Thought Leader Perspectives on Participant Protections in Precision Medicine Research

Catherine M. Hammack, Kathleen M. Brelsford, and Laura M. Beskow

Precision medicine research is thought to offer unparalleled opportunities to address pressing questions related to health and disease, but large-scale data collection and wide-spread data sharing present challenges for protecting research participants.[1] Specifically, privacy concerns are magnified in the context of gene-environment interaction studies due to the nature of the data collected. The inherent identifiability of genomic data, as well as the depth and breadth of other data collected from mobile health devices and electronic health records, increases the risk of re-identification in such studies.[2] In addition, the longitudinal nature of these studies presents important considerations such as changing circumstances in individual participants' health, future research uses of stored materials, or even the socio-cultural context.

Privacy and confidentiality are central to individuals' decisions about taking part in research and concerns about these can be a main deterrent to participation.[3] Researchers must accurately describe to potential participants any reasonably foreseeable risks and procedures to maintain confidentiality.[4] To do so requires sufficient understanding of whether or not — and to what extent — participants are protected against the risks and potential harms associated with precision medicine research.

Catherine M. Hammack, M.A., J.D., *is an Associate in Health Policy in the Center for Biomedical Ethics & Society at Vanderbilt University Medical Center (Nashville, TN).* **Kathleen M. Brelsford, M.P.H., Ph.D.,** *is a Research Assistant Professor in the Center for Biomedical Ethics & Society at Vanderbilt University Medical Center (Nashville, TN).* **Laura M. Beskow, M.P.H., Ph.D.,** *is a Professor of Health Policy and Ann Geddes Stahlman Chair in Medical Ethics in the Center for Biomedical Ethics & Society at Vanderbilt University Medical Center (Nashville, TN).*

To this end, we conducted empirical research on the scope of confidentiality risks and protections applicable to precision medicine research, as well as how these are and should be described to prospective participants. Here we report key findings from in-depth interviews conducted with a diverse group of thought-leaders — prominent individuals at the forefront of precision medicine research who are uniquely positioned to identify critical issues in this swiftly changing environment. We focus on their perspectives on the protections associated with precision medicine research.

MATERIALS AND METHODS

Participants

We conducted semi-structured interviews with 60 nationally recognized thought leaders possessing a range of expertise and experience regarding confidentiality and genome research, including:

- ELSI research (ELSI): Scholars who study ethical, legal, and social issues in genome science
- Ethics (Ethics): *e.g.,* directors of centers for bioethics
- Federal government (Government): Individuals in relevant positions in the federal government
- Genome research (Research): Bench science and medical genomics researchers
- Health law (Law): *e.g.,* directors of centers for health law
- Historically-disadvantaged populations (Historically-Disadvantaged): Scholars who study issues related to Historically-Disadvantaged populations

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)
DOI: 10.1177/1073110519840493

- Human subjects protections (Human Subjects): *e.g.*, leaders of national organizations related to human subjects protections
- Informatics (Informatics): Bioinformatics, clinical and medical informatics experts
- Participant-centric approaches (Participant-Centric): Leaders in participant-centric approaches to research

We used a stratified purposive sampling approach to interview at least six thought leaders per stakeholder group, which is the minimum number of interviews expected to achieve saturation.[5] Prospective participants were identified based on leadership positions in prominent institutions, organizations, and studies, as well as authorship of highly influential papers on relevant topics. We used nominated expert sampling to identify additional thought leaders and further expand our sample.[6]

Instrument Development

We developed and piloted a semi-structured interview guide that included narrative and non-narrative elicitation techniques to explore confidentiality and privacy issues and solutions in gene-environment interaction research. To facilitate discussion, questions were framed around a hypothetical big data study — the "Million American Study"[7] (Box A) — that involved extensive characterization (including whole genome sequencing) of biospecimens, ongoing collection of information from electronic health records, and real time monitoring of lifestyle and behavioral information through mobile devices. Interview topics included, among other things, risks,[8] benefits and harms,[9] and the strengths and limitations of a range of general and specific approaches to protecting confidentiality. The final instrument (available upon request) consisted of 19 questions; here we report findings in responses to the following questions:

Box A

The Million American Study (a hypothetical scenario)

The Million American Study (MAS) is a large-scale research endeavor to improve understanding of health and to find new ways to predict, detect, diagnose, treat, and prevent disease. Specifically, the aim is to compile comprehensive information from a cohort of one million Americans in a repository that will serve as a rich research resource for a wide variety of studies for decades to come.

MAS will seek to enroll a representative sample of U.S. adults reflecting diversity in terms of race and ethnicity, age, and sex. Those who agree to participate will give consent for:

- Extensive characterization (including whole genome sequencing) of biospecimens, such as blood
- Ongoing access to clinical data (such as medications, test results, and imaging) from electronic health records
- Real-time monitoring of lifestyle and behavioral information, such as physical activity and environmental exposures, through mobile health devices

At the time of consent, participants will be offered choices about whether they are willing to be re-contacted for various purposes, for example to provide additional information or specimens, or to receive individual research results. Participants will be able to withdraw consent for future use of their specimens and data, with the exception that data generated in past studies cannot be withdrawn, nor can specimens and data be withdrawn from studies already begun.

Specimens and data will be stored in coded form in a federal repository. A robust data security framework will be in place, including administrative, technical, and physical safeguards. There will be a centralized governance process, comprising participant representatives, researchers, health care providers, government officials, and other stakeholders to ensure overall accountability and responsible project management.

Multiple tiers of access to MAS data — from open to controlled — based on data type, data use, and user qualifications will be employed. For example, certain information, such as some aggregate results, will be publicly available. Access to other information will be available to qualified researchers from academic, non-profit, and for-profit entities, in the U.S. and around the world, through application to a Data Access Committee. For approved projects, Material Transfer Agreements will be used to ensure that data and specimens are used and shared for authorized purposes only, and that privacy and security safeguards are maintained.

Information will be publicly available concerning how MAS cohort data and specimens are being used, including information about ongoing studies and summaries of research findings.

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

Imagine that your family members and close friends are all at a gathering together. The conversation turns to the "Million American Study" that has been in the news recently. Everyone is eager to hear your thoughts about whether they should consider signing up to be in this study.

