An International Framework for Data Sharing: Moving Forward with the Global Alliance for Genomics and Health

Vasiliki Rahimzadeh, Stephanie O.M. Dyke, and Bartha M. Knoppers

The Global Alliance for Genomics and Health is marshaling expertise in biomedical research and data sharing policy to propel bench-to-bedside translation of genomics in parallel with many of the BioSHaRE-EU initiatives described at length in this Issue. Worldwide representation of institutions, funders, researchers, and patient advocacy groups at the Global Alliance is testament to a shared ideal that sees maximizing the public good as a chief priority of genomic innovation in health. The Global Alliance has made a critical stride in this regard with the development of its *Framework for Responsible Sharing of Genomic and Health-related Data.*[1] This article first discusses the human rights pillars that underlie the *Framework* and mission of the Global Alliance. Second, it outlines the Global Alliance's use of data governance policies through a number of demonstration projects. Finally, the authors describe how the Global Alliance envisions international data sharing moving forward in the postgenomic era.

Introduction

THE GLOBAL ALLIANCE for Genomics and Health (GA4GH) is marshaling expertise in biomedical research and data sharing policy to propel bench-to-bedside translation of genomics in parallel with many of the BioSHaRE-EU initiatives described at length in this Issue. The mission of the GA4GH—herein referred to as the Global Alliance—is to accelerate progress in human health through establishing a common framework for policy harmonization and catalyzing projects that demonstrate the value of open data sharing.[2] Global Alliance stakeholders collaborate in four Working Groups (Regulatory and Ethics [REWG]; Clinical; Data; and Security) to ensure participants make informed choices toward responsibly and securely sharing their genomic and clinical data.[3,4] They do so against a background of data secrecy that encourages disciplinary silos rather than integration,[5,6] and has hitherto thwarted the development of data tools that enhance interoperability. The Global Alliance is challenging existing norms in this regard.[7] Collaboration through data sharing is key,[8] while fostering and catalyzing interoperability is the necessary vehicle for future ethics governance in an increasingly data-driven era of biomedical research.

Worldwide representation of institutions, funders, researchers, and patient advocacy groups from 43 countries at the Global Alliance is testament to a shared ideal that sees maximizing the public good as a chief priority of genomic innovation in health. Only recently have the scientific and ethical merits of data sharing prompted evolution in data governance policies pursuant to these shared ideals.[9] The Global Alliance has made a critical stride in facilitating open data sharing with the development of its *Framework for Responsible Sharing of Genomic and Health-related Data.*[1] Its constitution makes adhesion to the *Framework* the only obligatory criterion for membership in the Global Alliance. BioSHaRE-EU was one of 15 consortia and research institutions that contributed to the drafting of the *Framework*. The purpose of the *Framework* is to provide a principled and practical approach for the responsible sharing of genomic and health-related data. The primary goals aim to:

- Protect and promote the welfare, rights, and interests of individuals from around the world in genomic and health-related data sharing, particularly those who contribute their data for biomedical research.
- Complement laws and regulations on privacy and personal data protection, as well as policies and codes of conduct for the ethical governance of research.
- Foster responsible data sharing and oversight of research data systems.
- Establish a framework for greater international data sharing, collaboration, and good governance.
- Serve as a dynamic instrument that can respond to future developments in the science, technology, and practices of genomic and health-related data sharing.

Centre of Genomics and Policy, Montreal, Canada.

- Serve as a tool for the evaluation of responsible research by research ethic committees and data access committees.
- Provide overarching principles and policies that can be referred to in legally-binding tools such as data access agreements.[1]

The Global Alliance actualizes a paradigmatic shift in meeting the challenges of data-driven research and medicine. Since 2013, its approach to data sharing has complemented traditional bioethics, focused largely on protection from harm, by adopting human rights as its foundation. The article elaborates on these human rights pillars that underlie both the mission and objectives of the Global Alliance. Second, it outlines the use of its data governance *Framework* and policies through a number of demonstration projects. Finally, the article describes how the Global Alliance envisions international data sharing moving forward in the postgenomic era.

