Indiana Health Law Review

Volume XVIII **2021** **Number 1**

ARTICLES

AM I MY COUSIN'S KEEPER? A PROPOSAL TO PROTECT RELATIVES OF GENETIC DATABASE SUBJECTS

ROBERT I. FIELD,* ANTHONY W. ORLANDO,** & ARNOLD J. ROSOFF***

I. INTRODUCTION

We live in a brave new world full of exciting possibilities for medical and human advancement. Mapping of the genome has taken our ability to understand individual humans to a substantially more detailed and intimate level than has ever before been possible. At the same time, advances in computer science, artificial intelligence, and the universality of internet connectedness have ushered in the era of "big data" and enabled us to examine and understand as never before what all of us have in common and what makes each of us unique.[1] By aggregating data from and about large numbers of people and getting a broader view of the human collective, we can understand better and in greater detail what makes for the best health and health care of an individual human, especially when we can take it down to the intimate genetic level.

These advances have spawned the era of "precision medicine," in which we can diagnose and treat human conditions and ailments with much greater effectiveness. It is a time of great possibilities but also of great perils, particularly as regards the confidentiality of each individual's medical status and human

* Robert I. Field, J.D., M.P.H., Ph.D., is Professor of Law at the Kline School of Law, Professor of Health Management and Policy at the Dornsife School of Public Health at Drexel University, and an Adjunct Senior Fellow of the University of Pennsylvania's Leonard Davis Institute of Health Economics.

** Anthony W. Orlando, M.Sc., M.P.W., Ph.D., is Assistant Professor of Finance, Real Estate, and Law at the College of Business Administration at California State Polytechnic University, Pomona.

*** Arnold J. Rosoff, J.D., is Professor Emeritus of Legal Studies and Health Care Management at The Wharton School of the University of Pennsylvania and a Senior Fellow of Penn's Leonard Davis Institute of Health Economics. The authors would like to thank the attendees of the 2019 ASLME Health Law Professors Conference for their helpful comments. The authors also thank Aminata Jalloh for research assistance.

1. For an overview of "big data" in health care, see Wullianallur Raghupathi & Viju Raghupathi, *Big Data Analytics in Healthcare: Promise and Potential*, 2 HEALTH INFO. SCI. & SYS. art. 3 (2014).

characteristics and potential. We can achieve miraculous things in this new world, but we must proceed with caution, sensitivity and wisdom—or we may do great and lasting damage to long-cherished notions of individual privacy. Moreover, since the benefits potentially achievable depend on public trust in the data-sharing that is required, we must thread a needle between making sure that people are sufficiently aware of the risks to privacy but are not so frightened by these concerns that they shy away from informed participation in this grand endeavor.

Gathering, compiling, and analyzing human genetic data have become a large and growing societal enterprise. Governmental and private organizations are engaging to a considerable extent, and individual citizens have jumped on the bandwagon, patronizing direct-to-consumer ("DTC") genetic analysis services such as 23andMe, AncestryDNA, and the like. In all of these contexts, people are being asked to consent to various uses of their genetic data, and, being aware of the great good that can be gained from sharing that data, many are strongly motivated to do so. The enthusiasm about what we can learn and achieve is widespread and commendable, but we must take care not to throw caution to the wind. The confidentiality of personal health information is precious, and once compromised, it may be irretrievable.

To take full advantage of the invaluable opportunities presented by breakthroughs in genetic understanding coupled with the potential of big data research, it is important to understand the risks and benefits of this data-sharing and what is needed—both in individuals' own actions and those of governmental and private entities—to adequately protect personal information. What makes this situation exceptional and calls for special treatment is that sharing of genetic data not only puts the individual subject's data privacy at risk but also that of their genetic relatives, most of whom will not know of or have any ability to affect the subject's choices regarding how broadly to share this highly sensitive information. In this Article, we propose an approach to balancing these goals so that medical science can continue its forward march while mitigating risks to the privacy we all should hold dear.

Section II of our analysis examines the different contexts in which individuals' genetic data and related bits of personal health information are gathered, compiled, and used, dividing these into three broad categories: clinical, research, and proprietary. Section III considers the potential risks to personal privacy in each category and the types of protections in place to guard against those risks. Section IV focuses on the least regulated area of the genetic database ecosystem, proprietary databases, which have proliferated in recent years and for which only a minimal external regulatory oversight presently exists. Section V contains our call for the creation of a new regulatory mechanism that we call "Data Protection Review Boards" ("DPRB"). Our concept steps off from the model of Institutional Review Boards ("IRB") that oversee the protection of human subjects in research. DPRBs would review data-sharing arrangements between database companies and other private entities with an exclusive focus on privacy risks to data subjects and their identifiable relatives. They would operate outside of formal government structures and represent a cross-section of stakeholders, including privacy experts, attorneys, ethicists, subjects, and the entities seeking to share data. We identify various key elements to be addressed

in their formation and operation. Section VI summarizes our proposal and the need to act before individuals' control over their information is irretrievably lost.

II. THE USES AND USERS OF GENETIC DATABASES

Not all genetic databases are created equal. Organizations vary widely in their reasons for collecting genetic data, their relationship with the suppliers of the data, their planned use of those data, the value they can derive from the data, and the laws, stakeholders, norms, and expectations governing their data operations. These diverse factors create very different incentives and contexts motivating the privacy protections these organizations erect around genetic data they collect.

In terms of purpose, the spectrum of genetic databases can be divided into three broad categories: clinical, research, and proprietary. Each category corresponds to a qualitatively different legal framework for the collection, use, and sharing of data. We therefore use this taxonomy to distinguish databases where the law already affords a degree of privacy protection from those where we see a need for new legal protocols and safeguards.

A. Clinical

Clinical genetic databases are created and maintained to assist in the provision of health care. Patients give consent for their health care providers to collect their genetic information, typically with the understanding that the data will primarily be used for their own diagnosis and treatment. These databases have opened the door to personalized precision medicine that uses genetic information to guide diagnosis and treatment and even to repair deleterious mutations in specific genes. It also permits clinicians to target therapies for specific diseases, such as cancer, and to specific pathogens, eliminating the need for blunt tools with significant side effects, such as growing resistance to broad-spectrum antibiotics.[2] Because genes shape the body's internal systems, they also offer the additional benefit of illuminating a person's vulnerabilities prospectively. Thus, they can be *predictive* as well as *prescriptive*, not only aiding the fight against current illnesses but also raising awareness about the risk of future threats.

Because of this dual nature of genetic data—informing both the present and the future—the effects of their use may extend beyond the current patient-physician relationship and treatment agenda. Patients may wish to—and indeed, have a right to—know about future health risks that cannot be managed well with current treatment alone. Thus, health systems are torn between ensuring that the information is used wisely under expert care and empowering patients to address their own risks on their own terms. One early adopter in the former camp is NorthShore University HealthSystem in Illinois, which has two genetic testing programs: the Genetic and Wellness Assessment, a survey that patients take

---

2. Megan Molteni, *DNA Tests Could Help Docs Detect Infectious Diseases Faster*, WIRED (Oct. 22, 2018), https://www.wired.com/story/dna-sequencing-detect-infectious-disease [https://perma.cc/K29X-SPBH].

before their physical exam, and MedClueRx, an at-home kit administered by the pharmacogenomics clinic to facilitate "gene-powered prescription."[3] These programs primarily provide doctors with information about how patients metabolize different drugs and point to further testing if needed.[4]

A more expansive example comes from Geisinger Health System in Pennsylvania and New Jersey, which does whole genome sequencing for all patients as part of their primary care.[5] Geisinger's reporting system, GenomeCOMPASS, gives information directly to consumers concerning eighty specific genes, provides even more detailed information to doctors, and connects patients with a geneticist if they desire counseling to understand their results.[6] Both of these health systems are atypical in their use of genetic data, however, pushing past common limitations of electronic health record interoperability and the lack of physician training in incorporating genetics into their diagnosis and treatment.[7]

The Geisinger system goes above and beyond the current recommendations of the American College of Medical Genetics and Genomics ("ACMG"), which has been advising physicians about the communication of genetic testing results to patients since 2012.[8] As of this writing, the ACMG has identified fifty-nine genes that patients have a right to know about if they so choose, even though these findings may be "secondary" to their actual "diagnostic results."[9] As the list grows, so do the risks of misunderstanding and misinterpretation, with the attendant risks to health. Notwithstanding this, the ACMG does not recommend that genetic counseling be mandated alongside the communication of these test results, arguing that it would be too time-consuming to explain all fifty-nine findings.[10]

Many experts have noted, however, that genes tend to cluster, making it less daunting to explain the results than the ACMG's position would make it seem.[11] Other critics have challenged the ACMG's "all or nothing" approach, which does not allow patients to opt out of receiving information about *specific* genes, giving them less "autonomous choice."[12] They also argue that the ACMG encourages

---

3. Dava Stewart, *Genetic Testing as Part of Primary Care and Precision Medicine Is Underway at NorthShore University HealthSystem and Geisinger Health*, DARK DAILY (Oct. 8, 2018), https://www.darkdaily.com/genetic-testing-as-part-of-primary-care-and-precision-medicine-is-underway-at-northshore-university-healthsystem-and-geisinger-health/ [https://perma.cc/3RTN-AJSL].

4. *Id.*

5. *Id.*

6. *Id.*

7. *Id.*

8. Susan M. Wolf, *The Continuing Evolution of Ethical Standards for Genomic Sequencing in Clinical Care: Restoring Patient Choice*, 45 J.L. MED. & ETHICS 333, 333 (2017).

9. *Id.* at 334.

10. *Id.* at 337.

11. *Id.*

12. *Id.* at 334.

physicians to explain the risks of opting out but not the risks of opting in.[13] Patients may not be aware, for example, of the rate (or even the *possibility*) of false positives and act on the test results as if they were certain.[14] They also may not be aware of the many ways their genetic data can be used against them if the data fall into the wrong hands, as we discuss in Section IV below.[15]

Clinical databases have a measure of quality control in the form of the Clinical Laboratory Improvement Amendments of 1988 ("CLIA"), which directs the federal Department of Health and Human Services ("DHHS") to oversee the operation of clinical laboratories.[16] Such laboratories process biological samples from patients in connection with their care. The Centers for Medicare and Medicaid Services ("CMS") within the DHHS performs this function. As a result, laboratories that conduct analyses using genetic databases for patient care must meet standards that do not apply to other kinds of databases and other uses of them.

B. Research

Research genetic databases are created and maintained for scientific discovery and the development of new and improved medical practices. Subjects give consent for academic, government, or private research organizations to collect their genetic information, typically with the understanding that their data will be used for scientific research to benefit society. As these databases grow, so will the statistical power of these research studies. Not only will researchers be able to report their findings with greater confidence, but they will also be able to explore the heterogeneity of treatment effects, illuminating more precisely how results differ across subsamples of the populations. The more thinly they can slice the population into subsamples, the more useful their findings will become to precision medicine.

To this end, the federal government has undertaken two ambitious data collection efforts: the All of Us Research Program ("All of Us") conducted by the National Institutes of Health ("NIH") and the Million Veteran Program ("MVP") conducted by the Department of Veterans Affairs. Launched by the Obama administration, each program is recruiting one million volunteers to share their personal health information, including their genetic data. All of Us is especially

---

13. *Id.* at 337.

14. A large body of evidence demonstrates that many, if not most, physicians are woefully unskilled at calculating, understanding, and communicating the correct probability of illness associated with a patient's test result. *See, e.g.*, Daniel Morgan, *What the Tests Don't Show: Doctors Are Surprisingly Bad at Reading Lab Results. It's Putting Us All at Risk.*, WASH. POST (Oct. 5, 2018), https://www.washingtonpost.com/news/posteverything/wp/2018/10/05/feature/doctors-are-surprisingly-bad-at-reading-lab-results-its-putting-us-all-at-risk/ [https://perma.cc/HNM9-HEBW].

15. *See* Wolf, *supra* note 8.