…

8. Imagine now that your family and friends ask about the kinds of confidentiality **protections** that would be in place for the Million American Study to minimize the risks and potential for harm. What would you tell them about the usefulness, strengths, and limitations of the following general approaches:
   a. Technical **data security** measures (*e.g.,* computer passwords, encryption, audit trails)
   b. Laws/rules/procedures intended to **restrict access** to research data (*e.g.,* data access committees, Certificates of Confidentiality)
   c. Laws/rules/procedures intended to **prevent misuse** of research data (*e.g.,* data use agreements, anti-discrimination laws)
9. Given your understanding of their strengths and limitations, how reassured do you think your family and friends should be by each of the following protections, and why?
   a. How would you rate* the level of reassurance your family and friends should feel based on the **Common Rule** requirements for consent and IRB oversight?
   b. How would you rate* the level of reassurance your family and friends should feel based on the **HIPAA Privacy Rule**?
   c. How would you rate* the level of reassurance your family and friends should feel based on **GINA**?
10. Are there other specific protections that you think should be reassuring to family/friends? If so, please describe.

** On a 5-point rating scale from "Not at all reassured" → "Very reassured"*

The Duke University Health System and the Vanderbilt University Medical Center Institutional Review Boards deemed this research exempt under 45 CFR 46.101(b)(2) (2009).

Data Collection

We emailed prospective interviewees an invitation to participate and a study information sheet. Before the interview, participants received additional study information, including a description of the "Million American Study" and an outline of the interview topics.

Interviews were conducted between September 2015 and July 2016. All interviews were conducted by telephone by three members of the research team. At the beginning of each interview, we reviewed the study information sheet and obtained the participant's verbal agreement to participate. With participants' permission, interviews were audio-recorded and professionally transcribed. Interviews ranged from 30 to 120 minutes in length, with an average length of approximately one hour. Participants were offered $100 compensation for their time.

Data Analysis

Transcribed interviews were uploaded into qualitative research software NVivo 11 and a standardized iterative process was used to develop a codebook.[10] Specifically, two team members first created a structural codebook to index interview questions and corresponding responses based on the interview guide. They then independently reviewed four transcripts to identify substantive content for inclusion in an initial codebook of thematic or content codes which they jointly developed and refined. Next, they independently applied codes to a fifth transcript and then compared the results to revise codes and code definitions as needed. They followed this iterative process with additional transcripts until they achieved at least 80% inter-coder agreement in code application. The remaining transcripts were then divided between the two coders; each independently coded every sixth interview to ensure inter-coder agreement remained at a minimum of 80%.

Once all data were coded, the team systematically generated narrative summaries of relevant codes to explore the range of thematic responses and to identify additional sub-themes.[11] Narrative summaries were reviewed by at least one other team member, who read the corresponding NVivo code reports to identify and confirm agreement in sub-theme identification and the synthesis itself.

RESULTS

Participant Characteristics

We interviewed 60 thought leaders, representing a wide array of perspectives and demographic diversity (Table 1).

Views on General Approaches to Protecting Confidentiality in Precision Medicine Research

We asked thought leaders to discuss the usefulness, strengths, and limitations of three general approaches to protecting confidentiality: technical data security

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

Table 1
Participant Characteristics (n = 60)

	n	(%)
Perspective		
ELSI research	6	(10.0)
Ethics	7	(11.7)
Federal government	7	(11.7)
Genome research	7	(11.7)
Health law	6	(10.0)
Historically-disadvantaged populations	7	(11.7)
Human subjects protections	7	(11.7)
Informatics	6	(10.0)
Participant-centric approaches	7	(11.7)
Gender		
Female	31	(51.7)
Male	29	(48.3)
Race		
American Indian or Alaska Native	2	(3.3)
Asian	5	(8.3)
Black or African American	3	(5.0)
Native Hawaiian or Other Pacific Islander	1	(1.7)
White	49	(81.7)
Ethnicity		
Hispanic or Latino	2	(3.3)

measures, procedures intended to restrict access to data, and procedures intended to prevent misuse of data.

Technical Data Security Measures
Nearly all interviewees characterized technical data security measures — such as computer passwords, encryption, and audit trails — as important and necessary, although not sufficient:

> I don't think there's anything that can't be hacked, but I would think those things would be very important; that there is definitely encryption and all of those different security things — absolutely. (50, Human Subjects)

Only a few interviewees described such measures as providing strong or very strong protection for our Million American Study. Of these, some favorably compared their effectiveness to that in a clinical context, with one opining that technical measures "will be protective at levels that exceed all of the levels that your clinical information typically is safeguarded at" (53, Informatics). Another drew positive comparisons to commercial contexts:

> The mechanisms available to a study like this are very comprehensive, and if well-governed, will be extraordinarily secure, and will probably meet or beat anything that we see happening in the consumer financial — you know, credit card support or online or social media world, by an order of magnitude, if they are maintained and if they are sustained and tested. (03, Informatics)

Others cited the historical success of data security measures:

> Folks are working every day to make [technical measures] better. And particularly in the research realm, there isn't a long harrowing history of misuse in that context. There's no Edward Snowden of health records that I'm aware of. (40, Human Subjects)

Only a few interviewees argued the opposite, saying that the protection provided by technical security measures is weak:

> Given that every other day in the news you hear about the federal government releasing social security data, the chances of [a breach] happening are very, very high. I would tell family that it's probably going to occur … I'm sure [hackers] would be just dying to get into this database. So passwords, encryption, means nothing anymore when it comes to 12-year-olds who have figured out how to get into every system in the world. Nothing is infallible. (36, Historically-Disadvantaged)

The vast majority of thought leaders recognized both the strengths and weaknesses of technical measures, characterizing them as 'good but not perfect':

> I would say that nothing is perfectly safe. There are no foolproof systems. There is always going to be some sort of residual risk. And if you can't tolerate that, then don't be part of it. But the risk is likely to be very small. (12, Law)

> I think that technology is good. It's not perfect. I think it's a reasonable risk. We trust it every day in other parts of our lives, so I would be supportive of it, but make sure people realize it's not a guarantee. (34, Human Subjects)

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

Regardless of how strong or weak interviewees perceived these protections to be, most recognized two major limitations. First, data security measures serve as a barrier, but primarily identify problems in hindsight; "Audit trails are ex-post-facto … If stuff already went out, that's not protective to me. That's protective in terms of making it a better system." (34, Human Subjects)