Human Rights

The globalization of genomic research requires an equally globalized approach to regulation and governance. It is the Global Alliance's international subscription to Article 27 of the UN Declaration of Human Rights that all humans have a right to benefit from the fruits of medical progress and that all scientists have the right to be recognized for their contributions[10] that anchor the *Framework*. The legal force of human rights entrenches and makes actionable the shared ideals of collaboration in data governance that include, but are not limited to, privacy, access, storage, and sharing. As such, the human rights approach "reach[es] beyond the moral appeals of bioethics and can provide a more robust governance framework for the regulation of genomic research."[11] Invoking human rights as the foundation for international data sharing efforts further coheres with the collective heritage of the human genome itself[12] that transcends claims of genomic sovereignty and biopolitics.[13] Moreover, scientists (and institutions) should be duly recognized for their contribution(s) to the greater understanding of the genome and their application toward improving human health.

Taken together, the "right to benefit from," and the "right to be recognized for" ground the data sharing initiatives and policies of the Global Alliance and translate into four foundational principles of the *Framework*:

- Respect Individuals, Families, and Communities.
- Advance Research and Scientific Knowledge.
- Promote Health, Well-Being, and the Fair Distribution of Benefits.
- Foster Trust, Integrity, and Reciprocity.[1]

Policies that activate these principles enable translational genomic research to proceed in the harmonization of privacy, antidiscrimination, and fair access procedures in the context of a human rights approach to data sharing.[11,14] This is particularly true for the demonstration projects described in the subsequent section.

Demonstration Projects

Ethics interoperability and harmonization are manifest in the Global Alliance work products, including consent tools and policies, as well as other policy tools for data sharing and data access/security infrastructure.[15–17] These products are developed and piloted in demonstration projects, of which the Global Alliance launched three in 2014: the Beacon Project, the BRCA Challenge, and Matchmaker Exchange. Common to each of these projects is the categorization and management of "sensitive data." Recommendations on the ELSI status of "sensitive" data are under development to support the various demonstration projects. These distinguish between degrees of sensitive of data recognized in policy and law. The demonstration projects utilize various tiers of data access. While open and controlled access have been the most familiar mechanisms of data access, the Registered Access Task Team of the REWG is currently developing the concept of, and legal implementation for the middle access tier, "registered access." The Task Team was established, in part, to determine from both practical and legal standpoints how registered access can be operationally different from open and controlled access, and with the intent to responsibly enable sharing of some forms of sensitive data. The data sharing objectives and spectrum of data access mechanisms for each demonstration project are discussed in turn below.

The goal of the Beacon Project is to bridge connections between researchers who share the need for specific genomic information. Beacons are simple public web services designed to respond to simple researcher queries such as "Do you have record of any genomes with an 'A' at position 100,735 on chromosome 3?"[18] Similar to its use in navigation, a beacon serves as a genotypic signal or flag denoting a particular variant that any researcher can further investigate by accessing the dataset. To date, over 47 Beacons have been created worldwide. Consent Codes describing data use conditions that are based on consent have also been produced to facilitate appropriate secondary use of data.[19]

The BRCA Challenge enables large-scale international sharing of data on BRCA variants and their rapid interpretation for clinical purposes. Eventually, the BRCA Challenge data sharing platform will couple sequence variation to scientific and medical evidence that, together, will improve patient diagnosis and prevention.[20] The BRCA Challenge fosters responsible collection, curation, and sharing of cancer data internationally. It relies on expert support from a task team of the REWG, the BRCA Challenge Ethics, and Legal and Advocacy Task Team to operationalize elements of the *Framework* in the areas of consent, liability, health privacy, data protection and access, clinical practice, genetic testing, and patient advocacy. Inaugural activities for the task team include establishing the terms and conditions for the various tiers of data access to the eventual BRCA variant portal for different types of users (e.g., clinician, researcher, curator) with the intent to maximize global participation and utility.

Finally, the Matchmaker Exchange demonstration project[21] emerged from a meeting convened with the support of the International Rare Diseases Research Consortium (IRDiRC), as well as the Clinical Genome Resource (ClinGen) program.[22] It "creates a federated platform (Exchange) to facilitate the matching of cases with similar phenotypic and genotypic profiles (matchmaking) through standardized application programming interfaces (APIs) and procedural conventions."[23] The Regulatory and Ethics Working Group developed a tiered consent policy for the project that is dependent on context (research or clinical care), as well as the risk of identifiability that certain data can potentially present.[23]

As their name suggests, the demonstration projects of the Global Alliance produce evidence-in-action that underscore the international need for greater data sharing in genomics that can eventually reach clinical implementation.

