PERSPECTIVE

Time for a change: considering the rights of study participants to ownership of their personal research-grade genomic data

Convergent Science Physical Oncology

PERSPECTIVE

Time for a change: considering the rights of study participants to ownership of their personal research-grade genomic data

RECEIVED
2 October 2018

ACCEPTED FOR PUBLICATION
12 December 2018

PUBLISHED
31 December 2018

Alice Lu[1,2], Peter Kuhn[3,4,5] and Ida Deichaite[1]

[1] Moores Cancer Center, University of California San Diego, La Jolla, CA, United States of America
[2] Pritzker School of Law, Northwestern University, Evanston, IL, United States of America
[3] Bridge Institute, University of Southern California, Los Angeles, CA, United States of America
[4] Bridge Institute, University of Southern California, 1002 W Childs Way, MCB-351, MC:3502, Los Angeles, CA 90089-3502, United States of America
[5] Author to whom any correspondence should be addressed.

E-mail: alicelu2019@nlaw.northwestern.edu (A Lu), pkuhn@usc.edu (P Kuhn) and ichaite@ucsd.edu (I Deichaite)

Keywords: genomic data ownership, biorights, raw research grade genomic data

Abstract

Determining the ownership of a patient's personal genomic data is important because it impacts how data is governed and shared, which has both clinical and research implications for precision oncology. The 21st Century Cures Act enacted in December 2016 defined the ownership of clinical genomic data, but the governance of research-grade genomic data remains a hotly contested topic. The many stakeholders often have competing perspectives about ownership of raw and processed genomic data derived in research settings and how to weigh risks versus benefits of sharing this data with study participants. A growing number of research studies, policy recommendations, and ethics reviews have not been enough to influence changes in practice. Most genomic research is conducted in academia, which is guided by Institutional Review Board-approved protocols to protect study participants. The current standard is to limit the return of research-grade data to study participants, and give data ownership solely to the researchers or the research institution, since this data is not vetted for clinical purposes and is meant for research use only. However, these practices conflict not only with recommendations from peer-reviewed literature on best practices for addressing research study participants' needs but might indeed run counter to legal and ethical guidelines about data ownership. For example, patient-participants faced with poorly understood or incurable diseases such as certain cancers want, and could potentially benefit from, having access to their personal genomic data in this rapidly evolving field. This commentary highlights the gap between the status quo as approved by the IRB and the literature suggesting that study participants should be given access to their personal genomic data. There is an opportunity to facilitate a more effective and ethical way to collect genomic data for research use across institutions.

Background

The fast-evolving fields of precision oncology and translational research have great potential clinical benefits for existing and future patients. As such, patients with aggressive terminal diseases such as cancer should be able to reap the clinical benefits of the translational science that is being conducted as part of research in which they participate, in the fastest timeline possible. However, the growing field of genomics is in many ways outpacing the development of accompanying regulation that would allow these study participants to directly benefit from the studies they participate in.

One of the most contested topics is the ownership of research-grade genomic data and by extension really a wide array of patient-derived research data. This data is unique and not to be confused with genomic data that is derived from clinical studies, laboratories that are CLIA-certified (Clinical Laboratory Improved Amendments) or medical devices that are PMA-certified (pre-market approval), which are beyond this discussion. The ownership of the raw and processed genomic data that arises from research is important because it dictates who can access this data and how it can be shared, downloaded, integrated with existing health data, or otherwise used. In most cases

with cancer patient-participants, the tumor sample that they provide for extraction of genomic data is also a unique sample that is difficult to obtain, often requires surgery, and cannot be replicated. This makes the governance of the ownership and usage of the original research genomic data critical because the research participant cannot simply go to another research lab or clinical lab if they wanted to obtain the same genomic data for themselves.