Second, they described limitations imposed by human behavior, saying that "the human is always the weak link" (41, Ethics). Observations concerning human action extended from the skill and motivation of hackers …

> There's a lot of protections that are currently in place or currently being thought of, but I also think that there's more and more sophisticated ways of hacking. (30, Historically-Disadvantaged)

> These measures … are important but imperfect. Given the right motivation, people can get around anything. (08, ELSI)

…to the fallibility of those tasked with implementing or adhering to data security procedures:

> Is it foolproof? No, there are always students or interns or people that are not – it's not a foolproof. People are sloppy. (05, Research)

> The limitations of course are that you can't control human behavior, and so most of the security breaches that we hear about deal with people transferring data to unencrypted environments or losing laptops or things like that. (15, Ethics)

> I would also emphasize to family and friends that it always comes down to individuals. And that individuals, either for reasons of sloppiness or because of intentional disregard for the rules, are probably the weakest link in any sort of a security setup. That you can have the best security and firewalls and everything else in the world, but if somebody downloads information onto their laptop or their flash drive and then leaves it sitting around … Almost all the data breaches that have occurred have occurred because of that type of poor human factor. (58, Research)

Some noted particular concern given the context of widespread data sharing: "It's really hard to make and keep promises about [things like] audit trails if you're going to share the data." (23, Participant-Centric)

Thought leaders' comments included a few proposed solutions with regard to data security, such as issuing "stand-alone, encrypted, tamper-proof, anonymized" devices, rather than transmitting real-time data from mobile devices (42, Law); decentralization, so that data are stored "on a bunch of different servers" (19, Ethics); and removing identifiers "into their own silos that require separate access" to reduce the risk of inadvertent exposure (57, Informatics). More generally, they hoped that those planning the study "would be giving a lot of attention to those kinds of issues" (26, Human Subjects) and seeking out gold-standard techniques "that get you a lot of bang for your buck" (51, Informatics). One suggested including expertise from those outside the usual sphere of biomedical research:

> I'll use the example of Lockheed Martin — they've been keeping national secrets for a number of years … I really think there are probably other groups out there that can prevent breaches and do better than people we've got working in the space traditionally. (06, Participant-Centric)

Restricting Access to Research Data

Thought leaders commonly characterized laws, rules, and procedures intended to restrict access to research data — such as data access committees and Certificates of Confidentiality — as "another layer of protection" (10, Historically-Disadvantaged). Perceived benefits included screening prospective users "to ensure legitimate access" (08, ELSI) and to "keep out the bad actors" (24, Government). Interviewees also recognized, however, the potential for cumbersome bureaucracy and procedures that could impede science:

> I think in some cases they are overly restrictive. They impose a substantial burden on researchers and it makes it harder to actually do the science when you have to jump through all of these hoops and check all of these boxes and fill out all this paperwork. So it's a double-edged sword. (43, Participant-Centric)

In terms of the protections afforded by restricting access to data, thought leaders typically described them either as weak …

> I don't think that works very well. It doesn't make me feel any better. I mean, I don't have a lot of faith in that. (06, Participant-Centric)
> They're not nearly as effective as the right technical things … You're dealing with a gigantic study

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

> with zillions of people, so the biggest Achilles heel is who has access and how you do that. (44, Research)

...or perhaps useful but no panacea:

> I think they reduce the likelihood of [bad] things happening compared to not have such laws and rules in place. But there are no guarantees. (20, Research)

> They do protect people's information to an extent. None of them are foolproof. (27, Government)

> It's a patchwork that has been fairly thoughtfully put together but still has gaps. (55, Government)

With regard to data access committees in particular, some interviewees did not view them as protecting confidentiality, but rather as guarding "more against stigmatization, community harms, those types of things" (50, Human Subjects). Another noted that the level of protection provided depends on the quality of the committee:

> I think you can have poorly operating data access committees that are not that careful, and they can greatly increase the risk. Or you can have really excellent data access committees who can really provide another strong link in ... using [information] in the proper ways and not using it in improper ways. (42, Ethics)

With regard to Certificates of Confidentiality specifically, thought leaders described them as "an extra step, but I don't think it's any guarantee" (50, Human Subjects), noting in particular that "their legal effect is really unclear and they've not been fully litigated before, so it would be better than nothing but it's not perfect" (29, Human Subjects). Interviewees were especially uncertain of Certificates' protections in the context of multi-site research:

> Part of the difficulty is how Certificates travel with data sets and who is enforcing them. There is some worry given how fluid the circuits of information flow are these days. But often times, promises of Certificates don't travel well. Not because of bad intentions, but there are just limitations in trying to keep track of what people want and what people are intending to do with data. (32, Ethics)

Overall, thought leaders highlighted several limitations associated with the various approaches intended to restrict access to research data. They again commonly referenced human behavior as a major shortcoming, observing that "people are not perfect" (16, Ethics) and measures are "only as good as the people applying them" (26, Human Subjects). They also frequently discussed limitations related to monitoring and enforcement. As one stated, procedures to restrict access "are effectively useless unless you can guarantee compliance" (48, Law). Finally, some described limitations associated with delegated decision making (i.e., entities making decisions about data access on behalf of research participants), noting, for example, that data access committees would not always "make all the same decisions that you might make if you, the individual, were making them." (21, Government)

A common refrain was the need for balanced approaches that protect the data without unduly hindering beneficial research:

> If you restrict things and put in lots and lots of steps to restrict the data, to make sure it's safe and so forth, you're also going to kill 90% of the science. You want to know ... that they've actually balanced the protection with the goal of the database [itself]. (02, Government)

Preventing Misuse of Research Data

We also asked thought leaders about the usefulness of laws, rules, and procedures intended to prevent misuse of research data, such as data use agreements and anti-discrimination laws. A few perceived the prospect of data misuse as both real and likely: "Somebody will do it — whether intentionally or not, it will happen" (06, Participant-Centric). More commonly, however, interviewees suggested that the likelihood of tangible harm was low or theoretical:

> Misuses of data ... it is difficult to point to very clear examples that would be relevant to this. (03, Informatics)

> As far as I'm aware ... no research participant has ever been discriminated against because of participation in a genetic research study. People are actually more likely to die in car crashes going back and forth from the medical center. (20, Research)

In particular, some expressed doubt about intent or motive to target specific individuals:

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

> For the most part, nobody's going to care enough to go after you … If you were Bill Gates, that would be different, or Hillary Clinton. But my family and friends are not Bill Gates and Hillary Clinton or Oprah Winfrey or Michael Jordan or name your celebrity. Being an ordinary person has its advantages. (14, Law)

Irrespective of thought leaders' opinions concerning likelihood of misuse, the importance of addressing it was a prominent theme, often described in contrast to efforts to restrict access:

> Measures that work not to restrict access but to punish wrongdoers, to punish people who misuse information or who obtain access illicitly, those are the kinds of laws that I would prefer to see on the books because they don't interfere with appropriate data access. (24, Government)

> I think the one area that we could do a whole lot better of a job in is making it illegal to use data in ways that it was not intended … We ought to be more concerned about punishing people who do inappropriate things with the data rather than restricting access. (59, Government)

With regard to the strength of the protections afforded in actual practice, however, interviewees frequently observed that measures to prevent misuse rely on trust. For some, this was accompanied by skepticism:

> For people deterred by laws, these are effective. They don't make much difference for those who don't care [about laws]. (08, ELSI)

> [Measures to prevent misuse] are well-intended, but in the end, pretty useless. It has a high symbolic value, and it represents what we as a society feel should be norms that we have in common. But in terms of the real ability to prevent misuse, I'm very skeptical about it. I think it does not prevent people with bad intentions from doing things. (45, Participant-Centric)

With regard to data use agreements in particular, thought leaders found benefit in the opportunity to set expectations, and "to disclose to folks who are accessing [data]: what the implications are for misuse of the information" (01, Human Subjects):

> I think in their implementation, [data use agreements] keep people aware of just how important it is to treat the data with respect. They're incredibly bureaucratically cumbersome, so they remind us all the time just how important it is to follow the procedures and that not everybody can have the data. You have to show that you have the capability to keep it secure and that you've got a reason for having it and that it's not to be shared. (05, Research)

Some interviewees were generally comfortable relying on researchers to keep these agreements, noting that "there's the small percentage of researchers who are going to violate those types of rules and laws" (30, Historically-Disadvantaged) and that "investigators are very, very, very rarely interested in the private information of individuals, and they generally can be trusted as long as they are properly vetted" (42, Ethics). Others, however, were less comfortable:

> I think scientists have gotten way too accustomed to downloading the data, and pretending that there are no restrictions on it. (23, Participant-Centric)

> People sign their agreement to follow a certain set of rules and to abide by the agreement, but those are only as good as the people who sign them. (27, Government)

> My experience has been that data use agreements are rarely really verified or enforced. And by that … I mean the people who sign the data use agreement do not verify that the recipient is actually doing with the data what they intended or what they said. (55, Government)

With regard to anti-discrimination laws in particular, some thought leaders found the risk of discrimination largely hypothetical:

> When GINA was being written, people had a really hard time finding good examples of nefarious uses. So I would say it's there to protect but there's not a huge experiential literature on how this stuff has been inappropriately used. So while I don't think the laws are as strong as they potentially could be, I think they're protecting for very rare occurrences. (34, Human Subjects)

In general, however, interviewees' sentiments suggested that "the fact that we have GINA is a good thing" (10, Historically-Disadvantaged), but with notable concerns. First, many observed gaps in the protections provided:

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

> I would say that the non-discrimination laws are very weak … and that even though we have GINA, many people know that it has a number of loopholes in it. So I think I would tell [family and friends] not to hang too much on non-discrimination laws, but really depend more on security. (42, Ethics)

Second, some mentioned that focusing narrowly on the misuse of genetic information propagates notions of genetic exceptionalism:

> We have such a fear-mongering mentality when it comes to genetic data. People just look at genetics differently, and there are reasons for that. Genetic information is not just individual. It's about your family, your inherited line — I appreciate that there are risks with genetics that don't exist with other kinds of information. But to say that means that genetic research should be viewed through a different lens is a mistake, overkill for the situation. (24, Government)

Overall, thought leaders described a number of limitations to measures intended to prevent misuse. First and foremost was concern about enforcement, with many expressing that such measures "don't have any teeth" (06, Participant-Centric) and that "there are no real enforcement mechanisms [and] no clear penalties" (15, Ethics):

> Everyone hopes that all the researchers are just wonderful people and they will abide by 'please don't do that', but you're really banking on people adhering to things just because they feel like they should or because they have good morals, rather than it having any consequences. (35, Participant-Centric)

Some interviewees opined that attempts to prevent misuse are limited by being primarily reactive; for example, describing anti-discrimination laws as "probably helpful, but they don't keep information from getting out" (18, Law) and "additional protection if their information is out there floating around" (58, Research).

A few mentioned barriers to pursuing the penalties that do exist. As one explained, "It's hopefully a deterrent to somebody, but once it happens, now the burden's on me to wield the stick and penalize somebody" (26, Human Subjects).

Finally, some interviewees noted that attempts to prevent misuse are limited by the patchwork of laws — at both state and international levels — that create inconsistencies and gaps in protections:

> In the United States it's a bit crazy because so many laws are state-based, so it's not equal. We have the general GINA law, but as far as information be used for seeking for life insurance or long-term care, that's state-by-state. (50, Human Subjects)

> Well, there's Henrietta Lacks and her genome and her family members. I think that the publication of her genome on the Internet without her family's consent shows you how there are gaps in the law considering the international scale of data sharing and genomic research. (10, Historically-Disadvantaged)

In the end, calls for stronger enforcement and greater penalties were common across thought leaders' discussions of efforts to prevent misuse:

> I think the biggest intervention that we can make is to make things really egregiously illegal and to really penalize misuse with fines and other kinds of punishments … I would like to see that whole landscape shift and maybe we're moving in that direction but we still have a long way to go. (32, Ethics)

> So if you look at HIPAA as an example, the consequences to individuals and to institutions of data breaches under HIPAA are really significant. Whereas I don't think we've seen that type of enforcement on the research side. And I think we need to move in that direction. Because if there's not a significant penalty to individuals or institutions, it might lead some people that're a little bit morally shaky to say, "Well, the consequences aren't that big, and this is an important question. I'm just going to go ahead and do it." That's I think where we have a mismatch right now. (58, Research)

Reassurance Provided by Specific Protections in Precision Medicine Research

In addition to assessing general approaches to protecting confidentiality, we asked thought leaders to rate how reassured people should be by three specific legal protections: the Common Rule, the HIPAA Privacy Rule, and GINA.