16. 42 U.S.C. § 263a (2020).

intended to capture historically underrepresented minorities,[17] while MVP will reflect the population of veterans.[18] Both initiatives aim "to accelerate health research and medical breakthroughs, enabling individualized prevention, treatment, and care"[19] and "to learn how genes, lifestyle, and military exposures affect health and illness," respectively.[20] While they are not the largest genetic databases in the country, they are likely to be among the most detailed—and therefore heavily safeguarded with privacy protections.[21]

Unlike clinical databases, research databases do not have quality standards to ensure that the data themselves are reliable—in other words, that the laboratories are conducting high-quality tests. The CLIA standards only apply to clinical laboratories. As a result, the National Academies of Sciences, Engineering, and Medicine have recommended that the NIH develop a "quality management system" for research laboratories that analyze human biospecimens, and once the analyses have been completed, researchers should conduct "a peer-review process to assess the risks and benefits of results disclosure."[22] Whereas clinical databases are created primarily for the benefit of the patients, research databases have broader mandates, making them less likely to report the results back to the subjects who may benefit the most from them. Thus, these databases have earned the moniker of "helicopter research" for dropping into people's lives temporarily and then flying away with the valuable information they have gleaned, leaving the subjects no better off for their participation.[23]

C. Proprietary

Proprietary genetic databases are created and maintained for profit, which is earned by providing testing and related services for a payment and by sharing aggregated data with other organizations for a fee. Consumers give consent for companies to collect their genetic information, typically with little understanding of how their data will be used. As these databases grow, so do the opportunities for monetization. And grow they will, as this industry is projected to triple in size

---

17. The All of Us Research Program Investigators, *The "All of Us" Research Program*, 381 NEW ENG. J. MED. 668, 668 (2019).

18. John Michael Gaziano et al., *Million Veteran Program: A Mega-Biobank to Study Genetic Influences on Health and Disease*, 70 J. CLINICAL EPIDEMIOLOGY 214, 216 (2016).

19. NAT'L INST. HEALTH, https://allofus.nih.gov [https://perma.cc/5AS3-892C] (last visited Oct. 5, 2020).

20. *Million Veteran Program (MVP)*, U.S. DEP'T VETERANS AFF., https://www.research.va.gov/MVP/default.cfm [https://perma.cc/WM4N-VSAD] (last visited Oct. 5, 2020).

21. Eric Dishman, *I Handed over My Genetic Data to the NIH. Here's Why You Should, Too*, STAT (June 13, 2018), https://www.statnews.com/2018/06/13/entrusted-my-genetic-data-nih/ [https://perma.cc/C9P3-L6QU].

22. NAT'L ACADS. OF SCIS. ENG'G & MED. ET AL., RETURNING INDIVIDUAL RESEARCH RESULTS TO PARTICIPANTS: GUIDANCE FOR A NEW RESEARCH PARADIGM x, xxvi (2018).

23. *Id.* at ix.

over the next three years.[24] The uses of proprietary databases are multitudinous, bounded not by clinical or research needs but, rather, only by the entrepreneurial limits of human imagination. The main categories of such uses are discussed below.

1. Discovering Genetic Heritage

Some proprietary database companies exist to enable their customers (subjects) to use genetic testing to facilitate genealogy tracking, helping them to discover their genetic heritage. Even in this seemingly most innocuous of uses, however, there can be controversy, particularly among a population that does not fully understand the science behind the results—and therefore, often, their import.

The case of Senator Elizabeth Warren, in preparation for her campaign for the United States presidency, is instructive. Faced with criticism and skepticism of her claim to have Native American heritage, Senator Warren submitted her DNA to testing. The geneticist, Carlos Bustamante, immediately faced a data challenge: Native Americans are significantly underrepresented in genetic databases due in part to mistrust that their data will be misinterpreted or used for discriminatory purposes.[25] As a comparison to match Senator Warren's DNA, Bustamante substituted Colombian, Mexican, and Peruvian DNA, which he judged to be similar enough to detect her Native ancestry.[26] From this match, he concluded that Senator Warren indeed had a Native American ancestor, likely from the Cherokee Nation tribe, approximately six to ten generations ago, giving her DNA that is between 1/64th and 1/1024th Native American today.[27]

---

24. Tony Romm & Drew Harwell, *Ancestry, 23andMe and Others Say They Will Follow These Rules When Giving DNA Data to Businesses or Police*, WASH. POST (July 31, 2018), https://www.washingtonpost.com/technology/2018/07/31/ancestry-andme-others-say-they-will-follow-these-rules-when-giving-dna-data-businesses-or-police/ [https://perma.cc/G82G-MHTD].

25. *See* Lizzie Wade, *To Overcome Decades of Mistrust, a Workshop Aims to Train Indigenous Researchers to Be Their Own Genome Experts*, SCI. (Sept. 27, 2018), https://www.sciencemag.org/news/2018/09/overcome-decades-mistrust-workshop-aims-train-indigenous-researchers-be-their-own [https://perma.cc/V3MR-JSNF].

26. Glenn Kessler, *Just About Everything You've Read on the Warren DNA Test Is Wrong*, WASH. POST (Oct. 18, 2018), https://www.washingtonpost.com/politics/2018/10/18/just-about-everything-youve-read-warren-dna-test-is-wrong/ [https://perma.cc/N5CX-VKS3].

27. *Id.* Thereupon began a series of public misunderstandings and recriminations. Hearing these fractions and comparing them to widely reported averages, the Republican National Committee claimed that "Warren might even be less Native American than the average European American," which the *Washington Post* debunked as a misunderstanding of the test results—after admitting that the science was so little understood that the *Post* itself initially endorsed a similar mistaken interpretation. *Id*. The misunderstanding, according to the *New York Times* science reporter whose article was taken out of context, stems from the incorrect belief that each person inherits half of their DNA from each parent, one-quarter from each grandparent, and so on. Carl Zimmer, *Before Arguing About DNA Tests, Learn the Science Behind Them*, N.Y. TIMES (Oct. 18, 2018), https://www.nytimes.com/2018/10/18/opinion/sunday/dna-elizabeth-warren.html

Even with this level of genetic detail, however, ancestry is only one aspect of the complicated experience that defines any social constructed ethnicity, particularly an historically oppressed minority group. The Cherokee Nation took offense at the implication that Senator Warren belonged to or spoke on behalf of their cherished history and community, reminding the country that "being a Cherokee Nation tribal citizen is rooted in centuries of culture and laws not through DNA tests."[28] Senator Warren subsequently apologized, adding "I am not a person of color."[29] By then, however, it had become clear that the growth of proprietary databases had made it impossible for Native Americans to escape the misinterpretations and abuses that some feared when they chose not to share their data in the first place.

2. Clarifying Parentage

The inability to hide from the magnifying glass of modern genealogy is one of the most valuable and controversial aspects of proprietary databases. Once it becomes possible to identify people who do not wish to be found, their private lives can be threatened and exposed. Children who were adopted or conceived with donated gametes can learn about their biological parents, even if they never met them or knew their names.

On the one hand, the biological parent(s) might claim a right to remain anonymous, often based on a contract promising them anonymity when they put a child up for adoption or donated gametes. On the other hand, this anonymity can be psychologically harmful for children, severing the connection they intrinsically crave and the understanding of self that comes with such a connection. Without this anonymity, however, birth parents might be more cautious about permitting their children to be adopted, and gamete donation might become less frequent. From the child's perspective, were it not for a gamete donation, they would never have been born in the first place.[30] If these acts of procreation are valuable to society, the new transparency foisted upon potential parents by proprietary databases might come with costs for future generations.

Complicating this tradeoff further is the potential for abuse by "search angels" who use genetic tests and other personal information to help people find their birth families. Unfortunately, it is possible for unscrupulous purveyors to

---

[https://perma.cc/FHK6-B4CQ]. Thus, the public assumes that those fractions accurately reflect the *number* of Native American ancestors she had. On the contrary, she could have had—and indeed, likely *did* have—much more Native American *ancestry* that simply did not get retained proportionally in her DNA.

28. Asthma Khalid, *Warren Apologizes to Cherokee Nation for DNA Test*, NPR (Feb. 1, 2019), https://www.npr.org/2019/02/01/690806434/warren-apologizes-to-cherokee-nation-for-dna-test [https://perma.cc/YWQ9-LAF8].

29. *Id.*

30. Dani Shapiro, *How a DNA Testing Kit Revealed a Family Secret Hidden for 54 Years*, TIME (Jan. 3, 2019), https://time.com/5492642/dna-test-results-family-secret-biological-father/ [https://perma.cc/345Y-V24T].

pose as genetic experts, take advantage of people—and their limited understanding of this new science—and tell them the *wrong* conclusions about their ancestry.[31] If anyone deserves anonymity, it is *unrelated* individuals who are identified by these businesses.

3. Tracking Down Criminal Suspects

No one has a greater interest in remaining anonymous than criminals. Thus, proprietary genetic databases have increasingly found their way into law enforcement. The largest and most prominent databases, AncestryDNA and 23andMe, do not cooperate with law enforcement or share their data with public databases, but other databases are rapidly growing to fill the void.[32]

The most famous case involved genetic genealogist Barbara Rae-Venter, who helped police identify Joseph James DeAngelo as the Golden State Killer who committed at least fifty rapes and thirteen murders in California from 1974 to 1986.[33] A retired patent attorney with a Ph.D. in biology, Rae-Venter uploaded the suspect's DNA to the public database GEDmatch, which discloses in its terms of service that it cooperates with law enforcement and that consumers are allowed to submit the DNA of a third party without the person's consent.[34] Thus, no court order was required.[35]

The sample revealed the identity of a distant cousin of the suspect.[36] This enabled police to use other publicly available information to narrow the search.[37] While we can applaud the detective work and celebrate the outcome, the search's success demonstrates the ease with which a subject's genetic data can be linked to a web of people to whom they are related.[38]

Rae-Venter is not alone. In the wake of her success, the forensic genealogy market has expanded rapidly. By August 2018, the firm Parabon NanoLabs claimed that they had found eight perpetrators, including the Ramsey Street Rapist who committed at least six rapes in Fayetteville, North Carolina, between 2006 and 2008.[39] Another entrant, Interfinders International, declared its intention

---

31. Heather Murphy, *She Helped Crack the Golden State Killer Case. Here's What She's Going to Do Next.*, N.Y. TIMES (Aug. 30, 2018), https://www.nytimes.com/2018/08/29/science/barbara-rae-venter-gsk.html [https://perma.cc/QQ3H-BY7A].

32. Teri Figueroa, *Forensic Genealogists Shake Family Trees to Find Crime Suspects*, SAN DIEGO UNION-TRIB. (Jan. 2, 2019), https://www.sandiegouniontribune.com/news/public-safety/sd-me-forensic-genealogy-20180102-story.html [https://perma.cc/2KNU-EFKP].

33. *Id.*

34. *Id.*

35. Romm & Harwell, *supra* note 24.

36. Murphy, *supra* note 31.

37. *Id.*

38. Ironically, Rae-Venter did not want the police to disclose her name to the public out of fear for her safety, even though she had no qualms about pinpointing other individuals for police scrutiny based on their genetic profiles. *See id.*

39. Jacey Fortin, *In Serial Rape Case That Stumped Police, Genealogy Database Leads to*

to find suspects, heirs, and fraud.[40] Even Ancestry, which claimed that it would not cooperate with law enforcement, admitted that it gave data to law enforcement in thirty-one out of thirty-four cases in 2017, mostly related to credit card and identify theft.[41]

This newfound wealth of resources overcomes the constraints that police have faced historically in accessing their own criminal databases, primarily the Combined DNA Index System, for which they need a judge's approval in some states. Those same states do not restrict their access to public databases such as GEDmatch.[42] When many of these laws were written, it was nearly unimaginable that proprietary databases would become so extensive and forensic techniques so sophisticated.

In *Maryland v. King*, for instance, the Supreme Court ruled that DNA cheek swabs did not violate the Fourth Amendment on the grounds that they did "not reveal an arrestee's genetic traits."[43] Six years later, researchers reported in *Science* that they could successfully match 60% of Americans of European descent to their third cousin using "anonymized" genetic data on GEDmatch—and within two to three years, they predict that the match rate will increase to 90%.[44] Their test was less accurate for non-White Americans.[45] This raises two new risks for society: first, the increasing likelihood that law enforcement will have the power to view the full genetic profiles of the majority of Americans, and second, the greater likelihood of error with vulnerable minorities who already have a history of mistreatment at the hands of law enforcement.

These risks are at least partly mitigated when careful experts like Barbara Rae-Venter are involved, but forensic genealogists are increasingly teaching their techniques in do-it-yourself fashion to law enforcement officials.[46] What is to stop governments from abusing these new surveillance powers? For instance, authoritarian regimes could target democratic protestors and other perceived enemies, as they have done repeatedly throughout history.[47]

Arrest, N.Y. TIMES (Aug. 24, 2018), https://www.nytimes.com/2018/08/23/us/ramsey-street-rapist-dna.html [https://perma.cc/MG5F-ZVQ6].

40. Jessica Testa, *Nobody Was Going to Solve These Cold Cases. Then Came the DNA Crime Solvers.*, BUZZFEED (Sept. 22, 2018), https://www.buzzfeednews.com/article/jtes/dna-cold-case-crime-doe-project-genealogy [https://perma.cc/S68Z-Y6VW].

41. 23andMe, in contrast, refused all five requests that year. *See* Romm and Harwell, *supra* note 24.