Genomic data ownership involves a large number of stakeholders that have incongruent goals and are governed by different sets of policies and priorities [1]. These include: patients/participants and patient advocacy groups, research teams and ethics committees, clinical staff, policymakers, funding groups, and relevant for-profit, non-profit, and professional organizations. While some see personal genomic data as a private good, others believe that it should be held as common property due to the nature of genomic research and personalized medicine needing vast amounts of data [2].

Among these different influences is also the federal law and regulation which aims to ensure the rights and safety of participants in human subject research. The most salient is the Privacy Rule from HIPAA (Health Insurance Portability and Accountability Act of 1996) that mandates a legal obligation to protect study participant privacy and confidentiality. There is also the Common Rule from the DHHS (Department of Health and Human Services), and the Human Subject Protection Regulations and 21st Century Cures Act from the FDA [3, 4].

These regulations are set as minimum requirements (e.g. prohibiting re-identification of anonymized data) in order to protect human subjects. The federal government is largely hands off with enforcement giving individual Institutional Review Boards (IRBs) the responsibility for peer-reviewing and approving informed consent documentation and study protocols—including the regulation of return of personal genomic data to study participants [3, 5]. As a result, these IRBs are placed at the interface of research and clinical care—two fields that traditionally have separate oversight mechanisms and very different norms, missions, and ethical frameworks. While the language of these regulations does permit individual IRBs to make discretionary judgment calls about the amount of risk imposed by a study's protocol, the IRBs tend to take conservative interpretations to pursue 'minimal risk' and 'practicability' [6]. When this is applied as a blanket restriction against returning research-grade data to study participants, the IRB is neglecting the perhaps even greater risk to participants of not being able to access their genomic data. This is especially true for study participants who are also patients of poorly understood and/or incurable diseases that rely on cutting-edge research to find new treatments and curative therapies.

Problem statement: ethical and legal concerns

In the past few years, there has been extensive discussion on the unique ethical challenges that genomic research brings [7, 8]. Many believe that the status quo IRB-approved protocols fail to adequately address ethical concerns about study participants rights, particularly in regards to 'future use of participant samples or data' and 'disclosure of results' [9]. IRB-approved protocols tend to be ambiguous on genomic data ownership and the return of research, incomplete for consideration of exceptions, and often reliant on traditional (dated) models of privacy and informed consent.

One specific controversy is the recent concept of 'biorights'—should study participants maintain an ongoing right to control their individual research samples and do they have the right to benefit directly from the research [10]? Would the potential benefits from receiving this data (e.g. patient education, productive communication between researchers and participants, and more information to inform clinical decisions) outweigh the potential risks (e.g. increased psychological distress, taking action on false positives, or other unintended consequences of sharing research-grade data)?

When it comes to research protocols that involve genomic data, the IRB-approved language defines the participant as a data donor (in this case donating material from which genomic data is derived), and informed consent as the vehicle for this essentially irrevocable donation to the research institution. However, IRB professionals have not reached a consensus regarding whether genomic studies should disclose incidental findings to study participants [11].

The IRB has traditionally approved informed consent language stating that there is no promised return of research or direct benefit of participating in the study for the study participant. This is important for preventing any risks that may arise from re-identifying data or allowing research-grade data that is exploratory in nature to be used for non-research purposes. However, IRB committee opinions often differ from the desires of the study participants whom they are working to protect [12]. Studies have repeatedly shown that even healthy study participants desire to know 'a broader array of genetic results than what is recommended', 'learn as much as possible about their health through the study', and gain control over deciding what data they want to receive and what data they want to share with others, including other study investigators who could make use of their data [13–15]. It seems that IRB-approved protocols are not currently consistent with what the public wants, in protecting the best interests of study participants and increasing access to data for advancing genomic research.

It is important to weigh the interests of the study participants with the risks of giving them access to their research-grade data that is not clinically reli-

able or valid. It can be argued that all research-grade data should be available to the patient upon request. Research-grade data is only ever 'research use' and must not be used for clinical decision making, but can in fact contribute to opinion forming much like any Google search does. A treating physician and a patient, i.e. decision makers, will distinguish between a web-result from a news group and one from a peer reviewed paper in much the same way. In fact, there is a strong argument that the education of how to evaluate research-grade data (or news group) data is critical for today's patients and physicians alike.

