as an illustration of the governance functions we have identified. The table lists the main aims of each project, the trade-offs that are considered in the governance framework, and the degree to which each framework can be seen to fulfil the five functions of good governance. Importantly, these examples illustrate differences in transparency with regard to the information they provide about their governance approach.

There are points of similarity and difference across these governance frameworks. For instance, PEER allows for participants to provide consent or decline specific studies whereas a PGP consent form notes, "You may dislike or be upset by some … uses" of PGP data; and "Neither you nor the PGP will be able to restrict or specify the type of research or other purposes for which your cell lines will or will not be used." The other frameworks involve some form of centralized access to data, allowing the governance process to determine whether the proposed use of data is acceptable; however, criteria for making this determination are generally not specified. The H3Africa model aims to develop research capacity on the African continent, and thus gives H3Africa's researchers a much longer exclusive period of data access and use (23 months) than other projects in, say, Europe or North America. This may prioritize greater equity in collection, utilization, and benefits of genomic data. The All of Us Research Program enacts a more open model for data access – pushing a "registered" rather than "controlled" access mechanism. It also has made a concerted effort to target traditionally under-representative groups to participate, promoting equity. The 100,000 Genomes Project is noted both for its Participant Panel and independent Ethics Advisory Committee, which illustrates a bottom-up effort for stakeholders to feed into the decisions made by the Access Review Committee (the DAC) and help promote publicly accepted uses of genomic data. However, as is the case in other governance approaches, there is a lack of certainty about the capacity of an ethics advisory committee to effectively monitor and enforce ethical norms.

Conclusion

Good governance of genomic data should address several key functions and consider the trade-offs inherent in addressing the rights and interests of different stakeholders. Different contexts will result in different emphases in prioritizing the issues. As a result, there is

no single "best" governance framework, but some are certainly better than others. For example, we note that failing to account for, and sustain, the five functions of good governance may significantly compromise a project's ongoing social license to operate[35]. In addition, how one governance function is addressed may influence others. For example, if a repository has robust governance that adequately addresses all functions identified above, then secondary use with broader consent may be more acceptable. Because of these complexities, we argue that effective governance must be sensitive to relevant contextual factors and may legitimately vary. Nevertheless, governance systems should be transparent about how (or whether) they address each key function, how particular trade-offs were made, and who had input in those decisions. Transparency should extend to how governance committees or advisory boards are formed and what decision-making authority each holds; yet this information is often not readily available. Indeed, we see transparency as a meta-function of good governance which, unlike the other functions, is not something that can legitimately vary by context or be balanced against other dimensions of good governance. An important issue we have not addressed is what entities and what mechanisms would be involved in oversight with respect to adherence to various principles of governance or governance frameworks. This issue requires detailed analysis of the complex considerations of integrating data governance frameworks within various levels of existing legislation and policy in local contexts, and as a result, is beyond the scope of this article. Although our focus has been on genomic data, we believe that many of the considerations are also relevant to other forms of personal health data. Finally, we note that the private and philanthropic sector is playing an increasingly important role in facilitating human genomic data collection and sharing. Though our focus in this article is primarily on publicly funded projects, our core messages apply equally to other sectors.

Acknowledgements and funding information:

Kieran C. O'Doherty is supported by the Canadian Institutes of Health Research and the Social Sciences and Humanities Research Council.

Heidi Beate Bentzen is supported by the Research Council of Norway grant number 238999 and NordForsk grant number 81105.

Pascal Borry is funded through Horizon 2020 project euCanSHare (an EU-Canada joint infrastructure for next-generation multi-Study Heart research, project number 825903).

Michael Burgess is supported by the Canadian Institutes of Health Research, the Social Sciences and Humanities Research Council, The National Science Foundation, the Patient-Centered Outcomes Research Institute, and the Research Council of Norway.

Jantina De Vries is supported by the National Human Genome Research Institute of the National Institutes of Health under Award number 1U54HG009790 and by the Wellcome Trust under Award number 219600_Z_19_Z.

Eric Juengst is supported by the National Human Genome Research Institute of the U.S. National Institutes of Health under award number R01HG010661.

Kazuto Kato is supported by the Japan Agency for Medical Research and Development (AMED); grant number JP19kk0205012.

Bartha Maria Knoppers is funded by Genome Canada, Genome Quebec, and the Canada Research Chair in Law and Medicine.

Barbara Koenig is supported by the Eunice Kennedy Shriver National Institute of Child Health and Human Development and the National Human Genome Research Institute of the National Institutes of Health under Award Number U19HD077627.

Spero M. Manson is supported by 1U54MD011240 (Manson, PI) [National Center on Minority Health and Health Disparities, National Institutes of Health]; and 1P30DK092923 (Manson, PI) [National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health].

Amy Lynn McGuire is supported by NIH-NCI R01CA237118 (Cook-Deegan and McGuire, PI) and NIH-NHGRI R01HG008918 (McGuire and Cook-Deegan, PI).

Dianne Nicol, Lisa Eckstein and Don Chalmers are supported by Australian Research Council (ARC) DP 180100269.

Adrian Thorogood is supported by Genome Canada, Genome Quebec, and the Canadian Institutes of Health Research.