42. Heather Murphy, *How an Unlikely Family History Website Transformed Cold Case Investigations*, N.Y. TIMES (Oct. 16, 2018), https://www.nytimes.com/2018/10/15/science/gedmatch-genealogy-cold-cases.html [https://perma.cc/S749-6B2U].

43. Maryland v. King, 569 U.S. 435, 438 (2013).

44. Yaniv Erlich et al., *Identity Inference of Genomic Data Using Long-Range Familial Searches*, 362 SCI. 690, 690-94 (2018).

45. *Id.* at 690.

46. Murphy, *supra* note 31.

47. Noa Yachot, *History Shows Activists Should Fear the Surveillance State*, ACLU (Oct.

4. Identifying Unidentified Bodies

Even when individuals do not seek anonymity, proprietary databases can implicate others. The DNA Doe Project, for example, has tried to avoid controversy by helping law enforcement identify *innocent* individuals—namely, unidentified dead bodies—who are not at risk of harm, but complications can still arise. What happens when they deduce the identity of a dead baby, leading law enforcement to find the mother? Unintentionally, they have identified a suspect in the child's death and disposal—a suspect who did not consent to the use of her genetic information by law enforcement.[48] And how do the families of the deceased know that they can trust the traumatic news they are receiving from this squad of volunteer genealogists? While genealogy has professional certification, genetic genealogy does not.[49] The genetic genealogist's findings do not come with any independent screening *ex ante* or verification *ex post*.

5. Finding Lost Relatives

The risks are no less profound for users who seek to locate lost relatives and reunite families. Though the aim once again can be admirable, the execution opens the door to new threats. For this reason, 23andMe and MyHeritage encountered significant public criticism when they donated genetic testing kits to the Trump administration to help them return immigrant children to their parents.[50] Especially in an administration hostile to immigrant rights, it is possible to imagine genetic data being used to track and discriminate against individuals long after they have left government custody. Officials were never required to prove that genetic testing was more effective than the collection of names, documents, and photographs that they were already using; rather, it normalized a more extensive system of data collection without evidence of its effectiveness.

Reliance on genetic profiling may even work *against* reunification in cases where the guardians are not genetic relatives. All the while, these data are being collected without informed consent, which is impossible to get from young children separated from their guardians.[51] Though genetic testing has been used successfully to reunite families in other countries, some of those same countries, such as China and Kuwait, have used their new data collection capabilities to target marginalized ethnic groups.[52] The threat is real.

27, 2017), https://www.aclu.org/blog/national-security/privacy-and-surveillance/history-shows-activists-should-fear-surveillance [https://perma.cc/WRY6-EZCJ].

48. Testa, *supra* note 40.

49. *Id.*

50. Ava Kofman, *DNA Testing Might Help Reunite Families Separated by Trump. But It Could Create a Privacy Nightmare.*, INTERCEPT (June 27, 2018), https://theintercept.com/2018/06/27/immigration-families-dna-testing/ [https://perma.cc/274R-CHG2].

51. *Id.*

52. *Id.*

6. Proprietary Uses for Clinical

Because there are so few restrictions on the uses to which these databases may be directed, it is not surprising to find the line blurring between proprietary and non-proprietary. Virtually all personal health information got into a patient's records through the actions of members of the health care team and those who worked closely with them. However, others can be involved, as well. For example, a partnership between 23andMe and the Michael J. Fox Foundation has offered free testing to 10,000 people with Parkinson's disease, releasing data to third parties outside the medical system.[53] At least when genetic tests are ordered by physicians, as in the partnership between Ancestry and the diagnostic testing group PWNHealth, CMS has regulatory authority.[54] To date, however, its intervention has been minimal. CMS does not require that doctors report results for all fifty-nine genes recommended for analysis by the ACMG but rather only results for seventeen genes. It does not consider whether patients overestimate the importance of these genes or even have the capacity to understand the nuances of the results. It does not recommend best practices to deliver emotionally challenging news, beyond the online videos and access to genetic counselors that most doctors already provide. Of course, such recommendations would be difficult to devise, since experts have not yet determined which practices are best for communicating this nascent form of health information.[55]

7. Proprietary Uses for Research

The same line-blurring is evident in research, which has become especially lucrative for some proprietary database companies. The most active participant appears to be 23andMe, largely due to the size of its database and the fact that 80% of its consumers have consented to the use of their genetic data for research.[56] It has entered into more than fifty academic collaborations, including a $60 million deal with Genentech,[57] a $300 million deal with GlaxoSmithKline,[58] and partnerships with Alnylam Pharmaceuticals, Biogen, Janssen, Lundbeck, and Pfizer.[59] By 2015, the average customer's genome had already been used in 230

---

53. Valerie Gutmann Koch & Kelly Todd, *Research Revolution or Status Quo?: The New Common Rule and Research Arising from Direct-to-Consumer Genetic Testing*, 56 HOUS. L. REV. 81, 93 (2018).

54. 42 C.F.R. § 410.32 (2020).

55. *See* Matthew Herper, *Ancestry Launches Consumer Genetics Tests for Health, Intensifying Rivalry with 23andMe*, STAT (Oct. 15, 2019), https://www.statnews.com/2019/10/15/ancestry-health-launch/ [https://perma.cc/6CTC-HULA].

56. Jennifer Abbasi, *23andMe, Big Data, and the Genetics of Depression*, 317 JAMA 14, 15 (2017).

57. *Id.*

58. Romm & Harwell, *supra* note 24.

59. Molteni, *supra* note 2.

different studies.[60] Just as significant, if not more so, as part of its research collaborations 23andMe asks those who have provided genetic material and consented to the use of their genetic data to also answer extensive surveys about their health status, co-morbidities, diet, lifestyle, sexual orientation, mental-emotional state, and other intimate personal matters.[61] These surveys also contain many questions about whether the responder's family members have or have had certain health characteristics and conditions. This is necessary for fully meaningful genetic analysis because health effects are the product of not only one's genetic makeup but also environmental factors, including how that person lives their life.

On another point of concern, experts have questioned the validity of some of the findings of these collaborations, since companies such as 23andMe do not examine patients in a traditional clinical or research setting—but rather tend to rely on self-reported measures, which have high error rates.[62] In an extensive survey questionnaire to which one of the authors responded for one commercial database, a large number of questions asked for specifics of past illnesses and conditions and indicated that, if the responder could not provide exact information, a "best guess" would be sufficient. Despite these concerns, research based on commercial genetic databases is proliferating. A survey of the medical literature from 2011 to 2017 found 181 publications that used private genetic data, with the annual publication record growing from four to fifty-seven studies per year over that time period.[63] Of these publications, 86% had an academic collaborator, and 45% had NIH funding, signifying just how far proprietary genetic databases have crossed the line into the established world of academic research.[64]

III. GENETIC EXCEPTIONALISM: SPECIAL THREATS TO PRIVACY

From time to time in the course of technological progress, it becomes necessary to update our concept of the right to privacy. When in 1888 George Eastman introduced the world to his new "snap camera," he placed into the hands of each consumer the power to record, to preserve, and thus to distribute every visible moment, no matter how fleeting or acontextual, with all the embarrassment, indignity, or incrimination it might arouse. It was partly in response to this new temptation that Samuel Warren and Louis Brandeis penned

---

60. Michael Grothaus, *How 23andMe Is Monetizing Your DNA*, FAST COMPANY (Jan 5, 2015), https://www.fastcompany.com/3040356/what-23andme-is-doing-with-all-that-dna [https://perma.cc/XJ94-HY9Y].

61. *See 23andMe Research Surveys and Questions*, 23ANDME, https://customercare.23andme.com/hc/en-us/articles/212881977-23andMe-Research-Surveys-and-Questions [https://perma.cc/8PTR-HVWL] (last visited Nov. 12, 2020).

62. Abbasi, *supra* note 56.

63. Kayte Spectator-Bagdady et al., *Genetic Data Partnerships: Academic Publications with Privately Owned or Generated Genetic Data*, 21 GENETICS MED. 2827, 2828 (2019).

64. *Id.*

their famous treatise on *The Right to Privacy*, wherein they described this new right as an emergent application of other rights long enshrined in the common law.[65] We begin our investigation of genetic privacy with this classic formulation, and we show how it is insufficient to address the risks of the technologies in our own time. From there, we propose a new concept of privacy, constructed to guard against four potential harms that may befall genetic relatives who are unwittingly ensnared in the database's grasp: wrongful publicity, discrimination, expropriation of profit, and violation of public trust.

A. Toward a New Concept of Privacy

"That the individual shall have full protection in person and in property is a principle as old as the common law," write Warren and Brandeis, and therein do they find many aspects of the right to privacy already in force.[66] In intellectual property rights, they find protection against publishing what creators do not want published without agreed-upon compensation.[67] In the Fifth Amendment, they find the right not to express personal thoughts.[68] In the rights not to be assaulted or defamed or imprisoned without due process—violations that all have "the quality of being owned or possessed"—they find the right "to be let alone."[69] If "[t]he common law has always recognized a man's house as his castle, impregnable," they argue, then surely that inviolability extends not just to physical intrusions but also to the intrusions of "idle or prurient curiosity."[70]

When these words were written in 1890, it was possible to draw a direct line from each privacy violation to the individual being harmed. When a voyeur peers into a man's house, he sees only the contents of that one house. He does not see into the man's neighbor's house or his cousin's or his distant relative's in a faraway land. When the state imprisons someone, its jail cells do not fill up with that person's entire extended family. When a person is assaulted, there are no bruises or broken bones to be found anywhere else in his family tree. The right to privacy, as originally conceived and applied, was meant to protect the individual alone. It did not envision a technological context that would allow a person to reveal what many other people might not want revealed, a secret so buried that this person does not even know that they are revealing it—indeed, that the revealer does not even know that the secret exists. The technology here is the ability to collect and analyze genetic information, and the intrusion it allows is not just individual but, rather, collective. To guard against this intrusion, it is not enough to require the consent of the person who owns the property or volunteers the information, as in the case of intellectual property rights or the Fifth

---

65. *See* Samuel D. Warren & Louis D. Brandeis, *The Right to Privacy*, 4 HARV. L. REV. 193 (1890).

66. *Id.* at 193.

67. *See id.* at 200-03.

68. *See id.* at 198-99.

69. *Id.* at 205

70. *Id.* at 220.

Amendment or the homeowner opening their door to prying eyes. For contained within those hidden truths are someone else's hidden truths, and by Warren and Brandeis's logic, the same "full protection" must apply to their person and property as well.[71]

The common law has advanced considerably in this direction in recent decades. Although it does not yet recognize a collective right to privacy, it acknowledges a person's "reasonable expectations of privacy" as a standard to be protected by the state. This test, famously articulated by Justice Harlan in his concurring opinion in *Katz v. United States*, requires that the person "have exhibited an actual (subjective) expectation of privacy and . . . that the expectation be one that society is prepared to recognize as 'reasonable.'"[72] A decade later, the Court applied constitutional protection to information privacy, extending the reasonable expectations standard to "the individual interest in avoiding disclosure of personal matters."[73]

By these standards, genetic information is a strong candidate for privacy protection. It is one of the most personal matters in one's life, both in *substance*—DNA, the body's fundamental building blocks, the very essence of one's existence—and in *content*—revealing everything from detailed family history to the most serious medical conditions. The public has exhibited its expectation of genetic privacy through legislation, contractual agreements, and public opinion polls.[74] All that remains to establish a collective right to privacy is a demonstration that society ought to recognize these expectations as "reasonable," not only for the individual giving the DNA sample but for everyone else implicated by genetic relation. In the remainder of this section, we make this case by enumerating the potential harms that warrant such expectations of privacy.

B. Wrongful Publicity

The same technological advances that allow the health care system to collect greater quantities of genetic information simultaneously create more possibilities for criminals to steal and expose that genetic information. Data privacy laws vary widely across states, leaving millions of Americans far less protected than they could be.[75] Even if data managers deidentify the data, cyber experts have found

---

71. *Id.* at 193.

72. Katz v. United States, 389 U.S. 347, 361 (1967).

73. Whalen v. Roe, 429 U.S. 589, 599 (1977).

74. *See, e.g.*, Christi J. Guerrini et al., *Should Police Have Access to Genetic Genealogy Databases? Capturing the Golden State Killer and Other Criminals Using a Controversial New Forensic Technique*, 16 PLOS BIOLOGY art. e2006906 (2018); *see also* Scott Hensley, *Poll: Genealogical Curiosity Is a Top Reason for DNA Tests; Privacy a Concern*, NPR (June 1, 2018), https://www.npr.org/sections/health-shots/2018/06/01/616126056/poll-genealogical-curiosity-is-a-top-reason-for-dna-tests-privacy-a-concern [https://perma.cc/G68J-QCR2].