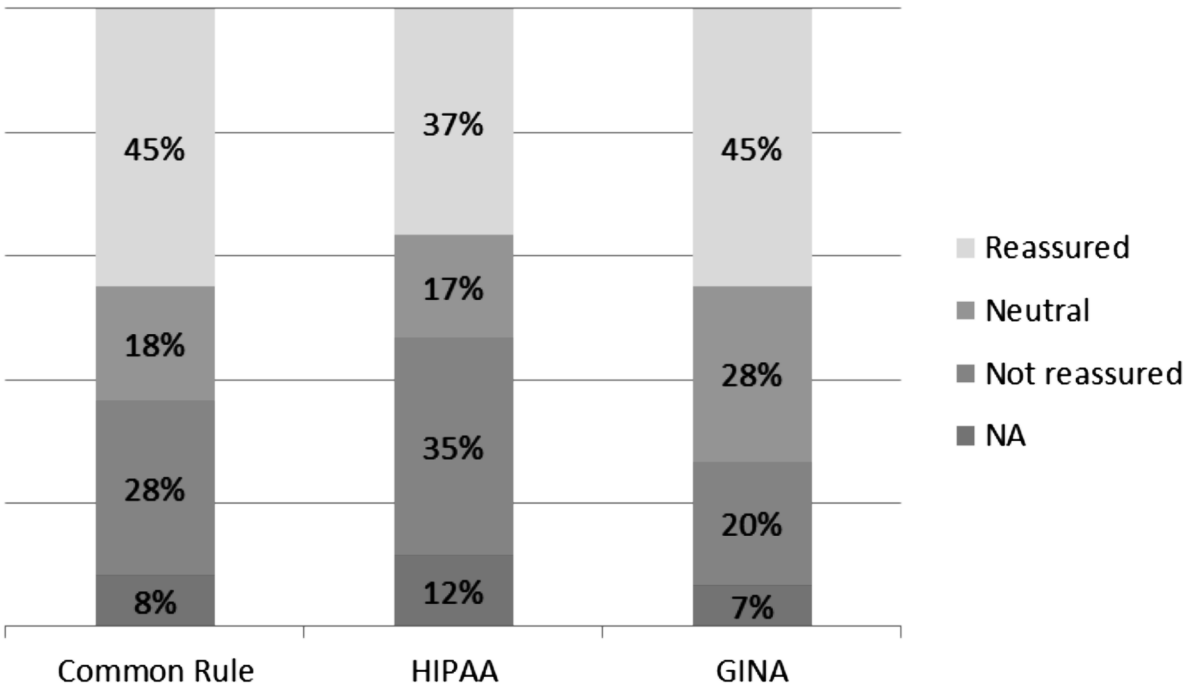
The Common Rule

Nearly half (45%) of all thought leaders believed their family and friends should be reassured by the Common Rule's requirements for consent and IRB oversight (Figure 1). They generally cited the Common

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

Figure 1

Reassurance Afforded by Specific Protections (n = 60)



Asked on a 5-point scale: 1=Not at all reassured, 2=Not too reassured (combined here to "Not reassured); 3=Neither reassured nor not reassured (labeled here as "Neutral"); 4=Reassured, 5=Very reassured (combined here to "Reassured")

NA = No answer (interviewee not able to assign numeric rating)

Rule's scope and purpose as a reason for reassurance, particularly insofar as "it is a system that is based very firmly in respecting autonomy and individual privacy" (03, Informatics). Indeed, one thought leader noted that this focus could sometimes come at a cost to the study:

> It tends to be overly restrictive … [IRBs] tend to be risk averse. So they tend to be beneficial to the individuals, rather than to the overall group or the overall goals of the study. They're much more likely to be risk averse at the expense of the study rather than be sort of less risk averse at the expense of the participants in the study. (48, Law)

About one fifth (18%) of interviewees said their family and friends should feel neutral — neither reassured nor not reassured — with respect to the Common Rule's protections, often citing variability between IRBs:

> Some IRBs will approve a chicken sandwich and their oversight is not meaningful. And there are other IRBs whose oversight is very tough … I see widely varying degrees of IRB sophistication and oversight and monitoring effectiveness. (29, Human Subjects)

> They should be really reassured that a really good IRB with a high-profile study will do a good job of doing what it can do. I guess they shouldn't think that it can do things that it can't do, though. It can only do what it's supposed to do. For the average IRB reviewing the average study involving genetic data and other data, that I would be much less reassured by. Not every IRB understands genomic data, understands data security protocols. Some of them are too permissive. Some of them are way too restrictive. There's variability all over the place. It's almost a crapshoot. (54, Ethics)

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

More than one-fourth (28%) of interviewees felt their family and friends should not be reassured by the Common Rule's requirements. Many referenced informed consent as a particular weakness:

> If we're just talking about consent requirements, I would say "zero." You should not be reassured at all because the consent requirements under the Common Rule can often just be addressed by: you have a consent form and you sign it. (19, Ethics)

Another weakness was relatively low standards regarding identifiability, "with the Common Rule standard … being 'readily identifiable to the investigator,' whereas HIPAA has the 18 criteria that need to be stripped in order for it to be considered de-identified." (01, Human Subjects)

Finally, some questioned the ongoing influence of the Common Rule for a long-term study like the Million American Study:

> You've already pretty much exhausted the requirements of the Common Rule by signing up for the study. Once you're in the study, the Common Rule does relatively little to protect the data that's gathered or the information that's acquired about you. (37, ELSI)

The HIPAA Privacy Rule

Over one-third (37%) of thought leaders felt their family and friends should be reassured by the HIPAA Privacy Rule (HIPAA) (Figure 1). Among these interviewees, some felt that HIPAA had generally been effective in raising awareness and setting expectations for the confidential handling of health data, making people "broadly cognizant of the standards and appropriately concerned about maintaining compliance" (01, Human Subjects). They also perceived HIPAA's criteria for de-identification to be strong:

> It does not prevent re-identification in all aspects and computational ways. But what it does do is make it hard, and it reduces the probability dramatically. (53, Informatics)

Some highlighted the threat of serious penalties that help deter HIPAA violations. As one interviewee noted:

> That's a pretty good rule. The only problem with that in my view is it doesn't give a private right of action, so it's up the Office of Civil Rights to bring in enforcement measures. But they have done a pretty good job in that area I think. (04, Law)