References

1. Widdows HThe Connected Self: The Ethics and Governance of the Genetic Individual (Cambridge University Press, 2013).
2. Taylor MGenetic Data and the Law: A Critical Perspective on Privacy Protection (Cambridge University Press, 2012).
3. Juengst ET & Meslin EM Kennedy Inst. Ethics J 29, 67–95 (2019). [PubMed: 31080178]
4. Cadigan RJet al.Genet. Med16, 738–740 (2014). [PubMed: 24763288]
5. Bezuidenhout LDev. World Bioeth19, 18–24 (2019). [PubMed: 29356295]
6. Corpas Met al.PLoS Comput. Biol14, e1005873 (2018). [PubMed: 29543799]
7. Hudson Met al.Nat. Rev. Genet21, 377–384 (2020). [PubMed: 32251390]
8. Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation).
9. Knoppers BMHUGO J. 8, 3 (2014). [PubMed: 27090251]
10. Fullerton SM & Lee SS-J BMC Med. Ethics 12, 16 (2011). [PubMed: 21943371]
11. O'Doherty KCet al.BMC Med. Ethics17, 54 (2016). [PubMed: 27600117]
12. Wang YHuman Population Genetic Research in Developing Countries: The Issue of Group Protection (Routledge, 2015).
13. DeVries RGet al.Life Sci. Soc. Policy12, 1 (2016). [PubMed: 26790860]
14. National Academies of Sciences, Engineering, and Medicine; Health and Medicine Division; Board on Health Sciences Policy; Roundtable on Genomics and Precision Health, Understanding Disparities in Access to Genomic Medicine: Proceedings of a Workshop (National Academies Press, 2018).
15. Claw KGet al.Nat. Commun9, 2957 (2018). [PubMed: 30054469]
16. Mills MC & Rahal C Commun. Biol 2, 9 (2019). [PubMed: 30623105]
17. Martin ARet al.Nat. Genet51, 584 (2019). [PubMed: 30926966]
18. Burke W, Trinidad SB & Schenck D Ethn. Dis 29, 669–674 (2019). [PubMed: 31889772]
19. Sterckx Set al.Med. Health Care Philos19, 177–90 (2016). [PubMed: 26280642]
20. Kaye Jet al.Hum. Genomics 201812, 13 (2018).
21. O'Doherty KC, Burgess MMResearch Ethics9, 55–77 (2013).
22. Dyke SOMet al.Eur. J. Hum. Genet26, 1721–31 (2018). [PubMed: 30069064]
23. Global Alliance for Genomics and Health, GA4GH Passports and the Authorization and Authentication Infrastructure, available at https://www.ga4gh.org/news/ga4gh-passports-and-the-authorization-and-authentication-infrastructure/.
24. Shabani M & Marelli L EMBO Rep. 20, e48316 (2019). [PubMed: 31126909]
25. Global Alliance for Genomics and Health and ELIXIR, GA4GH Genome Beacons, available at https://beacon-project.io/tags/queries.html.

26. Berger B & Cho H Genome Biol. 20, 128 (2019). [PubMed: 31262363]
27. Bonomi L, Huang Y & Ohno-Machado L Nat. Genet 52, 646–654 (2020). [PubMed: 32601475]
28. Mittos A, Malin B & De Cristofaro E Proceedings on Privacy Enhancing Technologies 1, 87–107 (2019).
29. Bonte Cet al.BMC Bioinformatics19, 537 (2018). [PubMed: 30572817]
30. Dixon-Woods Met al.BMC Med. Ethics18, 22 (2017). [PubMed: 28330487]
31. Lensink MAet al.Am. J. Bioeth19, 43–45 (2019).
32. Koenig BAHastings Cent. Rep44, 33–34 (2014).
33. National Institutes of Health (USA), NIH Guidance on Consent for Future Research Use and Broad Sharing of Human Genomic and Phenotypic Data Subject to the NIH Genomic Data Sharing Policy, available at https://osp.od.nih.gov/wp-content/uploads/NIH_Guidance_on_Elements_of_Consent_under_the_GDS_Policy_07-13-2015.pdf.
34. Budin-Ljøsne Iet al.BMC Med. Ethics18, 4 (2017). [PubMed: 28122615]
35. Harmon SHE, Yen S-Y & Tang S-M SCRIPTed 15, 103–129 (2018).

Table 1:

Governance frameworks of six large-scale international genomic research projects.

Table 1 provides an overview of the governance frameworks of six large-scale international genomic research projects. The frameworks have been categorized in line with the analysis in the paper, focusing on: (1) the model of consent (e.g. broad, open specific); (2) the model of data access (e.g. what types of data are made available and under what conditions); (3) the form(s) of compliance with relevant national laws and international agreements; (4) the ways in which harmful uses of data might be mitigated; (5) the extent to which, and if so how, equity in collection, utilization, and benefits of genomic data is deployed in the project; and, finally, (6) the extent to which consideration is given by the project to using genomic data for public benefit.