75. *See* Karen Turner, *The Equifax Hacks Are a Case Study in Why We Need Better Data Breach Laws*, VOX (Sept. 14, 2017), https://www.vox.com/policy-and-politics/2017/9/13/

that it is possible to reidentify much of it, a problem of which most Americans are not aware.[76] Genetic data theft takes this problem even farther than other data theft; while a medical identification number or a credit card number may change, a person's genome will stay with them to the day they die.[77]

Wrongful exposure of this kind is harmful in several ways. Human beings typically cannot establish "a close, relaxed, and frank relationship" without some degree of intimacy.[78] They strategically expose their information as a relationship deepens as a way to signal the importance they place on the person and the moment at hand.[79] By depriving them of the ability to control this exposure—control over the very identity they wish to create themselves—wrongful publicity robs them of this careful construction of fragile relationships. People may interact less, experiment with ideas less, and express their emotions less in a community where they do not feel that their persona is safe and accepted by the other members of the community.[80] If they cannot know what others know about them—and this is particularly true of genetic databases where they are tracked through distant relatives—then, in the words of one author, they are being "deliberately deceive[d] . . . about [their] world."[81]

For these reasons, the *Restatement (Second) of Torts* recognizes four privacy torts related to wrongful publicity: public disclosure of private facts, intrusion upon seclusion, the "false light" in which publicity puts a person, and appropriation of someone's "name or likeness."[82] Any violation typically must rise to the level of being "highly offensive to a reasonable person."[83] Examples include disclosing membership in a political group formed under the First Amendment's Freedom of Association Clause,[84] publicizing the names of abortion patients or in vitro fertilization patients,[85] and even subpoenaing deidentified late-term abortion records.[86] In the latter case, a federal appellate court ruled that the data could be identified by "persons of their acquaintance, or

---

16292014/equifax-credit-breach-hack-report-security [https://perma.cc/59ZZ-8B2R].

76. Scott R. Peppet, *Regulating the Internet of Things: First Steps Toward Managing Discrimination, Privacy, Security, and Consent*, 93 TEX. L. REV. 85, 128 (2014).

77. Jennifer Kulynych, *Is Privacy the Price of Precision Medicine?*, OUPBLOG (Mar. 26, 2017), https://blog.oup.com/2017/03/privacy-precision-medicine/ [https://perma.cc/CL8K-JDY7].

78. ALAN F. WESTIN, PRIVACY AND FREEDOM 31 (1967).

79. *See* Julie E. Cohen, *Examined Lives: Informational Privacy and the Subject as Object*, 52 STAN. L. REV. 1373 (2000).

80. WESTIN, *supra* note 78; *see also* Paul M. Schwartz, *Property, Privacy, and Personal Data*, 117 HARV. L. REV. 2055 (2004).

81. Stanley I. Benn, *Privacy, Freedom, and Respect for Persons*, *in* NOMOS XIII: PRIVACY 10 (1971).

82. RESTATEMENT (SECOND) OF TORTS § 652 (AM. LAW INST. 1977).

83. *Id.*

84. NAACP v. Alabama, 357 U.S. 449 (1958); Shelton v. Tucker, 364 U.S. 479 (1960).

85. Doe v. Mills, 536 N.W.2d 824 (Mich. Ct. App. 1995); Y.G. v. Jewish Hosp. of St. Louis, 795 S.W.2d 488 (Mo. Ct. App. 1990).

86. Nw. Mem'l Hosp. v. Ashcroft, 362 F.3d 923 (7th Cir. 2004).

skillful 'Googlers,'"[87] a concern that can apply equally to family history and other genetic information. The court observed:

> Even if there were no possibility that a patient's identity might be learned from a redacted medical record, there would be an invasion of privacy. Imagine if nude pictures of a woman, uploaded to the Internet without her consent though without identifying her by name, were downloaded in a foreign country by people who will never meet her. She would still feel that her privacy had been invaded.[88]

It is not too far of a stretch to imagine a future where her genetic information will allow that same foreign voyeur to construct an even more detailed portrait of her than just a few nude photos.

C. Discrimination

In 1924, Dr. Albert Sidney Priddy petitioned the Virginia State Colony for Epileptics and Feebleminded to forcibly sterilize an 18-year-old, mentally deficient inmate named Carrie Buck.[89] In an 8-1 decision, the Supreme Court ruled in favor of Dr. Priddy and upheld the Virginia law that authorized the sterilization.[90] Justice Oliver Wendell Holmes Jr. argued in the majority opinion that the "public welfare" was more important than Buck's right to procreate.[91] She was, he said, too "promiscuous," as her mother before her had been and, he predicted, her daughter after her, leading to his famous conclusion: "Three generations of imbeciles are enough."[92] Although *Buck v. Bell* has been widely criticized over generations, it has never been expressly overturned by the Supreme Court.

The bygone logic of eugenics may seem repugnant by modern standards, but the crusade to discriminate against undesirable genes continues to this day. Consider the case in 1998 when the Lawrence Berkeley Laboratory tested job applicants for sickle cell trait without their knowledge.[93] Or in 1999 when Terri Sergeant was fired one month after her employer learned of her genetic disposition for lung disease—the firm's health plan was self-insured.[94] Or in 2001 when a company was sued for discriminating against workers who were "more likely" to develop carpal tunnel syndrome.[95] This initial suit was dismissed

---

87. *Id.* at 929.

88. *Id*.

89. Buck v. Bell, 274 U.S. 200, 205 (1927).

90. *Id.* at 207-08.

91. *Id.* at 207.

92. *Id.*

93. Norman-Bloodsaw v. Lawrence Berkeley Lab., 135 F.3d 1260 (9th Cir. 1998).

94. Anita Silvers & Michael Ashley Stein, *Human Rights and Genetic Discrimination: Protecting Genomics' Promise for Public Health*, 31 J.L. MED. & ETHICS 377, 380-82 (2003).

95. EEOC v. Woodbridge Corp., 263 F.3d 812, 813 (8th Cir. 2001).

because genetic discrimination itself was not illegal at the time.[96]

These incidents are not isolated. Research shows that a significant percentage of individuals with a genetic predisposition for Huntington's disease experience discrimination in employment (6.5%), insurance (25.9%), and relationships (32.9%).[97] Across North America, a genetic test revealing sudden arrhythmia death syndromes is associated with a 60% likelihood of being rejected by insurers (disability, health, life, or travel)—and for those who do get insurance, a 39% likelihood of higher premiums.[98] In one survey, 46% of respondents worried that genetic testing would lead to stigmatization based on the resulting diagnoses, leading many of them to consider avoiding it.[99]

They have good reason to be concerned. The history of health care is replete with dangerous, damaging, ill-informed stigmas. HIV-related stigma is so severe that it leads to a significant increase in depression and alcohol abuse in patients.[100] It is so pervasive that researchers have created a global HIV Stigma Index to track it over time,[101] and courts have ruled that HIV-related disclosures "cause a violation of the family's privacy much greater than simply revealing any other aspect of their family medical history."[102] But medical problems and a genetic predisposition to such problems need not rise to the level of HIV infection to expose a person to potential discrimination. Despite statutory protections against discrimination discussed in Section IV below,[103] cancer survivors report experiencing negative stereotypes in the workplace about their ability to do their job, their productivity and reliability, and the future cost of their illness to the company.[104] If they apply for a new job, they are less likely to be viewed as

---

96. *See* EEOC v. Woodbridge Corp., 124 F. Supp. 2d 1132, 1139 (2000). Discrimination in employment or insurance on the basis of one's genetic makeup has been outlawed by the Genetic Information Nondiscrimination Act of 2006. *See* discussion *infra* Section IV(A)(1).

97. Cheryl Erwin et al., *Perception, Experience, and Response to Genetic Discrimination in Huntington Disease: The International RESPOND-HD Study*, 153B AM. J. MED. GENETICS PART B 1081, 1082 (2010).

98. Saira Mohammed et al., *Genetic Insurance Discrimination in Sudden Arrhythmia Death Syndromes: Empirical Evidence from a Cross-Sectional Survey in North America*, 10 CIRCULATION: CARDIOVASCULAR GENETICS (2017), at 1, https://www.ahajournals.org/doi/pdf/10.1161/CIRCGENETICS.116.001442 [https://perma.cc/F5NP-RLJV].

99. Davit Chokoshvili et al., *Public Views on Genetics and Genetic Testing: A Survey of the General Public in Belgium*, 21 GENETIC TESTING & MOLECULAR BIOMARKERS 195, 195 (2017).

100. *See* Kaylee B. Crockett et al., *Experiences of HIV-Related Discrimination and Consequences for Internalised Stigma, Depression and Alcohol Use*, 34 PSYCHOL. & HEALTH 796 (2019).

101. Barbara A. Friedland et al., *Measuring Intersecting Stigma Among Key Populations Living with HIV: Implementing the People Living with HIV Stigma Index 2.0*, 21 J. INT'L. AIDS SOC'Y 115, 115 (2018).

102. Doe v. Barrington, 729 F. Supp. 376, 385 (D.N.J. 1990).

103. *See* discussion *infra* Section IV.

104. Mary Stergiou-Kita et al., *The "Big C" – Stigma, Cancer, and Workplace Discrimination*, 10 J. CANCER SURVIVORSHIP 1035, 1035 (2016).

competent, with significantly negative hiring results.[105] Although the majority of the population does not appear to discriminate against cancer patients or survivors, the rate is high enough to represent a significant challenge for anyone trying to get a mortgage (31%), receive "the best possible care" (17%), and even feel accepted in normal social situations (17%).[106] Because genetic testing can reveal predispositions for many types of cancers, it exposes the patient to all of these risks, as well as innumerable others. Unlike a single diagnosis, a genetic test is a Pandora's box of potential risk factors, introducing more stigmas than are even known to exist yet. It is not hyperbole, therefore, to agree with Jessica Roberts that we risk creating a future world with a "genetic underclass."[107]

D. Expropriation of Profit

Historically, patients and research subjects have not had ownership rights in their biological specimens once extracted from their bodies or in any resulting personal health information. Once extracted, the specimens and data belong to the organization, whether a university, hospital, or research institute, that collects them. The data are the organization's to sell, share, or use for research purposes. The "donors" no longer have the right to access their own data, to exclude others from accessing them, or to commercialize them. Therefore, courts have found that they are not entitled to a share in the profits earned from research that uses those data.[108]

John Moore famously challenged this commercialization when the UCLA Medical Center created and patented a profitable cell line using T-lymphocytes that were collected from his spleen, which was removed as part of his treatment for hairy-cell leukemia.[109] Moore characterized this use of his biological samples as a *conversion*, an unlawful expropriation of personal property, arguing that the doctor, his assistant, UCLA, a related genetics institute, and the Sandoz Pharmaceutical Company had violated his property rights.[110] The Supreme Court of California disagreed.[111] In their judgment, patients do not have an ownership

---

105. Larry R. Martinez et al., *Selection BIAS: Stereotypes and Discrimination Related to Having a History of Cancer*, 101 J. APPLIED PSYCHOL. 122, 122-23 (2016); *see also* Mary Stergiou-Kita et al., *Stigma and Work Discrimination Among Cancer Survivors: A Scoping Review and Recommendations*, 84 CANADIAN J. OCCUPATIONAL THERAPY 178 (2017).

106. Charlotte Vrinten et al., *Cancer Stigma and Cancer Screening Attendance: A Population Based Survey in England*, 19 BMC CANCER art. 566 (2019), at 1.

107. Jessica L. Roberts, *The Genetic Information Nondiscrimination Act as an Antidiscrimination Law*, 86 NOTRE DAME L. REV. 597, 597 (2011) [hereinafter *The Genetic Information Nondiscrimination Act*].

108. *See* Jessica L. Roberts, *Progressive Genetic Ownership*, 93 NOTRE DAME L. REV. 1105 (2018) [hereinafter *Progressive Genetic Ownership*].

109. Moore v. Regents of Univ. of Cal., 51 Cal.3d 120, 125-29 (Cal. 1990).

110. *Id.*

111. *Id.* at 125.

interest in their body parts after removal.[112] In contrast to a wrongful-publicity case where a person retains an ownership interest in their own unique likeness, wrote Justice Edward Panelli for the majority, "the particular genetic material which is responsible for the natural production of lymphokines . . . is no more unique to Moore than the number of vertebrae in the spine or the chemical formula of hemoglobin."[113] Federal patent law reinforces this perspective, bestowing protection on "the product of 'human ingenuity,' but not naturally occurring organisms."[114] The goal of ownership, in this view, is to reward "inventive effort . . . not the discovery of naturally occurring raw materials."[115]

Of course, there are other reasons for property rights. People may not be able to patent naturally occurring raw materials, but they certainly cannot mine them without the consent of the owner of the mine. Here, ownership does not serve to reward "inventive effort" but rather to protect the land from unproductive use and unrestrained exploitation—to prevent a "tragedy of the commons."[116] By analogy, we can understand why Justice Stanley Mosk dissented that "scientists or industrialists" ought not to have "the right to appropriate and exploit a patient's tissue for their sole economic benefit - the right, in other words, to freely mine or harvest valuable physical properties of the patient's body."[117] Imagine how much more widespread this exploitation becomes when they can appropriate and exploit valuable property from not just one patient but, rather, from dozens or even hundreds with a single sample from a single genetic test.