However, even interviewees who rated HIPAA as reassuring recognized that the strength of its protections could have both positive and negative implications for researchers and patients:

> It kind of stifles a little bit of research and sharing. But it also is designed to help protect. So it's a double-edged sword. (38, Government)

> At least it's terrified everybody who has anything to do with it. And so [family and friends] should feel pretty protected and in fact worried they're a little too protected so that they themselves can't get access to what they need. (35, Participant-Centric)

HIPAA's positive aspects notwithstanding, more than half of thought leaders believed their family and friends should feel either neutral (17%) or not reassured (35%). The scope and limits of HIPAA protections was an area that these interviewees recognized as ripe for potential misunderstanding and false reassurance:

> HIPAA is in many ways reassuring to many folks. I think that most people don't quite understand [the] rules around "covered entity" and what actually falls under that rubric. Most people assume that things are protected when they're not. (32, Ethics)

Many of these interviewees pointed out that HIPAA's primary focus is not research: "HIPAA really is not a research regulation; it's a consumer protection act for healthcare in general, and research kind of came along for the ride" (26, Human Subjects). Thus, they cautioned against relying on HIPAA once data are moved into the research domain:

> I'm not sure once you put PHI [protected health information], an electronic health record, into a [research] repository, whether HIPAA applies anymore. HIPAA has to do with getting access to it but you've given permission. (22, ELSI)

> HIPAA addresses clinical data, but a researcher isn't a HIPAA-covered entity, and … we're really not talking about clinical data in a research environment. Once the clinical data get there, they aren't governed by HIPAA. So that shouldn't really give them much assurance at all. (51, Informatics)

Some saw even the protections offered by meeting HIPAA de-identification standards as limited, given

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

the possibility of triangulating among several sources of complex data:

> There's always somebody left in the data set that's identifiable, just because of statistics. We have no way of knowing in advance whether you're the outlier who's identifiable in any given slice of data, and because this data's going to be sliceable, we're all going to eventually be the identifiable person in some slice somewhere. (23, Participant-Centric)

> So here's one of the problems. There are all kinds of data. It's not clear how they're going to be kept or by whom. Different researchers will have different kinds of access to the data or will access different elements of the data, right? So if researchers are accessing robust EHR information ... they should probably tell you that that's not possible to de-identify under HIPAA. (19, Ethics)

The Genetic Information Non-discrimination Act (GINA)

Like the Common Rule, nearly half (45%) of thought leaders believed that their family and friends should be reassured by GINA (Figure 1). Some of these interviewees perceived the risk of genetic discrimination as mostly theoretical: "I don't actually think that there are that many employers or insurers who want to discriminate on this basis" (14, Law). Others gained confidence from a perception of GINA as covering what matters most — "the most likely areas where you could get discriminated against is employment and health insurance" (51, Informatics) — as well as being effective in practice:

> We actually have looked [at] how many cases have come through the EEOC [Equal Employment Opportunity Commission] and there are almost none have gone to court. There certainly have been a few that have been adjudicated ahead of time. So they should be pretty confident that GINA works. (35, Participant-Centric)

However, even among interviewees who felt their family and friends should be reassured, several acknowledged the gaps in protection ...

> GINA will be quite effective about protecting you against a certain kind of misuse of the information, and that's misuse by your employer-sponsored insurance plans, your health benefits plans. So that's great ... But it doesn't keep your information from being misused by racists or used in other ways that you might find offensive. (37, ELSI)

... as well as redundancies in coverage:

> I think GINA's fine, but I don't think GINA actually does anything. I think most of the same protections are already within the Affordable Care Act and the Americans with Disabilities Act. (13, Informatics)

Compared to those who found reassurance in GINA's protections, a slightly greater proportion of thought leaders believed that their family and friends should feel either neutral (28%) or not reassured (20%). Their sentiments reflected themes mentioned earlier when we asked about general measures intended to prevent misuse of research data, including gaps in coverage, redundancies in protection, and enforcement challenges. With regard to the latter, interviewees especially highlighted the difficulty of people proving — or perhaps even knowing — they have been discriminated against based on genetic information:

> The bad thing is sometimes it's just hard to prove discrimination based on GINA. If someone goes to a database ... and looks at your DNA and they decide to fire you without giving you why they did that, then okay, how can you even tell? (07, Research)

These concerns led thought leaders who were less reassured by GINA to describe it using terms ranging from "aspirational" (41, Ethics) to "misleading" (04, Law). One suggested that GINA may actually cause harm because it legitimizes exceptionalist ideas:

> [There is] a misunderstanding of the value of genetics and, ironically, what GINA did is legitimize this exception. 'Genetic information is special, so special that we need to have this new law even though it's not very penetrant and most of this information isn't very particular at all, we're going to create a special law for it.' You can see that it's hard to answer: 'How reassured should they be that mostly useless information may get disclosed and ignored by insurance companies?' I don't know. (28, Law)

Views on Other Protections and Solutions

Thought leaders identified several additional protections and solutions potentially applicable to large scale gene-environment interaction studies, including

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

the Americans with Disabilities Act and, in particular, the Affordable Care Act:

> There are two other things I think make a big difference. One is the reinvigoration of the American Disability Act … but I think real biggie is the Affordable Care Act, which, by eliminating pre-existing conditions, really makes access to healthcare coverage more available. These are complicated questions, but there are policy moves that can be, that have already been made and policy moves that can be made in the future that will be helpful. (04, Law)

> Laws that require health care for everyone regardless of their genetics makes it okay for people to want to hand over their DNA because if we know that there's a law that says you can't discriminate against me, then it makes me more likely to want to contribute my DNA. So those laws make it possible for people to feel a little bit safer. Even if they may not always work, it's still a good thing. (10, Historically-Disadvantaged)

> I think if your worry is we uncover something that becomes a pre-existing condition and therefore I'm worried that I won't be able to get health insurance, well the Affordable Care Act mostly solves that problem. (16, Ethics)

Interviewees mentioned other specific laws, regulations, and guidelines, such as Fair Information Principles, Federal Trade Commission security regulations, the HIPAA Security Rule, and professional codes of conduct and licensure standards. One interviewee referenced state laws as a potential source of additional protection:

> This area of the law is a creature of state statute and state legislation. So, there are lots of state laws out there that deal with this stuff, depending on what state you're in, some more than others. (24, Government)