Traditional IRB-approved protocol regarding the governance of genomic data sharing has been long critiqued by researchers, study participants, and leaders of genomic research centers alike [10, 13, 15–17]. It has even been questioned whether the IRB is the appropriate body to oversee what results are returned to study participants because providing oversight on the clinical implications may be beyond the IRB's scope of practice [16]. The inability to give study participants access to their personal genomic data altogether, due to IRB interpretation of the federal regulations, can be seen as a major bottleneck for achieving the full potential of cancer genomic research.

Returning research-grade data to study participants

While many studies are already returning clinically actionable incidental findings to study participants, largely in accordance to the guidelines of the ACMG (American College of Medical Genetics and Genomics), these studies have varying mechanisms for divulging information on any genomic findings outside of the 56 genes recommended by the ACMG [16, 18, 19]. A recent analysis of IRB-approved informed consent documents for genome-wide association studies found that the majority either do not provide study participants full access to their personal genomic data or do not address the return of research [20]. There are two main issues at play.

First, many research institutions are unable to overcome the policies that prohibit them from re-identifying data or directly communicating with data donors, due to restrictive internal policies. This is likely due to issues of liability for the researcher or research institution regarding participant privacy, risks, and concerns for malpractice [16]. There is also a lack of standard operating procedures regarding how study investigators should provide this information to study participants. Given that the federal regulations and other guidelines can be interpreted differently by each institution, to both support and deny any return of research-grade genomic data, studies differ widely in whether they return raw research-grade data and how they do so [16, 21]. Some of the primary questions that researchers face when making these decisions include: participants' well-being, researcher responsibility, institution policies, recommendations from external entities, and which stakeholder they believe should be the decision-maker (participants, principal investigators, IRBs, and/or professional organizations) [22].

Second, many research institutions are unwilling to provide research-grade data to study participants at their request because this data is not derived from a CLIA-certified laboratory or PMA-certified medical device [16, 23]. In order to allow laboratory data to be returned to study participants, the status quo posits that the following should be established: 'analytical validity, clinical validity, clinical utility, clinical relevance and actionability of incidental genomic findings', as well as 'disclosure protocol as approved by the relevant institutional review board' [24]. Only clinical laboratories certified by CLIA adhere to protocols stringent enough to satisfy the former requirement. Accordingly, there is a general consensus in research laboratory protocols that when returning genomic results, these results should be either derived from or confirmed by CLIA-certified laboratories [21]. However, this notion is not fully grounded in empirical evidence or federal law. In terms of empirical evidence, there is insufficient data weighing the benefits and risks for participants receiving return of research-grade results for purposes other than the intended use of the research. Although CLIA data may be more stringently vetted, the downside is that its tests and the results are not as advanced because it takes time to get CLIA approval so CLIA data inherently lags behind cutting-edge research data. Our current understanding of genomics is unable to assign clinical significance to a majority of the information generated from genomic data, so CLIA data recipients can only view data for genes that are presently deemed important to them. If patients were able to access their full genome, they could then be afforded the opportunity to take charge of their clinical care by comparing their genomic data to up-and-coming literature and findings in genomic research. In terms of the law, a legal analysis of the situation posits that the CLIA regulations do not necessitate CLIA certification requirements when it comes to return of results for research studies. According to the 21st Century Cures Act, 'nothing in this subsection shall be construed to limit the access of an individual who is a subject of research to information about himself or herself collected during such individual's participation in the research' [4]. In fact, banning non-CLIA-certified laboratories from returning research-grade results to study participants may be at odds with the First Amendment [25, 26].