Nor is uniqueness a necessary qualification for ownership. In most supply chains in most industries, the inputs are homogeneous. One brick is the same as another brick, one barrel of oil identical to the next. We do not deny the suppliers their fair recompense because they lack originality. On the contrary, we reward them for supplying these necessary inputs and give them an incentive to supply even more in the future. We discourage theft. Why should a hammer or a nail be more deserving of compensation than the human genome? We shudder to think of the possibility that research subjects might fail to supply the inputs we need for lifesaving experiments. As Justice Allen Broussard noted in partial dissent, why should they not demand a say in what happens to their donation, just as organ donors have in transplantation?[118]

It is not clear, however, that property law is the best way to resolve these challenges. There is a difference between a thing and what information we can

---

112. *Id.* at 153-54.

113. *Id.* at 139. While this may be true as a legal matter, it was the special physical properties of Moore's T-lymphocytes that made them uniquely valuable. Thus, they were unique in a physical sense.

114. *Id.* at 141-42.

115. *Id.* at 142 (emphasis omitted); *see also* Diamond v. Chakrabarty, 447 U.S. 303 (1980) (holding that genetically modified living organisms can be patented).

116. For the modern formulation of our understanding of this problem, see Garrett Hardin, *The Tragedy of the Commons*, 162 SCI. 1243 (1968).

117. *Moore*, 51 Cal.3d at 174 (Mosk, J., dissenting).

118. *Id.* at 154-55 (Broussard, J., concurring and dissenting).

learn from the thing. As James Boyle points out, "I could stare at my genetic code all day and not even know it was mine."[119] But someone else could, and Boyle cannot begin to fathom what that person might come to know about him. Property law is not designed to protect him from this revelation. Only privacy serves to prevent "disclosure of intimate, embarrassing, or simply 'personal' *socially constructed facts* about ourselves," especially when those facts are unknowable to the person donating the "property."[120]

A patient like Moore might have more control over the future use of his body parts, for example, if the physician disclosed the potential profitmaking opportunities prior to donation. Moore's physicians did not, and the court ruled that they breached their fiduciary duty and failed to obtain informed consent as a result.[121] As Justice Mosk points out, however, these particular privacy protections do not "reach a major class of potential defendants: all those who are outside the strict physician-patient relationship with the plaintiff."[122] Nor do they reach a major class of potential *plaintiffs*, though Justice Mosk could not have known this given the early state of genetics at the time. We know that Moore is not the only person whose genetic information is contained in his biological samples. In fact, his data are so valuable precisely *because* they implicate so many different people with common genetic characteristics. The profit is in the collective.

E. Violation of Public Trust

Americans have been losing faith in collective institutions for over four decades.[123] Nevertheless, the medical profession still ranks high on measures of public trust.[124] This trust is critical to the success of the health care system. Patients who do not trust physicians are less likely to go to regular checkups, get necessary treatments, or take critical medication.[125] They may dissuade others from engaging with the health care system, and they may turn to dangerous alternative treatments instead. A decline in trust on par with other collective institutions could trigger nothing short of a public health crisis.

---

119. JAMES BOYLE, SHAMANS, SOFTWARE, AND SPLEENS: LAW AND THE CONSTRUCTION OF THE INFORMATION SOCIETY 105 (1996).

120. *Id.* (emphasis added).

121. *Moore*, 51 Cal.3d at 148.

122. *Id.* at 180 (Mosk, J., dissenting).

123. Bill Bishop, *Americans Have Lost Faith in Institutions. That's Not Because of Trump or "Fake News,"* WASH. POST (Mar. 3, 2017), https://www.washingtonpost.com/posteverything/wp/2017/03/03/americans-have-lost-faith-in-institutions-thats-not-because-of-trump-or-fake-news/ [https://perma.cc/4E89-N2KA].

124. Cary Funk et al., *Trust and Mistrust in Americans' View of Scientific Experts*, PEW RES. CTR. (Aug. 2, 2019), https://www.pewresearch.org/science/2019/08/02/trust-and-mistrust-in-americans-views-of-scientific-experts/ [https://perma.cc/E2WF-HVFU].

125. *See* Johanna Birkhäuer et al., *Trust in the Health Care Professional and Health Outcome: A Meta-Analysis*, 12 PLOS ONE art. e0170988 (2017).

Confidentiality is a critical component of this trust. Patients need to feel comfortable divulging their most personal, intimate, and sometimes embarrassing details in order to arm their physician with the necessary information to diagnose, treat, and advise them. The Supreme Court has recognized that this type of "confidential relationship [is] necessary for successful treatment" in psychotherapy, as in other forms of medical practice.[126] That is why, as far back as the Hippocratic oath, physicians have sworn that they "will not divulge . . . all [that] should be kept secret,"[127] and a majority of American states have enshrined this principle in law as a "physician-patient privilege."

Imagine how many more people might be scared away from the medical system if they learn that their information is being abused even *before* they volunteer it because they are genetically related to someone who donated to a proprietary database. This is a likely outcome if they witness the other harms we have discussed in this section. Already there are reports that African-Americans are opting out of genetic testing because they fear police abuse,[128] and we saw the concerns of Native Americans come true in the Elizabeth Warren case described in Section I above. Law enforcement already ranks low on the surveys of public trust, and we can see how that tension spills over into community relations, public safety, and the resulting negative effects on residents.[129] Do we want to see the same damage done to our health care system, the same retreat from our research studies, the same fear, the same questioned motives?

IV. CURRENT GENETIC PRIVACY PROTECTIONS AND THEIR LIMITATIONS

In this new and broader landscape of vulnerability to privacy invasions based on genetic information, existing legal protections are based on a patchwork of laws that leave substantial gaps. The most important ones were written decades ago, when genetic databases and their uses were but a distant vision. The major laws can be divided into two main categories: (1) restrictions on use, which proscribe certain uses of information once it has been obtained, and (2) restrictions on access, which limit the persons or entities with whom data may be shared and under what circumstances. Their coverage falls far short of being comprehensive, and some important risks remain unaddressed.

---

126. Jaffee v. Redmond, 518 U.S. 1, 10 (1996) (footnote omitted).

127. *Quoted in* DANIEL J. SOLOVE & PAUL M. SCHWARTZ, INFORMATION PRIVACY LAW 483 (5th ed. 2015).

128. Amy Dockser Marcus, *For Some African-Americans, Genetic Testing Reopens Past Wounds*, WALL STREET J. (July 14, 2018), https://www.wsj.com/articles/for-some-african-americans-genetic-testing-reopens-past-wounds-1531566000 [https://perma.cc/25FE-7DDV].

129. *See* Andrew McCall, *Resident Assistance, Police Chief Learning, and the Persistence of Aggressive Policing Tactics in Black Neighborhoods*, 81 J. POL. 1133 (2019); *see also* Cheryl Boudreau et al., *Police Violence and Public Perceptions: An Experimental Study of How Information and Endorsements Affect Support for Law Enforcement*, 81 J. POL. 1101 (2019).

A. Use Restrictions

1. The Genetic Information Nondiscrimination Act

The move to ban discrimination based on a person's genetic makeup was not immediate or unanimous. From the time the first genetic antidiscrimination bill was introduced in Congress in 1994, it took fourteen years to get the majority of both chambers onboard. The business community worried about the cost of frivolous lawsuits. Insurance companies worried about losing their underwriting ability, predicting adverse selection and rising costs to consumers.[130] Opponents even argued the law would prevent patients from learning about genetic tests that could save their lives. The most significant challenge to enacting a law, however, was the lack of evidence that genetic discrimination was, in fact, a real and serious risk. In general, law is a response to felt needs; society must perceive a problem, or at least share a strong sense that a problem is imminent, before it is moved to enact a law to correct or head off the problem. Situations in which some are foresighted enough to anticipate an approaching problem and persuasive enough to convince others of the need to take prospective preventative action are rare. The early emergent concern about genetic discrimination is one example of this cultural phenomenon.

The Genetic Information Nondiscrimination Act ("GINA") was finally passed by Congress in 2008.[131] It prohibits discrimination in health insurance and in employment for firms with fifteen or more employees.[132] Its health insurance protection was rendered largely superfluous by the Patient Protection and Affordable Care Act ("ACA"), which was enacted two years later[133] and prohibits insurers from excluding people from coverage because of preexisting health conditions, and it does not apply to other forms of insurance, such as those covering disability, life, and long-term care.[134]

Previous antidiscrimination laws had been retrospective. The Civil Rights

---

130. "Underwriting" in insurance is the process of identifying and evaluating the various risk factors and deciding the appropriate premium to charge for covering the risk, or whether to issue coverage at all. The ACA's prohibition against excluding pre-existing conditions from coverage goes against generations of insurance industry theory and practice. Without the protections afforded by the ACA and GINA, the risk of genetic discrimination would be a clear and present danger. *See* Arnold J. Rosoff & Anthony W. Orlando, *Employers and Health Insurance Under the Affordable Care Act*, 24 ANNALS HEALTH L. 470 (2015).

131. 42 U.S.C. ch. 21F (2020).

132. *What Is Genetic Discrimination?*, MEDLINEPLUS, https://medlineplus.gov/genetics/understanding/testing/discrimination/ [https://perma.cc/L4NR-CHLE] (last visited Nov. 12, 2020).

133. If either Congress or the Supreme Court were to nullify the ACA's prohibition of discrimination in health insurance on the basis of having a pre-existing condition, GINA would, of course, no longer be superfluous in this regard.

134. *What Is Genetic Discrimination?*, *supra* note 132.

Act,[135] the Age Discrimination in Employment Act,[136] the Rehabilitation Act,[137] and the Americans with Disabilities Act ("ADA")[138] each sought to end a form of discrimination that had a long history. With a few exceptions, genetic testing was too new, too undeveloped, and too limited to have a history of abuse. GINA was therefore preemptive. It envisioned the kinds of discrimination that could happen and attempted to prevent them *ex ante*. To its supporters, this approach was admirably foresighted; to its detractors, it was a recipe for unintended consequences.[139]

This unique context led, in part, to a unique design. The history of discrimination has created subordinated groups, leading Congress to craft "anti-subordination" statutes like the Civil Rights Act that go beyond merely outlawing negative differential treatment. These laws actually allow positive differential treatment to uplift the subordinated groups. After decades of inequality, they recognize that discrimination is built into the system in ways that are often not obvious. As a result, they allow "disparate impact" actions where claimants can challenge policies that have discriminatory effects even if intentional discrimination against the claimants cannot be established.

GINA, in contrast, does not yet have an overtly recognizable "genetic underclass" to protect. It therefore prohibits employers and health insurance companies from classifying people with respect to genetic information.[140] This type of "anti-classification" statute only prevents negative differential treatment and allows "disparate treatment" actions only where it is clear that the claimants have been intentionally targeted.[141]

Because of its limited scope, GINA allows any kind of genetic discrimination not explicitly listed in the law, including education, housing, insurance other than health coverage, and mortgage lending. Had Carrie Buck lived in the age of GINA, it would not have protected her from sterilization, since reproductive rights are not mentioned in the law.[142] For that matter, no actual health conditions are covered by the law.[143] Once a genetic predisposition develops into an actual illness, GINA no longer applies.[144]

2. The Americans with Disabilities Act

The ADA, passed in 1990, prohibits discrimination based on disability in

---

135. Civil Rights Act of 1964, Pub. L. No. 88-352, 78 Stat. 241.

136. Age Discrimination in Employment Act of 1967, Pub. L. No. 90-202, 81 Stat. 602.

137. Rehabilitation Act of 1973, Pub. L. No. 93-112, 87 Stat. 355.

138. Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 327.

139. *See* Jessica L. Roberts, *Preempting Discrimination: Lessons from the Genetic Information Nondiscrimination Act*, 63 VAND. L. REV. 439 (2010).

140. *See What Is Genetic Discrimination?*, *supra* note 132.

141. *The Genetic Information Nondiscrimination Act*, *supra* note 107, at 633.

142. *See* 42 U.S.C. ch. 21F (2020).