Name of the project	Aim	Consent	Data Access	Compliance with relevant national laws and international agreements	Mitigating harmful uses of data	Equity in collection, utilization, and benefits of genomic data	Use of data for public benefit
Human Heredity and Health in Africa (H3Africa) Initiative (https://h3africa.org/)	-To facilitate study of genomics and environmental determinants of common diseases, with the goal of improving the health of African populations. -Generates new data	Tiered and Broad	-Centralized data access control - Priority is given to H3Africa researchers (23 months of exclusive access); to researchers that can demonstrate capacity building in Africa thereafter.	Researchers requesting access must confirm legal compliance	Ethics and Regulatory Issues Working Group and Community Engagement Working Group established, but less effort directed to soliciting the opinion of the general public and research participants, which can inform what should be considered as harmful or objectionable data uses	Capacity building for Africa-based scientific effort	Emphasis on research that benefits African populations
Personal Genome Project (PGP) (https://www.personalgenomes.org/us)	-To facilitate genomic research -Generates new data and organizes existing data	Open	Identifiable data are publicly available. In addition, cell lines are made available.	-IRB review	Deliberately not addressed.	-Participants are provided access to their own data -All data are publicly accessible	The project seeks to advance scientific progress as a public good by making data publicly accessible data.
All of Us Program (https://allofus.nih.gov)	-To facilitate the development of precision medicine, through development of a large, well characterized research cohort. -Generates new data	Broad	-Aggregate data available in a public browser -Registered users approved by a designated committee can access de-	Research must comply with all laws governing NIH-funded research.	Registered researchers may study any topic that meets "criterion for allowable use."	Recruitment strategies emphasize recruitment of groups typically underrepresented in research.	The goal of the project is to promote research to improve healthcare and population health; alignment of criteria for allowable use with this goal is implied.

Name of the project	Aim	Consent	Data Access	Compliance with relevant national laws and international agreements	Mitigating harmful uses of data	Equity in collection, utilization, and benefits of genomic data	Use of data for public benefit
			identified individual-level data				
Program for Engaging Everyone Responsibly (PEER) (http://geneticalliance.org/programs/biotrust/peer/faqs)	-To provide communities (disease, environmental, social) methods to answer research questions advanced by the community itself; to accelerate health-related discovery, both disease-specific and cross condition; and promote trust. -Generates new data and organizes existing data	Specific	-Aggregate data available to vetted researchers, who can contact participants for access to individual data - Each individual determines how much data to share, and with whom – granular and dynamic approach. -Individual data remain identifiable	-GDPR and CCPA compliant -Compliant with SEC requirements (share participation limited to US residents) -IRB review	-Ethics Committee defines allowable research -Participants consent to research involving individual level data	-Participants share in benefits of research	-Emphasis on research addressing questions of importance to participating communities
100,000 Genomes Project (https://www.genomicsengland.co.uk/about-genomics-england/the-100000-genomes-project)	-To sequence 100,000 whole genomes from NHS England patients with rare diseases and their families, and patients with common cancers, to enable new scientific discovery and medical insights and kickstart the development of a UK genomics industry. -Generates new data	Broad	-Centralized data access through Access Review Committee (ARC) - All information related to individual identification is removed before data release.	-Review by Genomics England legal counsel and Ethics Advisory Committee	-Acceptable uses determined by ARC review	-Genomics England Clinical Interpretation Partnership (GeCIP) has been created to bring together funders, researchers, NHS teams and trainees to analyze the data and help ensure benefits for patients and an increased understanding of genomics. The data will also be used for medical and scientific research.	Ethics Advisory Committee, acts to identify, define, and respond to ethical issues in the Project, and helps to ensure the Project is delivered in the interests of the public and of participants. −100,000 Genomes Project Participant Panel ensures that the data collected by the project is being used in the best interests of the participants and is looked after with respect.
Taiwan Biobank (https://www.twbiobank.org.tw/new_web_en/	-To determine the effects of genetic and environmental factors and interactions on common diseases, and to develop personalized medicine. Biological samples from 200,000 healthy participants aged 30-70 linked with lifestyle, family history,	Broad	-Centralized access through Data Release Group of Taiwan Biobank -Anonymized data (individual genotype and phenotype data) are made available to bona fide researchers upon application	-Biobank must operate in accordance with the Human Biobank Management Act 2010	-The project receives independent ethics advice from an IRB and an Ethics and Governance Council -Biobank custodians must act in compliance with medical and research ethics,	Under the Human Biobank Management Act 2010, research uses must be authorized (by the EGC), and the principles of fairness and equality shall apply to data access	Under the Human Biobank Management Act 2010, research uses must be authorized (by the EGC), and the principles of fairness and equality shall apply to data access

Name of the project	Aim	Consent	Data Access	Compliance with relevant national laws and international agreements	Mitigating harmful uses of data	Equity in collection, utilization, and benefits of genomic data	Use of data for public benefit
	and health information. -Generates new data				as per the Human Biobank Management Act 2010		

Abbreviations: CCPA = California Consumer Privacy Act; GDPR = EU General Data Protection Regulation; IRB = Institutional Review Board; NHS = National Health Service (UK); NIH = U.S. National Institutes of Health; SEC = U.S. Securities and Exchange Commission