It is time for a paradigm shift. Could offering full return of research to participants be more ethically sound and confer more direct benefits for study participants? Ethically, retaining data control is a precondition to 'autonomy, self-determination, privacy, trust, transparency, and accountability' that must be preserved in human subject research to protect participants [13]. Therefore, an individual's control over

their personal genomic data should not be diminished unless there is adequate empirical evidence to do so (i.e. a direct causal relationship between their control rights and decrease in research dataset quality), which there currently is not [13]. Patient-participants faced with life-threatening diseases such as cancer have an even greater urgency with wanting access to their personal genomic data. Such patient-participants do not want to be 'unnecessarily hampered by restrictive rules that prevent, in the name of privacy, a patient from benefiting more directly from data they contribute' [15]. For these patient-participants, having access to their personal research-grade genomic data can be the difference between life or death, given how the rapidly evolving field of genomics and research breakthroughs can impact their clinical outcomes.

Case study: harvard returns raw research-grade genomic data to its study participants

One large-scale genomic sequencing initiative, the personal genome project (PGP) founded by Harvard Medical School, is creating a radical paradigm shift by embracing an improved model of consent that is 'dynamic' and 'open' in order to better address these ethical and legal concerns [27]. The study was founded in 2005, received high visibility for its participatory model and been described as bold, radical, and more democratic for its provision of raw research-grade genomic data to its study participants [13, 28, 29]. However, its practices were not widely adopted and such open access to raw research-grade genomic data remains highly unusual among other genomic research institutes. Three best practices from the PGP's participatory model are identified here to demonstrate the feasibility of these strategies in successfully minimizing risks of liability and participant harm while maximizing direct benefits to study participants [12, 26].

Learning module and quiz

Studies have shown that such use of digital education modules for the consenting process increases participant understanding for what they are consenting to [30]. A quiz then further ensures that study participants sufficiently understand the risks and benefits associated if they choose to receive and act on research-grade genomic data. There is also the collateral benefit of participants improving their health literacy.

Dynamic consent

The PGP researchers believed that informed consent should not simply be one initial static form, especially when federal regulation and the NIH are both endorsing the utilization of broad consent (i.e. one-time consent for unspecified future use of data) [31]. The PGP open access online consenting process allows participants to continuously control how their data can be used and shared after the first initial signing, to account for the fluctuating choices of these participants and address the ethical shortcomings of broad consent [13]. While researchers may fear that participants will then choose to restrict secondary usages of their data, preliminary empirical evidence shows that a large percentage of patients who participate in such research consider contributing to scientific knowledge as their primary goal [15, 32].

Communication between participants and researchers

The PGP created a proactive community of study participants by using the dynamic consenting process as a method for researchers to directly communicate with and support their study participants. There should be a push for reflecting on researcher-participant relationships and rethinking the boundary between them [33].

All in all, the PGP shows that there is a feasible, effective, and cost-efficient way for a research study to gather non-anonymized genomic data, return raw data to study participants, and allow communication between participants and researchers. The resulting increase in communication engaged thousands of participants to be proactive and continuously donate supplementary health data to the researchers. This model allows for both the accelerated advancement of research and public health, and the potential improvement of clinical outcomes for study participants. These results support the hypothesis that the benefits could outweigh the risks for giving patient-participants ownership of their genomic data.

Recommendations for future research

There are potential benefits to giving study participants access to their personal genomic data. It is time for IRB committees to reimagine their traditional stances on the return of research, and to appropriately update the necessary regulations. Successful case studies, such as Harvard's personal genome project, can be used as models for how research institutions can break down barriers that are traditionally built between researchers and their study participants.

Ultimately, we envision a system that will educate study participants and allow them to maintain access to, and control of, their research-grade genomic data. This will address the health disparity that is being created where only a minority of patients and study participants are known to have received access to their research-grade data. Additionally, this system would address the economical inefficiency of the translational research field of genomics where large amounts of identical data are generated repetitively. Study participants would instead have the ability to directly contribute research-grade genomic data from their own unique samples to future studies that they are qualified for. This is an approach that would reduce costs for future studies aiming to generate the same data,

potentially improving clinical outcomes for study participants. Moreover, this model provides an ethical framework for data sharing to advance research that bypasses multi-institutional regulatory barriers.