143. *See id.*

144. *Id.*

employment and "public accommodations," which include most facilities and services available to members of the public.[145] Job offers, promotions, and other terms of employment may not be denied or curtailed because of a disability.[146] The same applies to a denial or limitation of access to a private or governmental facility, such as an office, store, entertainment venue, hotel, or hospital.[147] Furthermore, employers and proprietors must make "reasonable accommodations" to enable a disabled person to perform a job or access a facility.[148] What is "reasonable" in this regard is generally assessed in terms of cost.[149]

A disability is defined as a condition that substantially limits one or more major life activities, having a history or record of having such a condition, or being perceived as having one.[150] Major life activities are defined by regulations to include major functional capabilities, such as walking, standing, speaking, and concentrating.[151] They also include major physiologic functions, such as immune response, digestion, brain activity, circulatory function, endocrine function, and reproduction.[152] A substantial limitation in these and similar functions triggers the ADA's protections.

The ADA would clearly protect an individual with a genetic trait that limits a major life function and has manifested itself.[153] However, the protection afforded to an individual with a trait that merely confers an enhanced probability of developing such a limitation in the future is less clear. An employer's concern over a current or prospective employee's enhanced probability of developing a genetic condition may trigger the law's protections for individuals *perceived* as having a disability, but the courts have required that the perception must concern "a substantially limiting impairment," not the predisposition to incur such an impairment.[154] Although the law was amended in 2008 to expand the definition of disability in response to these decisions, it still does not cover discrimination based solely on a genetic trait in an individual who is asymptomatic.[155] Therefore, should an employer acquire genetic information on a current or prospective employee suggesting that they stand an enhanced probability of developing a disability in the future, the ADA would offer no protection should the employer

145. Americans with Disabilities Act of 1990, Pub. L. No. 101-336, 104 Stat. 327.

146. *See id.*

147. *See id.*

148. *Id.*

149. *See id.*

150. 42 U.S.C. § 12102(1) (2020).

151. 42 U.S.C. § 12102(2)(A) (2020).

152. 42 U.S.C. § 12102(2)(B) (2020).

153. *See* 42 U.S.C. § 12102 (2020).

154. *See* Sutton v. United Air Lines, Inc., 527 U.S. 471 (1999).

155. ADA Amendments Act of 2008, Pub. L. No. 110-325, 122 Stat. 3553; *see also* Carly B. Eisenberg, *Genetic Predispositions v. Present Disabilities: Why Genetically Predisposed Asymptomatic Individuals Are Not Protected by the Amended ADA*, 15 B.U. J. SCI. & TECH. L. 131, 149 (2010).

discriminate based on it. Legal recourse might lie in challenging the legitimacy of the employer's receipt of the information, especially because employers are prohibited from requesting genetic information in most circumstances.[156] However, unauthorized receipt of information may be difficult to prove.

3. Insurance Underwriting under the ACA and the Health Insurance Portability and Accountability Act

The ACA includes strong protections against the use of most health-related information in underwriting for individual health insurance. Insurance companies may not even require customers to answer questions about health status, including preexisting conditions, on an application for coverage.[157] This prohibition applies not only to genetic information but to all medical information. Insurers may only ask applicants about their age, sex, geographic location, and use of tobacco.[158] If an insurer were to gain information about an applicant's genome, it could not use that information to deny a policy or to set the premium for it.[159]

A similar protection applies for group insurance under employer plans. The Health Insurance Portability and Accountability Act ("HIPAA") prohibits underwriting of applicants who switch jobs and leave one employer risk group for another.[160] The new insurer may not ask about health status or adjust coverage based on it, as long as the worker had enough continuous months of coverage with the prior employer. Other provisions of HIPAA regarding the privacy of medical information are discussed below. While the ACA and HIPAA provide important protections against discrimination based on genetic information in offering and pricing health insurance, this protection is limited to those two contexts; these laws do not ban the use of genetic information in other forms of insurance or in any other situation.

---

156. U.S. EQUAL EMP. OPPORTUNITY COMM'N, EEOC-NVTA-0000-4, FACT SHEET: GENETIC INFORMATION NONDISCRIMINATION ACT (2014) ("There are six very limited circumstances under which an employer may request, require, or purchase genetic information: Where the information is acquired inadvertently, in other words, accidentally; As part of a health or genetic service, such as a wellness program, that is provided by the employer on a *voluntary* basis; In the form of family medical history to comply with the certification requirements of the Family and Medical Leave Act, state or local leave laws, or certain employer leave policies; From sources that are commercially and publicly available, including newspapers, books, magazines, and electronic sources (such as websites accessible to the public); As part of genetic monitoring that is either required by law or provided on a *voluntary* basis; and By employers who conduct DNA testing for law enforcement purposes as a forensic lab or for human remains identification.")

157. 42 U.S.C. § 300gg (2020).

158. 42 U.S.C. § 300gg(a)(1)(A)(iv) (2020) (stating that insurers are permitted to charge people who use tobacco products up to 50% more for coverage in individual and small-group policies).

159. *See* 42 U.S.C. § 300gg(a)(1) (2020).

160. 29 U.S.C. § 1181 (2020).

4. Genetic Property Laws

Courts have generally refused to recognize that data subjects, whether patients or research participants, have an ownership right to their genetic information.[161] However, a number of states have enacted statutes that grant them such a right.[162] As of this writing, Alaska, Colorado, Florida, Georgia, and Louisiana have passed such laws, and several other state legislatures have considered joining them.[163] The laws grant individuals limited property interests in their own biological information.

However, genetic property laws are not without risks of their own. Requiring informed consent from everyone is no small challenge. If enough people decline to consent, the research databases may not be large or representative enough to be useful, and scientific advance will be unfortunately hampered. This "tragedy of the anti-commons" is a classic collective action problem.[164] When the benefits of participating in research are diffuse and speculative, and the risks of data-sharing can fall heavily on the individual, self-interest may lead these individuals to decline opportunities to participate in genetic research, to the detriment of medical progress and the population as a whole. Conversely, stronger property rights, or other data privacy protections, might make people more comfortable participating in research, particularly if they stand to benefit financially, or in some other significant way, from sharing their data. Researchers, in turn, might

---

161. Among the most influential decisions are: *Moore v. Regents of the University of California*, discussed above; *Greenberg v. Miami Children's Hospital Research Institute*, in which a federal district court rejected property claims by parents who had donated blood and tissue samples for development of a prenatal genetic test; and *Washington Univ*ersity *v. Catalona*, in which the Eighth Circuit Court of Appeals found that research subjects had no property interest in tissue samples they had donated for research and, therefore, no say over whether the lead researcher could take them to another university. Moore v. Regents of the University of California, 51 Cal.3d 120 (Cal. 1990); Greenberg v. Miami Children's Hospital Research Institute, 264 F. Supp. 2d 1064 (S.D. Fla. 2003); Washington University v. Catalona, 437 F. Supp. 2d 985 (E.D. Mo. 2006).

162. As an example of a case involving such a law, in 2013, Alaska resident Michael Cole bought a DTC genetic test online from Family Tree DNA. He swabbed his cheek, signed the release form, and sent in the sample. The company analyzed his DNA and sent him a link with the results so he could research his ancestry. It also gave him the option to join online forums where he could meet consumers with similar ancestry. What it did not tell him—and the release form did not explain—was that it would expose his genetic information publicly on these forums. In Alaska, that is illegal. Under the state's Genetic Privacy Act of 2004, Alaskans own both their DNA samples and any test results based on them. No company can disclose a consumer's genetic "property" without his or her informed consent. As of this writing, the case is still working its way through the courts. *See Michael Cole v. Gene by Gene, Ltd*., No. 1:14-cv-00004-SLG, 2019 WL 2571244 (D. Alaska June 21, 2019).

163. *Progressive Genetic Ownership*, *supra* note 108, at 1128.

164. The "tragedy of the anti-commons" occurs when one person can prevent others from using a common resource. It is the inverse of the more well-known "tragedy of the commons" by which everyone has an incentive to overuse a common resource.

feel more willing to share their databases for collaboration if they do not have a profit incentive to withhold the data from their competitors.

Laws that recognize and protect property interests of data subjects who submit genetic samples are also of no benefit to relatives of those consumers who share their valuable genetic traits. They face a risk of identification and must contend with the porous protections provided by anti-discrimination laws. Yet, they stand to derive no financial benefit from the sharing of information about their genes.

B. Access Restrictions

1. HIPAA

The risk of unwanted disclosure and abuse of patient medical information was partially anticipated by HIPAA, which restricts the use and disclosure of Protected Health Information ("PHI"), defined as "individually identifiable health information . . . that is: (i) Transmitted by electronic media; (ii) Maintained in electronic media; or (iii) Transmitted or maintained in any other form or medium."[165] Specifically, regulations implementing HIPAA, known as the HIPAA Privacy Rule, prohibit "Covered Entities"—health care providers, health care clearinghouses, and health plans—as well as their "business associates" from disclosing PHI, unless it has been deidentified, except in certain defined circumstances.[166] Thanks to HIPAA, genetic data collected in a clinical setting should be protected from most prying eyes if it contains personally identifiable information.

This protection is far from complete, however. First, research has shown that deidentified genetic data can often be reidentified, exposing one's personal health information to the outside world.[167] Second, HIPAA contains numerous exceptions to the disclosure restrictions. These include disclosure without a patient's consent (a) to other clinicians for treatment, (b) to insurance companies and other payers for payment, (c) for administrative activities, (d) to public health authorities, and (e) in response to a warrant or court order.[168] Data can also be disclosed in identifiable form for research purposes with the subject's authorization, as discussed below.[169] Finally, patients have limited recourse to enforce HIPAA and seek remedies. They can file a complaint with the Office for Civil Rights of the DHHS, but they do not have a private right of action.[170]

---

165. 45 C.F.R. § 160.103 (2020).

166. 45 C.F.R. § 160.310 (2020).

167. Katherine Drabiak, *Caveat Emptor: How the Intersection of Big Data and Consumer Genomics Exponentially Increases Informational Privacy Risks*, 27 HEALTH MATRIX 143, 167 (2017).

168. 45 C.F.R. § 164.512 (2020).

169. 45 C.F.R. pt. 160 (2020).

170. Leslie E. Wolf et al., *The Web of Legal Protections for Participants in Genomic Research*, 29 HEALTH MATRIX 1, 16.

Within the Privacy Rule is a set of standards specifically designed to protect electronically stored patient data, known as the Security Rule.[171] Under it, all Covered Entities must assess their security risks and put administrative, physical, and technological safeguards in place to mitigate them.[172] While these measures enhance protection against hackers, they do not alter the rules that permit sharing of information that is lawfully obtained.

Under this regulatory scheme, the Privacy Rule bars access to genetic information under many circumstances. For example, a health care provider could not provide patient genetic data to a marketing firm without the patient's consent. However, the protection applies only to data obtained in a health care context and affords several paths to circumvention. Also, it does little for a patient's relatives who may be identifiable in genetic data that are shared.

Beyond HIPAA, few laws govern clinical genetic databases, and the ones that do are often conflicting. Across the country, some states extend the meaning of Covered Entities to include employers, researchers, or other recipients of health information, but there is no uniform common practice.[173] Only the laboratories where the test results originate are governed by federal quality standards, set by the Clinical Laboratory Improvement Act ("CLIA"),[174] but enforcement is uneven. While HIPAA requires all laboratories to return test results to the patient, CMS prohibits laboratories from doing so if they are not CLIA-certified. Thus, HIPAA's insistence on transparency conflicts with CLIA's caution regarding quality, leading the National Academies of Sciences, Engineering, and Medicine to call for new legislation to resolve this contradiction.[175]

Perhaps most pertinent to genetic databases, HIPAA was not designed to address the rights of genetic *relatives* whose health risks might also be implicated in a patient's test results. As far back as 1988, the American Society of Human Genetics advised that physicians should alert non-patients that they are at risk when "[t]he patient is unwilling to inform his or her at-risk family member(s), the at-risk family member(s) are easily identifiable, harm is likely to occur, and the condition is treatable or prevention measures and surveillance will reduce the risk."[176] Ethicists continue to wrestle with this type of disclosure outside the immediate bounds of the physician-patient relationship, especially when the at-risk family member is a minor who cannot legally consent or "opt in" to receiving information about themselves. One must weigh the value of delaying until the child is old enough to decide against the "risk of adult-onset conditions" that require advance "life-planning."[177] To date, no consensus exists regarding this tradeoff.

---

171. 45 C.F.R. pt. 160 (2020).

172. 45 C.F.R. §§ 164.304, 164.310, 164.312 (2020).

173. Wolf et al., *supra* note 170, at 54.

174. 42 U.S.C. § 263a (2020).