Future Directions

The Global Alliance continues to further international data sharing priorities and governance in partnership with BioSHaRE-EU collaborators and others at the frontlines of genomic medicine. Lessons from the above demonstration projects, in part, inspired the future directions that the Global Alliance addresses in both the short and long-term future of data governance. Now in its third year, the Global Alliance has grown in membership and projects, which now also include pediatrics, clinical cancer genomics, e-Health, and dementia research. It has diversified its sectors and geographic reach and strengthened the collaborative ethos conducive to a sustainable environment of data sharing. This expansion of data sharing spheres is evident in the progress of the REWG achieved thus far (including the proposal of a Data Sharing Lexicon establishing common terms for data sharing across jurisdictions and research contexts) and in the new Task Teams planned for 2016. These include the Individual Access Task Team, which explores participant rights to their own raw genomic or health-related data; the Ageing and Dementia Task Team, created to set out principles and policies for data sharing among the world's aging population living with dementia; and finally, the Pediatric Task Team, which develops policies, tools, and guidelines to accelerate the sharing of genomic research data involving children, in particular whole genome sequence data to be used in newborn screening.

In addition to working on the *Framework*, BioSHaRE-EU was particularly active in creating one of the newest REWG Task Teams to explore ethics review equivalency (ERE). The ERE Task Team is developing models for mutual recognition across jurisdictions in the research ethics review process. The policy harmonization focus is on research most affected by barriers to data sharing, namely data-intensive sciences and internationally collaborative projects. The ERE Task Team will then pilot these models with Global Alliance demonstration projects and the International Rare Disease Research Consortium. Two forthcoming initiatives are planned for the Global Alliance in the wake of its collaboration with BioSHaRE-EU. The first involves the "machine-readability" of consent. The REWG Automatable Discovery and Access (ADA) Task Team is working to transform human-readable information (e.g., consent forms, photos, pdf, etc.) into an information format that can be read and processed by a computer. The ADA Task Team will drive these efforts with members from the Data Working Group and the Security Working Group, as well as from the demonstration projects. Second, recommendations on the ELSI status of sensitive data—including the development of the registered access tier—will be field-tested and further refined with Matchmaker Exchange and the BRCA Challenge.

These forthcoming initiatives exemplify two continuing priorities of the Global Alliance: relevance and sustainability. No doubt the emergence of specialized international databases and research consortia testifies to the scientific imperative to share data. International initiatives to streamline data collection are more numerous and are gaining policy traction as a result. These emerging trends in the ways of doing biomedical research post sequencing of the human genome underscore the relevance of an international collective on data sharing, such as that which the Global Alliance fosters. In the spirit of collaboration, the Global Alliance plans to first identify and then align with major data collection and sharing efforts within the international community. The Global Alliance, by virtue of its dynamic nature, can also communicate strategically and efficiently with key audiences. The diversification of stakeholder involvement, particularly through its increasing engagement with patient advocacy groups, will sustain the Global Alliance as a thought leader on governance issues in "big data." Finally, while not incorporated or funding research itself, the Global Alliance envisions a sustainable future for the data governance mechanisms it catalyzes by expanding circles of engagement. Through continued networking, the Global Alliance is expanding organizational capacities and establishing new demonstration projects.

Conclusion

The success of the Global Alliance and BioSHaRE-EU's participation in its data sharing initiatives are the result of innovative science and policy. Unified under a shared objective of global data sharing, BioSHaRE-EU has demonstrated through solution-based evidence and collaboration that a million data points are always better than one. Global Alliance Working Groups and Task Teams are oriented toward solving what were, only less than two decades ago, complex and intractable barriers to data sharing. The integration of science alongside policy—and the utility of such interdisciplinary initiatives through demonstration projects and tools—allows the Global Alliance to leverage the value of data sharing against a (fading) culture of data hoarding. While policy does not grow on trees,[24] continued collaboration can reflect the priorities of the international community in pursuit of improved health benefits for all.

Acknowledgments

Author Disclosure Statement

No conflicting financial interests exist.