The importance of flexible and adaptable oversight and governance was another common theme:

> Technology evolves, consequences of genomic research is evolving. We don't know all the full implications, the benefits and the risks … New issues will be constantly coming up that could benefit or threaten my privacy or my future and welfare. So there's going to have to be continuous mechanisms to adapt to these new developments. (02, Government)

> The real issues are the governance of those research databases. That's where the real action is in the policy world, not in the consent phase with individual participants … I'm much more interested in, for example, figuring out some kind of a community advisory board, some kind of mechanism for people who can really spend the time to think of the tradeoffs and make decisions, rather than going back to individuals. (41, Ethics)

Thought leaders highlighted the engagement of relevant communities and involvement of participants in governance processes:

> If there was true community engagement and participant engagement in the leadership … then there should be some openness to that being a more trustworthy environment. It needs to have governance that reflects the constituency that it would support. (03, Informatics)

> I think governance and having an IRB that's reflective of participants and not just researcher-driven. Because I think there are a lot of things you can do through the participant engagement side of things that build trust and make sure that policies are the right policies. (06, Participant-Centric)

> One of the interesting features of the [Million American Study] is having representatives from the cohort, actual research subjects, participating in the governance structure. And so having people like us on the data access committees and data use committees—that might be reassuring to the public: to think that we've had a hand in the game … Once people get onto the inside and have a sense of what kind of information is being produced and where it's going, their concerns will be heightened, and it will be easier to put in place protections and policies to protect that information. (37, ELSI)

Some, however, expressed concerns about the long-term effectiveness of such structures:

> In my experience, committees tend to be very interested at the beginning and they start to become less and less interested over time. They get distracted by other projects. They get

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

> distracted by compensation. The people that were initially involved move on and other people replace them. This is intended to be a long-term study and my experience with committees and committee structure is they start out great and they tend to mature into being not very active and not very, not containing the kinds of thought and deliberation and conscious affirmative action that was there in the beginning. (17, Participant-Centric)

A few interviewees pointed to the study having the necessary infrastructure and support as a source of reassurance: "For me to sign up for this, I would really want to know is there, in fact, a robust infrastructure for doing this right?" (29, Human Subjects). For many, 'doing this right' included the use of a combination of approaches to protect participants: "The strongest strategy is to have both rules and procedures and policies on the one hand, and technical controls on the other" (16, Ethics).

Finally, thought leaders highlighted the crucial role of transparency and trust. They noted, for example, that "a requirement and expectation for transparency would also be reassuring" (21, Government), and urged clear explanations of the risks, as well as the strengths and limitations of available protections, as part of consent processes: "And if those risks scare people off, then those aren't the people who should be participating in something like this" (36, Historically-Disadvantaged). Some suggested that researchers should enroll as participants, thus having 'skin in the game':

> ... so we are exposed to similar risks as the people who participate in our study. Just something to signal that we are together—it's not anymore like researchers and human subjects, we're just participants, everyone is on the same level. (07, Research)

Interviewees emphasized that participants would need to be able to trust in scientific integrity as "a real, a kind of protection that exists" (28, Law), noting that "the ultimate reassurance is knowing about the research and having confidence in the integrity of the research process" (08, ELSI). Such foundational trust would be vital for endeavors like the Million American Study because, for participants:

> You need to be comfortable with unknown and in many ways unknowable risks. As well as unknown and unknowable benefits. The reason you're participating in the study is the same reason that people got on wagon trails in the West—there's some sort of pull to the unknown. We don't want to pretend that we can quantify or describe all of these things, or that we're protected from all these things. This is an inherently risky study, and part of the risk is that we don't know how risky it is. (23, Participant-Centric)

Discussion

Precision medicine research is rapidly taking a lead role in the pursuit of new ways to improve health and prevent disease; this is perhaps best exemplified by the recent launch of the *All of Us* Research Program, an unprecedented endeavor aiming to collect genomic, clinical, and lifestyle information from one million individuals.[12] However, these kinds of large-scale gene-environment interaction studies raise privacy and confidentiality concerns due to the identifiability of data and longitudinal nature of such studies. The success of such research depends on understanding the "web" of laws, regulations, policies, and procedures in place in order to elucidate how the risks should be explained, the extent to which participants are protected, and what more could be done to safeguard privacy and confidentiality.

Our study sought to address these issues by eliciting the perceptions and opinions of experts at the forefront of precision medicine research. In general, our interviewees agreed that all technical, legal, and regulatory restrictions on data access and use are subject to limitations which can affect the likelihood and consequences of risks associated with participation. Specifically, thought leaders described technical data security measures as necessary but insufficient due to challenges in human involvement and widespread data sharing. They saw the laws, rules, and procedures intended to restrict access to research data as either weak or useful but not foolproof, noting several limitations such as human involvement and delegated decision-making. Their assessment of the laws, rules, and procedures intended to prevent misuse was similar, though they noted additional issues such as lack of enforcement. Fewer than half of respondents were reassured by the Common Rule, GINA, or the HIPAA Privacy Rule as a singular protection, citing the lack of ongoing influence of IRBs beyond initial review, limited scope and applicability in the research context, and gaps in coverage and enforcement challenges, respectively.

Our study is descriptive in nature; it cannot definitively answer questions of what the web of protections (or each of its components) objectively is or what it should be. Rather, our results comprise subjective

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148.

understandings from a group of thought leaders with diverse expertise. Their insights illuminate the real-world application of the web, going beyond objective descriptions of what a protection is and, instead, informing how a protection functions in actual practice. Like all policy-relevant results, ours are limited by the time in which they were generated. We conducted interviews in 2015-16. Since then, policy changes have been proposed or enacted which may impact the actual and/or perceived level of protection afforded by certain laws, regulations, and procedures, such as the Common Rule,[13] GINA,[14] and the 21st Century Cures Act.[15] Nonetheless, these experts' insights highlight weaknesses in the basic web of protections afforded to participants in precision medicine research.

Our study is descriptive in nature; it cannot definitively answer questions of what the web of protections (or each of its components) objectively is or what it should be. Rather, our results comprise subjective understandings from a group of thought leaders with diverse expertise. Their insights illuminate the real-world application of the web, going beyond objective descriptions of what a protection is and, instead, informing how a protection functions in actual practice.