175. NAT'L ACADS. OF SCIS. ENG'G & MED. ET AL., *supra* note 22, at xxvii.

176. Alicia Latham Schwark & Michael F. Walsh, *Duty to Warn in the Era of Next Generation Sequencing*, 18 AM. J. BIOETHICS 79, 79 (2018).

177. Wolf, *supra* note 8, at 336.

2. The Common Rule

Paramount among the protections for research data is The Federal Policy for the Protection of Human Subjects, generally known as "the Common Rule."[178] The Common Rule is a regulation adopted by nineteen federal agencies that sponsor research involving human subjects to protect research participants.[179] It was revised in 2018 to ease restrictions on the use of data obtained in the course of research.[180] The Rule and its limitations are discussed in more detail in Section V below, where we present our proposal to build on its approach to provide wider privacy protection for all data subjects.

Since 1991, the Office of Human Research Protection ("OHRP") of NIH has administered the Common Rule to govern "all research involving human subjects conducted, supported, or otherwise subject to regulation by any federal department or agency that takes appropriate administrative action to make the policy applicable to such research."[181] In all such research, the Common Rule requires researchers to obtain written consent from the potential subjects after providing them with a description of the risks and benefits of their involvement in the research. Specifically, the researchers must communicate to the subjects "whether identifiers will be removed, whether biospecimens will be used for commercial purposes, whether the individual can expect to share in any profits, and whether clinically actionable results of genetic testing or genomic sequencing will be returned to the individual."[182]

To enforce these protections, the Rule requires that Institutional Review Boards ("IRBs"), committees located in each organization that receives federal research funding, oversee studies that use human subjects.[183] Risks to subjects can include the possibility of physical or psychological harm and also the possibility that the confidentiality of personal information will be compromised. To guard against such risks, IRBs are charged with reviewing research protocols before they are implemented and reviewing research as it proceeds to enforce the protections that they mandate.[184]

---

178. The Common Rule grew out of the Belmont Report of 1978, written by the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, which was created pursuant to the National Research Act of 1974. *See* THE NAT'L COMM'N FOR THE PROT. OF HUMAN SUBJECTS OF BIOMEDICAL & BEHAVIORAL RESEARCH, THE BELMONT REPORT: ETHICAL PRINCIPLES AND GUIDELINES FOR THE PROTECTION OF HUMAN SUBJECTS OF RESEARCH (1979), https://www.hhs.gov/ohrp/sites/default/files/the-belmont-report-508c_FINAL.pdf [https://perma.cc/8K2C-EFX2]. The Common Rule and its major revision in 2018 are discussed later herein. *See* discussion *infra* Section IV(B)(2).

179. The regulations of the Department of Health and Human Services are codified at 45 C.F.R. pt. 46 (2020).

180. *See* 45 C.F.R. § 46.101(*l*) (2020).

181. 45 C.F.R. § 46.101 (2020).

182. Wolf et al., *supra* note 170, at 22.

183. § 46.101.

184. 45 C.F.R. § 46.109 (2020).

IRB review applies only to research that is either federally funded or used in support of an application for approval of a new drug by the federal Food and Drug Administration ("FDA").[185] It does not apply to privately funded research or to data-sharing by commercial entities for other purposes, such as marketing. Moreover, IRBs focus their review on the protection of the research subjects themselves, so risks to relatives are rarely, if ever, considered.

Moreover, even the explicit protections that IRBs enforce are not complete. First, written consent is *not* required for nonidentifiable information or deceased individuals. Again, what the Common Rule envisioned as "nonidentifiable" in 1991 under then-current data technology may now be—or may soon become—identifiable in this era of "big data," with greatly advanced technological progress and available access to many more sources of personal data, both health and non-health.

Second, there has been considerable debate over whether genetic biospecimens qualify as "human subjects." In 2008, when researchers tried to publish an early study from one of these databases in *PLoS Genetics*, the OHRP decided that genetic data did not fall into this category, exempting them from the Common Rule's requirements.[186] In 2012, the DHHS clarified this ruling with the interpretation that genetic biospecimens *do* qualify as "human subjects" *if* there is other personally identifiable information attached.[187] Finally, in 2018, the revised Common Rule officially concluded that deidentified biospecimens do not require consent, but the federal government is tasked with regularly reassessing what constitutes "identifiable personal information" and "identifiable biospecimen" in light of the ever-changing technological possibilities.[188]

Even more troubling are the possibilities for use of subjects' genetic data after the primary research has been completed. A particular concern for many subjects is that their data could be accessible to law enforcement. However, researchers may obtain a Certificate of Confidentiality from the NIH to prevent the government from forcing them to disclose their data "in any Federal, State, or local civil, criminal, administrative, legislative, or other proceeding."[189] The NIH believes Certificates of Confidentiality are necessary for "getting people comfortable" with sharing their personal information with the All of Us program.[190] Historically, the DHHS Secretary had broad discretion to issue Certificates of Confidentiality, but the 21st Century Cures Act of 2017 made it mandatory for all federally funded research to obtain them.[191] Privately funded researchers have the option, but not the obligation, to apply for them under this

---

185. *See* § 46.101.

186. Greg Gibson & Gregory P. Copenhaver, *Consent and Internet-Enabled Human Genomics*, 6 PLOS GENETICS art. e1000965 (2010), at 1.

187. Koch & Todd, *supra* note 53.

188. 45 C.F.R. § 46.102(e) (2020).

189. Wolf et al., *supra* note 170, at 12 (footnote omitted) (quoting 42 U.S.C. § 241(d) (2016)); *see* discussion *infra* Section V(B).

190. Wolf et al., *supra* note 170, at 4.

191. *Id.* at 20.

law.[192]

Even with a Certificate of Confidentiality, however, it is still possible for researchers to share genetic data after the study is completed. If they obtain "broad consent" from the subjects that does not specify a one-time use, then those data can be used for secondary research without specific consent for the later studies.[193] Even states that extend the Common Rule protections to non-federally-funded research—California, Maryland, New York, and Virginia—allow secondary research without specific consent.[194] Whether subjects really understand this implication when they give broad consent—in other words, whether it is truly *informed* consent—is questionable. Thus is born the possibility of limitless uses of the same data.

3. The European Union General Data Protection Regulation

The European Union ("EU") moved ahead of the United States in aggressively protecting data on individuals in 2016, when it adopted the European Union General Data Protection Regulation ("GDPR").[195] That law grants data subjects rights to limit the storage and use of information concerning them that has been collected and stored electronically.[196] It applies to all personal data, including genetic information in both anonymous and anonymized forms, and gives subjects the right to access their data, to have data transferred to a different entity, or to have them deleted.[197] Companies that store data must also implement reasonable measures to protect data from loss and exposure.[198] They must notify government authorities within seventy-two hours and data subjects as soon as possible after learning of a breach and must perform regular data protection impact assessments and data protection compliance reviews to identify and address potential risks.[199]

The GDPR applies not only to companies based in the EU but also to any company that collects or processes personal data on EU citizens.[200] This effectively extends its protections to virtually all data collection companies based in the United States. Penalties for noncompliance can be severe, including fines

---

192. *Id.* at 30 n.129.

193. Mahsa Shabani et al., *From the Principles of Genomic Data Sharing to the Practices of Data Access Committees*, 7 EMBO MOLECULAR MED. 507, 508 (2015).

194. Wolf et al., *supra* note 170, at 27.

195. *See* Council Directive 2016/679, 2016 O.J. (L 119) 7 (EU).

196. *See id.*

197. *What Is GDPR, the EU's New Data Protection Law?*, GDPR.EU, https://gdpr.eu/what-is-gdpr/ [https://perma.cc/E42Z-F76K] (last visited Nov. 13, 2020).

198. *See id.*

199. *What Does GDPR Stand for? (And Other Simple Questions Answered)*, GDPR.EU, https://gdpr.eu/what-does-it-stand-for/ [https://perma.cc/CH45-2ZCC] (last visited Nov. 13, 2020); *see also Data Protection Impact Assessment (DPIA)*, GDPR.EU, https://gdpr.eu/data-protection-impact-assessment-template/ [https://perma.cc/K9NK-EF75] (last visited Nov. 13, 2020).

200. *What Is GDPR, the EU's New Data Protection Law?*, *supra* note 197.

of up to 4% of a company's annual revenue.[201] A private right of action is available to subjects who believe their data are being misused or transferred for purposes they do not approve of.[202] No current protection in American law is as strong. However, since it is based on foreign authority, the GDPR does not provide a remedy through American courts.

Despite the GDPR's position as the world's most aggressive effort to date to protect data stored and shared electronically, its protections have significant limits. Subjects must affirmatively assert their rights, and it is not clear how many will take the time and effort to do so.[203] Polling shows that details of the law's protections are still widely misunderstood.[204] Data subjects also may not realize how much data are maintained on them, the potential uses of those data, and the full extent of the risks that disclosure can present. Furthermore, the GDPR offers no protection to relatives of data subjects. If a subject permits their data to be retained, the possibility always exists that a genetic relative of that subject, either a known or unknown relative, might be identified.

4. California Consumer Privacy Act ("CCPA")

In addition to international and federal initiatives in the United States to protect personal privacy in the electronic age, protections have also been enacted by some states. In 2018, the California Legislature enacted the California Consumer Privacy Act ("CCPA"), which bolsters the rights of consumers whose data are collected by businesses.[205] The first major legislation of its kind in the United States, the CCPA empowers individuals in the State to demand that companies disclose the personal data that have been collected on them and, if the individual wishes, to have those data deleted.[206] The law's reach extends beyond technology companies, such as Facebook and Google, to retailers, such as Walmart and Target.[207] It went into effect January 1, 2020, and genetic data are one of the types of information covered.[208]

The CCPA was modeled on the GDPR, but it is considerably more limited in important respects.[209] It protects only California residents and applies only to for-profit companies, so it offers no protection concerning data maintained by

---

201. *Id.*

202. *See Everything You Need to Know About the "Right to Be Forgotten*," GDPR.EU, https://gdpr.eu/right-to-be-forgotten/ [https://perma.cc/4QSB-LFJ4].

203. *See id.*

204. *Do Consumers Know Their GDPR Privacy Rights?*, GDPR.EU, https://gdpr.eu/consumers-gdpr-data-privacy-rights/ [https://perma.cc/MHC7-F35M] (last visited Oct. 25, 2020).

205. *See* CAL. CIV. CODE § 1798.150 (2020).

206. *Id.*; *see* Kari Paul, *California's Groundbreaking Privacy Law Takes Effect in January. What Does It Do?*, GUARDIAN (Dec. 30, 2019), https://www.theguardian.com/us-news/2019/dec/30/california-consumer-privacy-act-what-does-it-do [https://perma.cc/4D54-QGWV].

207. *See* Paul, *supra* note 206.

208. *Id.*; CIV. CODE § 1798.150.

209. Paul, *supra* note 206.

nonprofit organizations and universities.[210] Companies are only required to provide the most recent twelve months of data in response to a request.[211] Moreover, they are exempt if they have complied with other applicable laws, such as HIPAA.[212] The CCPA's penalties are meager, with a fine of only $7,500 for violations,[213] and like the GDPR, the law offers no protection for relatives of data subjects whose genetic information may be identifiable in a database.

5. Proprietary Database Terms of Service

With the exception of the GDPR and CCPA, proprietary genetic databases are largely exempt from the laws that limit access. For the most part, they follow the contract approach of other, non-genetic, for-profit companies. Each company writes its own rules embodied in its "terms of service," to which consumers must agree before they use the product. Industry experts refer to arrangements as "clickwrap" when consumers are required to click on a specific part of a website signifying consent and as "browsewrap" when they implicitly consent just by using the website.[214] The Federal Trade Commission ("FTC") typically upholds these agreements under their "Notice and Choice," or "Notice and Consent," framework.[215] Research has shown, however, that consumers rarely read what they are agreeing to, calling into question whether they really are given notice *or* choice—and certainly nothing rising to the level of informed consent generally required in health care.[216]

Even if they did read the terms of service with care, consumers typically would not find all the facts they need to make a truly informed choice. A recent survey of ninety direct-to-consumer ("DTC") genetic testing companies found that 39% did not have any privacy policy available on their website.[217] Of the policies that were accessible, the majority gave vague information about confidentiality and security.[218] Almost all reserved the right to modify their

---

210. Rita Heimes & Sam Pfeifle, *New California Privacy Law to Affect More Than Half a Million US Companies*, IAPP (July 2, 2018), https://iapp.org/news/a/new-california-privacy-law-to-affect-more-than-half-a-million-us-companies/ [https://perma.cc/3GZH-LB68]; *see* CAL. CIV. CODE § 1798.140 (2020).

211. *See* Paul, *supra* note 206.