Future directions for validating and facilitating the return of research to participants include:

1. Increase engagement of study participants on IRB committees to represent public opinion and give input on expectations, risks and benefits for ethical decision making [31, 34].
2. Develop educational resources and dynamic consent models to provide study participants with a solid understanding of the benefits and risks associated with their participation in the genomic research study. This can ensure true informed consent and decrease the chances for conflict with the study participant's future perspectives.
3. Review current IRB-approved study protocols and consenting documentation to identify best practices on how to facilitate the communication of research-grade genomic data to study participants.
4. Conduct research to examine how the return of research-grade genomic data might impact study participant clinical outcomes, psychological outcomes, and willingness to participate in future research.
5. Create a Standard Operating Procedure for genomic data sharing, between researchers and study participants, in order to share best practices for the return of research.
6. Make the process for requesting genomic data clearly known to study participants and clarify who in each research institution is responsible for responding to such requests.
7. Amend federal laws to provide clearer guidelines for the return of research and more acknowledgement of public opinion.

Acknowledgments

We would like to acknowledge Bryce Olson for inspiring this work, Leisa Sutton and Katherine Vaughan for their edits, Georgia Sadler for her mentorship, and also Elvia Nunez for her help in publishing this work.

ORCID iDs

Peter Kuhn https://orcid.org/0000-0003-2629-4505
Ida Deichaite https://orcid.org/0000-0001-8839-0573

References

[1] Quintana Y 2015 Challenges to implementation of global translational collaboration platforms *MOJ Proteom. Bioinform.* **2** 65
[2] Montgomery J 2017 Data sharing and the idea of ownership *New Bioeth.* **23** 81–6
[3] Commissioner of the Educational Materials 2018 Comparison of FDA and HHS human subject protection regulations www.fda.gov/scienceresearch/specialtopics/runningclinicaltrials/educationalmaterials/ucm112910.htm (Accessed: 28 June 2018)
[4] Bonamici S 2016 Text—H.R.34—114th Congress (2015-2016): 21st Century Cures Act www.congress.gov/bill/114th-congress/house-bill/34/text (Accessed: 20 August 2018)
[5] Manion F J, Robbins R J, Weems W A and Crowley R S 2009 Security and privacy requirements for a multi-institutional cancer research data grid: an interview-based study *BMC Med. Inform. Decis. Mak.* **9** 31
[6] McKinney R E *et al* 2015 Use of altered informed consent in pragmatic clinical research *Clin. Trials* **12** 494–502
[7] Compagnone C *et al* 2017 Past, present, and future of informed consent in pain and genomics research: challenges facing global medical community *Pain Pract.* **17** 8–15
[8] Lunshof J E, Chadwick R, Vorhaus D B and Church G M 2008 From genetic privacy to open consent *Nat. Rev. Genet.* **9** 406–11
[9] Khan A, Capps B J, Sum M Y, Kuswanto C N and Sim K 2014 Informed consent for human genetic and genomic studies: a systematic review *Clin. Genet.* **86** 199–206
[10] Caulfield T and Murdoch B 2017 Genes, cells, and biobanks: yes, there's still a consent problem *PLoS Biol.* **15** e2002654
[11] Gliwa C *et al* 2016 IRB perspectives on obligations to disclose genetic incidental findings to research participants *Genet. Med.* **18** 705–11
[12] Kraft S A *et al* 2016 A comparison of IRB and patient views on consent for research on medical practices *Clin. Trials* **13** 555–65
[13] Dankar F K, Ptitsyn A and Dankar S K 2018 The development of large-scale de-identified biomedical databases in the age of genomics—principles and challenges *Hum. Genomics* **12** 19
[14] Hylind R, Smith M, Rasmussen-Torvik L and Aufox S 2018 Great expectations: patient perspectives and anticipated utility of non-diagnostic genomic-sequencing results *J. Community Genet.* **9** 19–26
[15] Schadt E E 2012 The changing privacy landscape in the era of big data *Mol. Syst. Biol.* **8**
[16] Wolf S M *et al* 2018 Navigating the research–clinical interface in genomic medicine: analysis from the CSER Consortium *Genet. Med.* **20** 545–53
[17] Foster M W and Sharp R R 2007 Share and share alike: deciding how to distribute the scientific and social benefits of genomic data *Nat. Rev. Genet.* **8** 633
[18] Fossey R *et al* 2018 Ethical considerations related to return of results from genomic medicine projects: the eMERGE network (phase III) experience *J. Pers. Med.* **8**
[19] Green R C *et al* 2013 ACMG recommendations for reporting of incidental findings in clinical exome and genome sequencing *Genet. Med.* **15** 565
[20] Pereira S, Oliver Robinson J and McGuire A L 2016 Return of individual genomic research results: what do consent forms tell participants? *Eur. J. Hum. Genet.* **24** 1524–1529
[21] Henderson G E *et al* 2014 The challenge of informed consent and return of results in translational genomics: empirical analysis and recommendations *J. Law Med. Ethics* **42** 344–55
[22] Klitzman R, Buquez B, Appelbaum P S, Fyer A and Chung W K 2014 Processes and factors involved in decisions regarding return of incidental genomic findings in research *Genet. Med.* **16** 311–7
[23] Center for Devices and Radiological Health 2018 Premarket approval (PMA)—PMA clinical studies www.fda.gov/MedicalDevices/DeviceRegulationandGuidance/HowtoMarketYourDevice/PremarketSubmissions/PremarketApprovalPMA/ucm050419.htm (Accessed: 20 August 2018)
[24] Ayuso C, Millan J M and Dal-Re R 2015 Management and return of incidental genomic findings in clinical trials *Pharmacogenomics J.* **15** 1–5