First, researchers have an important, widely-acknowledged ethical responsibility to minimize risks and harms to participants.[16] This is embodied in various legal obligations including the Common Rule's specific requirements that researchers minimize risks to, and protect the privacy and confidentiality of, participants.[17] Our findings highlight the need for researchers to have an in-depth awareness of the gaps and limitations of the current web of protections and give robust attention to these challenges throughout the design and conduct of precision medicine research, as well as in the development of consent materials and processes. This may suggest the need for effective training to ensure that researchers understand and appreciate the kinds of risks and harms they should strive to avoid in designing and implementing the study and consent process. Similarly, these results point to the need for IRBs to build and maintain a comprehensive understanding of these issues in order to protect participants.

Next, our data may aid law- and policy-makers (in addition to researchers and IRBs) in assessing and strengthening the current frameworks for governance, oversight, and enforcement. Further empirical investigation is needed to identify, develop, and implement effective models for these protective mechanisms. Additionally, our findings point to the importance of the myriad of state laws governing precision medicine research.[18] These laws may provide models or other guidance for future law- and policy-making at the federal level, and may fill gaps in federal protections for some participants (though choice of law issues remain).[19] Our data also highlight the need for attention and adaptation to the external context in which precision medicine research occurs, including changes in law and regulation, medical and technological advancements, and evolutions in the socio-political environments.

Finally, as most thought leaders agreed, no single protection is sufficient to guard against the risks and potential harms associated with participation in precision medicine research.[20] Instead, it is the combination of protections that may be most effective. But as interviewees noted, even in the context of multiple layers of protections, human involvement is a primary source of concern insofar as the strength of any protection — whether legal/regulatory, technical, or procedural — is subject to the level of attention and care provided by the humans who are implementing, following, monitoring, and enforcing it. The success of precision medicine research depends on the public's trust in the research enterprise. Because no legal, regulatory, technical, or other protection will ever be foolproof, it is incumbent on researchers, institutional review boards, law and policy makers, and other stakeholders to demonstrably earn and maintain the trust of research participants and the public by attending to these issues.[21]

Note

The authors have no conflicts to declare.

Acknowledgments

This work was supported by a grant from the National Human Genome Research Institute (R01-HG-007733). The content is solely the responsibility of the authors and does not necessarily represent the official views of NHGRI or NIH. Thanks to our colleagues Leslie E. Wolf, Erin C. Fuse Brown, and Kevin C. McKenna.

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)

References

1. Y. Erlich and A. Narayanan, "Routes for Breaching and Protecting Genetic Privacy," *Nature Reviews Genetics* 15, no. 6 (2014): 409-421.
2. Z. Lin et al., "Genomic Research and Human Subject Privacy," *Science* 305, no. 5681 (2004): 183; E. E. Schadt, "The Changing Privacy Landscape in the Era of Big Data," *Molecular Systems Biology* 8, no. 612 (2012): 1-3.
3. J. Oliver et al., "Balancing the Risks and Benefits of Genomic Data Sharing: Genome Research Participants' Perspectives," *Public Health Genomics* 15, no. 2 (2012): 106-114; A. L. McGuire et al., "To Share or Not to Share: A Randomized Trial of Consent for Data Sharing in Genome Research," *Genetics in Medicine* 13, no. 11 (2011): 948-955.
4. 45 C.F.R. § 46.116(a) (2009); 45 C.F.R. § 46.116(a) (2017).
5. G. Guest et al., "How Many Interviews Are Enough? An Experiment with Data Saturation and Variability," *Field Methods* 18, no. 1 (2006): 59-82.
6. E. E. Namey and R. T. Trotter II, "Qualitative Research Methods," in G. S. Guest and E. E. Namey, eds., *Public Health Research Methods* (Los Angeles: SAGE Publications, 2015): 447.
7. Adapted from F. S. Collins and H. Varmus, "A New Initiative on Precision Medicine," *New England Journal of Medicine* 372, no. 9 (2015): 793-795; M. J. Khoury and J. P. Evans, "A Public Health Perspective on a National Precision Medicine Cohort: Balancing Long-Term Knowledge Generation with Early Health Benefit," *JAMA* 313, no. 21 (2015): 2117-2118.
8. L. M. Beskow et al., "Thought Leader Perspectives on Risks in Precision Medicine Research," in G. Cohen, H. Lynch, and E. Vayena, eds., *Big Data, Health Law, and Bioethics* (New York: Cambridge University Press, 2018): 161-174.
9. L. M. Beskow et al., "Thought Leader Perspectives on Benefits and Harms in Precision Medicine Research," *PLoS One* 13, no. 11 (2018).
10. K. M. MacQueen et al., "Codebook Development for Team-Based Qualitative Analysis," *Cultural Anthropology Methods* 10, no. 2 (1998): 31-36.
11. E. Namey et al., "Data Reduction Techniques for Large Qualitative Data Sets," in G. Guest and K. M. MacQueen, eds., *Handbook for Team-Based Qualitative Research* (New York: AltaMira Press, 2008); G. Guest et al., *Applied Thematic Analysis* (Thousand Oaks: SAGE Publications, Inc., 2012): 130.
12. National Institutes of Health, United States Department of Health and Human Services, "All of Us" Research Program, *available at* <https://allofus.nih.gov> (last visited February 1, 2019).
13. Federal Policy for the Protection of Human Subjects, 45 C.F.R. § 46 (2017) (Jan. 19, 2017).
14. 29 C.F.R. § 1635 (2016).
15. 21st Century Cures Act, Pub. L. No. 114–255 (2016).
16. National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, *The Belmont Report: Ethical Principles and Guidelines for the Protection of Human Subjects of Research* (Washington, DC: Government Printing Office, 1978).
17. 45 C.F.R. § 46.111(a) (2009); 45 C.F.R. § 46.111(a) (2017).
18. L. E. Wolf et al., "The Web of Legal Protections for Participants in Genomic Research," *Health Matrix: Journal of Law-Medicine* 29 (2019).
19. *Id.*
20. Beskow et al., *supra* note 8; Beskow et al., *supra* note 9.
21. C. Grady and A. S. Fauci, "The Role of the Virtuous Investigator in Protecting Human Research Subjects," *Perspectives in Biology and Medicine* 59, no. 1 (2016): 122-131.

The Journal of Law, Medicine & Ethics, 47 (2019): 134-148. © 2019 The Author(s)