References

1. Global Alliance for Genomics and Health. Framework for responsible sharing of genomic and health-related data. https://genomicsandhealth.org/files/public/Framework%20 for%20Responsible%20Sharing%20of%20Genomic%20

and%20Health-Related%20Data%20-%20Version%2010%
20September%202014.pdf (accessed April 7, 2016).

2. Our Work|Global Alliance for Genomics and Health. Available from: http://genomicsandhealth.org (accessed July 21, 2015).

3. Kosseim P, Dove ES, Baggaley C, et al. Building a data sharing model for global genomic research. Genome Biol 2014;15:430.

4. Drazen J. Sharing individual patient data from clinical trials. N Engl J Med 2015;372:201–202.

5. Garrett SB, Koenig BA, Brown A, et al. EngageUC: Developing an efficient and ethical approach to biobanking research at the University of California. Clin Transl Sci 2015;8:362–366.

6. Sansone SA, Rocca-Serra P, Field D, et al. Toward interoperable bioscience data. Nat Genet 2012;44:121–126.

7. Hayden EC. Geneticists push for global data-sharing. Nature 2013;498:16–17.

8. Veltman JA, Lupski JR. From genes to genomes in the clinic. Genome Med 2015;7:78.

9. Dyke S, Hubbard T. Developing and implementing an institute-wide data sharing policy. Genome Med 2011;3:60.

10. General Assembly of the United Nations. Universal Declaration of Human Rights, 1948.

11. Knoppers BM, Harris JR, Budin-Ljøsne I, Dove ES. A human rights approach to an international code of conduct for genomic and clinical data sharing. Hum Genet 2014;133:895–903.

12. Knoppers BM, Joly Y. Our social genome? Trends Biotechnol 2007;25:284–288.

13. Hinterberger A, Porter N. Genomic and viral sovereignty: Tethering the materials of global biomedicine. Public Cult 2015;27:361–386.

14. Shabani M, Knoppers BM, Borry P. From the principles of genomic data sharing to the practices of data access committees. EMBO Mol Med. 2015;7:507–509.

15. Paten B, Diekhans M, Druker BJ, et al. The NIH BD2K center for big data in translational genomics. J Am Med Inform Assoc 2015;22:1143–1147.

16. Burton PR, Murtagh MJ, Boyd A, et al. Data safe havens in health research and healthcare. Bioinformatics 2015;31:3241–3248.

17. Baker D, Bobrow M, Burton P, et al. Genomic and clinical data sharing policy questions with technology and security implications: Consensus position statements from the data safe havens task team [Internet]. 2014. Available from: http://genomicsandhealth.org/files/public/FinalV2_Data Safe Havens Task Team_Deliverable_0.pdf (accessed August 13, 2015).

18. GA4GH. Beacon Project. Available from: http://ga4gh.org/ #/beacon (accessed August 13, 2015).

19. Dyke SOM, Philippakis AA, Rambla De Argila J, et al. Consent Codes: Upholding Standard Data Use Conditions. PLoS Genet. 2016;12(1):e1005772. DOI:10.1371/journal .pgen.1005772.

20. BRCA Challenge | Global Alliance for Genomics and Health. Available from: https://genomicsandhealth.org/work-products-demonstration-projects/brca-challenge-0 (accessed August 13, 2015).

21. Matchmaker Exchange API | Global Alliance for Genomics and Health. Available from: https://genomicsandhealth .org/work-products-demonstration-projects/matchmaker-exchange-api (accessed August 13, 2015).

22. Philippakis A, Azzariti D, Beltran S, et al. The matchmaker exchange: A platform for rare disease discovery. Hum Mutat 2015;36:915–921.

23. Matchmaker Exchange Tiered Informed Consent Proposal. Available from: www.matchmakerexchange.org/assets/files/ matchmaker_exchange_tiered_informed_consent_proposal .pdf (accessed July 28, 2015).

24. Knoppers BM. Does policy grow on trees? BMC Med Ethics 2014;15:87.

Address correspondence to:
Vasiliki Rahimzadeh, MSc
Centre of Genomics and Policy
740 Avenue Dr. Penfield Suite 5200
Montreal
Quebec H3A-OG1
Canada

E-mail: vasiliki.rahimzadeh@mail.mcgill.ca