[25] Evans B J 2014 The first amendment right to speak about the human genome *Univ. Pa. J. Const. Law* **16** 549–636
[26] Jarvik G P *et al* 2014 Return of genomic results to research participants: the floor, the ceiling, and the choices in between *Am. J. Hum. Genet.* **94** 818–26
[27] Whitley E A, Kanellopoulou N and Kaye J 2012 Consent and research governance in biobanks: evidence from focus groups with medical researchers *Public Health Genomics* **15** 232–42
[28] Ball M P *et al* 2014 Harvard personal genome project: lessons from participatory public research *Genome Med.* **6** 10
[29] Aungst H, Fishman J R and McGowan M L 2017 Participatory genomic research: ethical issues from the bottom up to the top down *Annu. Rev. Genomics Hum. Genet.* **18** 357–67
[30] Shelton A K, Freeman B D, Fish A F, Bachman J A and Richardson L I 2015 A computer-based education intervention to enhance surrogates' informed consent for genomics research *Am. J. Crit. Care* **24** 148–55
[31] Yoshizawa G, Sasongko T H, Ho C-H and Kato K 2017 Social and communicative functions of informed consent forms in East Asia and beyond *Frontiers Genet.* **8** 99
[32] Corpas M, Kovalevskaya N V, McMurray A and Nielsen F G G 2018 A FAIR guide for data providers to maximise sharing of human genomic data *PLoS Comput. Biol.* **14**
[33] Zarate O A *et al* 2016 Balancing benefits and risks of immortal data: participants' views of open consent in the personal genome project *Hastings Cent. Rep.* **46** 36–45
[34] Folayan M O, Oyedeji K S and Fatusi O A 2015 Community members' engagement with and involvement in genomic research: lessons to learn from the field *Dev. World Bioeth.* **15** 1–7