212. CAL. CIV. CODE § 1798.145 (2020).

213. CAL. CIV. CODE § 1798.155 (2020).

214. Andelka M. Phillips, *Reading the Fine Print When Buying Your Genetic Self Online: Direct-to-Consumer Genetic Testing Terms and Conditions*, 36 NEW GENETICS & SOC'Y 273, 278 (2017).

215. *See, e.g.*, Joel R. Reidenberg et al., *Privacy Harms and the Effectiveness of the Notice and Choice Framework*, 11 I/S: J.L. & POL'Y FOR INFO. SOC'Y 485 (2015).

216. *See, e.g.*, Jorge L. Contreras, *Genetic Property*, 105 GEO. L.J. 1, 6-7 (2016).

217. James W. Hazel & Christopher Slobogin, *Who Knows What, and When?: A Survey of the Privacy Policies Proffered by U.S. Direct-to-Consumer Genetic Testing Companies*, 28 CORNELL J.L. & PUB. POL'Y 35, 48 (2018).

218. *Id.*

privacy policies after consumers had agreed to them, and very few companies promised to notify consumers when their policies had changed.[219] Only 49% of the policies gave consumers any information about what would happen to their physical sample after it had been processed.[220] Regarding the resulting genetic data, 45% of the policies said the company would keep the data forever, and 42% did not give any information about how long it would keep them.[221] If consumers want to delete their genetic data in the future, 56% of the policies did not mention whether or how that could be done.[222] Over 70% did not give the consumer any information about the ownership, licensing, or commercialization of the data,[223] and an astounding 95% said nothing about potential data breaches.[224]

If consumers want to ask questions about privacy issues, only 67% of the policies gave contact information, and "subsequent communication with customer service representatives revealed that, with the exception of industry leaders, these representatives were generally poorly equipped to handle privacy-related inquiries."[225] Perhaps it is not surprising, therefore, that 43% of the research publications using these proprietary genetic databases did not specify any informed consent or disclosure procedure.[226] Quite often, the privacy policies are so vague as to disclose little meaningful information—if they even exist at all.

In July 2018, the industry leaders—AncestryDNA, 23andMe, Helix, MyHeritage, African Ancestry, and FamilyTreeDNA—agreed to voluntary guidelines to improve these policies,[227] although the changes come with several caveats. The companies vowed to require consumers to give "separate express consent" to share genetic data with third parties,[228] but the companies did not commit to report *every* instance of data-sharing. They volunteered to disclose the number of law enforcement requests they receive each year,[229] not including any requests bound by a gag order. They will allow consumers to delete their data, but not if researchers are already using it. Finally, they will attempt to ensure that customers submit their *own* data and not someone else's.[230]

Because the companies have announced these guidelines publicly, the FTC can penalize them for false or misleading advertising practices if they violate

---

219. *Id.* at 49.

220. *Id.* at 50.

221. *Id.* at 51.

222. *See id.*

223. *Id.* at 52.

224. *Id.* at 53.

225. *Id.* at 64-65.

226. Spectator-Bagdady et al., *supra* note 63.

227. *See* Carson Martinez, *Privacy Best Practices for Consumer Genetic Testing Services*, FPF (July 31, 2018), https://fpf.org/2018/07/31/privacy-best-practices-for-consumer-genetic-testing-services/ [https://perma.cc/X8Q7-6AM6].

228. Romm & Harwell, *supra* note 24.

229. *Id.*

230. *Id.*

them.[231] However, FTC enforcement has been spotty. To date, the agency has only taken action against DTC genetic testing companies twice.[232]

The majority of these policies, therefore, provide little, if any, privacy protection. Moreover, in cases where protections do exist, it would be difficult for a data subject to know when a company had violated them or to enforce the terms if it did. Relatives would not even know that a commercial database contained data that could identify them.

One other federal regulator has intervened in this industry, but only with regard to a limited range of activities. The FDA ordered 23andMe to stop marketing the health benefits of its tests in 2013.[233] However, two years later, the agency cleared the company to advertise its tests for "ancestry, wellness, traits, and carrier status for inherited disorders."[234] Since then, it has approved 23andMe's carrier status test for Bloom Syndrome, Genetic Health Risk ("GHR") tests for ten diseases, and a test for three mutations in BRCA genes which predispose carriers to breast cancer.[235] This approval comes with the ongoing requirement that the company demonstrate the analytical and clinical validity of the tests, as well as sufficient consumer understanding of their results.[236] The FDA has also exempted future GHR tests from premarket review.[237] A few states supplement this regulation of health-related tests by requiring a physician's order, but most states do not.[238]

C. Remaining Legal Gaps

This inconsistent and porous set of legal protections cries out for stronger measures. The need is made all the more urgent by the rapidly growing ability of data analysts to identify subjects in seemingly anonymous genetic databases.[239]

---

231. *Id.*

232. Jennifer K. Wagner, *FTC Takes Action to Protect Consumers from False Genetic Advertising Claims*, PRIVACY REP. (July 3, 2014), https://theprivacyreport.com/2014/07/03/ftc-takes-action-to-protect-consumers-from-false-genetic-advertising-claims/ [https://perma.cc/PT37-Q8RN].

233. Amanda Holpuch, *FDA Orders Genetic Company 23andMe to Cease Marketing of Screening Service*, GUARDIAN (Nov. 25, 2013), https://www.theguardian.com/science/2013/nov/25/genetics-23andme-fda-marketing-pgs-screening [https://perma.cc/P6S2-W5HX].

234. Abbasi, *supra* note 56.

235. *Direct-to-Consumer Tests*, U.S. FOOD & DRUG ADMIN., https://www.fda.gov/medical-devices/vitro-diagnostics/direct-consumer-tests [https://perma.cc/M7GM-FRZJ] (last updated Dec. 20, 2019).

236. *Id.*

237. *Id.*

238. Emily Mullin, *As Consumer DNA Testing Grows, Two States Resist*, MIT TECH. REV. (Sept. 28, 2017), https://www.technologyreview.com/s/608958/as-consumer-dna-testing-grows-two-states-resist/ [https://perma.cc/FXG9-FHZA].

239. Erika Check Hayden, *Privacy Protections: The Genome Hacker*, 497 NATURE 172, 173 (2013).

Moreover, there is little to protect relatives who did not submit genetic samples to a database and may not even know that they could be identified by samples that have been submitted to one.

Beyond the threats that remain for subjects, the presence of an inconsistent set of policies could compromise the ability of researchers and businesses that rely on commercial genetic databases to plan long-term compliance. Were this uncertainty to lead them to avoid data-sharing arrangements or lead potential subjects to decline to submit genetic samples, the potential value of genomic data aggregation could be squandered. In the next section, we present our proposal for a regulatory mechanism to fill this gap.

V. DATA PROTECTION REVIEW BOARDS: PROPOSED STRUCTURE AND FUNCTION

Of all the legal gaps in genetic privacy protection, the most glaring is the lack of an external oversight mechanism for proprietary databases. To fill this void, we propose the creation of a universal system to oversee data-sharing by the companies that compile and maintain them. As discussed, it would take the form of a regulatory entity called a Data Protection Review Board ("DPRB") built upon the model of IRBs. The IRB model has already been applied to contexts beyond the protection of human subjects in government-sponsored research, such as requests by patients for compassionate-use access to experimental pharmaceuticals and requests by researchers for access to repositories of biospecimens, identifiable clinical trial data, and government-held demographic data. IRB-like review has also been proposed as a mechanism for overseeing research access to electronic medical records maintained by health systems.[240] While IRBs have had their share of criticism, as we discuss below, they have served for nearly half a century as the principal device to safeguard research subjects' interests and rights. Thus, they are a good foundation to work from in fashioning protections for genetic data privacy.

We do not propose that DPRBs replace IRBs, which would remain in place for federally funded research and clinical trials in support of FDA new drug applications. Rather, DPRBs would fill the gap left by the Common Rule for proprietary database companies that fall outside its scope. The nature of DPRB review would be narrower than that of IRBs, which evaluate all potential risks to subjects. DPRBs would focus on privacy risks and oversee all data-sharing arrangements that proprietary database companies engage in, including those for marketing or other business purposes. They would also bring to bear more specialized expertise in genomics and privacy than IRBs may have.

The scope of DPRB review would be further limited to the sharing of data with other private entities. DPRBs would oversee partnerships entered into as part of a proprietary database company's business operations. Disclosure to governmental authorities, including for law enforcement purposes, is usually

240. *See, e.g.*, I. Glenn Cohen & Michelle M. Mello, *Big Data, Big Tech, and Protecting Patient Privacy*, 322 JAMA 1141 (2019).

nonconsensual and is subject to legal rules, such as constitutional limits on search and seizure;[241] thus, it is beyond the scope of this Article.

Our proposal further incorporates elements of another layer of protection that applies to requests for access to many biomedical databases, Data Access Committees ("DACs").[242] These bodies exist to minimize threats to personal health information privacy by evaluating requests for access to all kinds of potentially identifiable biomedical information.[243] As with our proposed DPRBs, DACs apply specialized expertise in database management but they function only in the context of data access for research.[244] For proprietary databases, their role would be subsumed by the new review mechanism, which would function in all contexts, not just research.

As a foundation for the presentation of our DPRB proposal, the following three sections examine the nature of the two existing mechanisms, IRBs and DACs, on which our new review structure rests, and explain their limitations in the context of genetic databases.

A. IRBs: History and Purpose

IRBs rest on an ethical foundation set forth by the Code of Helsinki published in 1964.[245] This document emerged from the work of nations around the world to protect human subjects after revelations of the research abuses in Nazi Germany in which human subjects were used in gruesome and painful experiments, often with fatal consequences. The Code of Helsinki required that all research participation be voluntary as evidenced by the free consent of subjects who are fully informed of the risks.[246] Consequently, the primary focus of IRBs is to assure that subjects receive full and comprehensible information prior to participation, that they freely consent to participation, and that they can withdraw their consent at any time.[247]

The concept of using IRBs to enforce the principles of voluntary research participation based on full information grew out of proposals dating back to the mid 1960s.[248] In 1953, the NIH's Clinical Center implemented a model of group peer review for studies involving healthy volunteers.[249] In 1965, the director of

---

241. *See, e.g.*, Natalie Ram, *Genetic Privacy After* Carpenter, 105 VA. L. REV. 1357, 1384-85 (2019).

242. Mahsa Shabani et al., *Genomic Databases, Access Review, and Data Access Committees*, *in* MED. & HEALTH GENOMICS 29, 29-30 (2015) [hereinafter *Genomic Databases*].

243. *Id.*

244. *See id.*

245. *See* World Med. Ass'n, *World Medical Association Declaration of Helsinki Ethical Principles for Medical Research Involving Human Subjects*, 310 JAMA 2191 (2013).

246. *Id.*

247. *See* Christine Grady, *Institutional Review Boards: Purpose and Challenges*, 148 CHEST J. 1148 (2015).

248. *Id.* at 1150.

249. *Id.*

that agency proposed that all research be evaluated by a panel of peers for ethical compliance.[250] In 1966, the United States Public Health Service applied the concept as a requirement for all federal health research, although it was not well enforced.[251]

The National Research Act, enacted in 1974,[252] elevated the requirement for ethics review into a legislative mandate and applied the term "Institutional Review Board" to the review mechanism.[253] The law was a response to the latest in a string of scandals concerning abuses of human subjects. That scandal involved a research project known as the Tuskegee study, which observed the natural course of syphilis in a group of poor African-American men from whom antibiotic treatment was withheld without their knowledge or consent.[254] The law also established the National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research to study ways to reduce the risk of future abuses.[255] Its findings were published in 1979 as the Belmont Report.[256] They were incorporated into regulations that today comprise the Common Rule.[257]

The FDA issued its own regulations in 1981 for IRB review of clinical research used to evaluate new drugs.[258] Unlike the Common Rule, these rules also apply to studies that are entirely funded by private companies.[259] The rules prevent companies from using research findings to support an application to market a new drug unless the research has been subject to IRB review.[260]

The Common Rule requires that IRBs be registered with the OHRP, which monitors them for regulatory compliance.[261] Institutions that receive DHHS research funding, such as universities, hospitals, and research institutes, must provide assurance to the OHRP that they will comply with requirements for IRB review.[262] "The FDA requires registration of IRBs," and evidence of IRB review

---

250. *Id.*

251. *Id.*

252. National Research Act of 1974, Pub. L. No. 93-348, 88 Stat. 342.

253. ROBERT I. FIELD, HEALTH CARE REGULATION IN AMERICA: COMPLEXITY, CONFRONTATION, AND COMPROMISE 216-17 (2007).

254. DeNeen L. Brown, *"You've Got Bad Blood": The Horror of the Tuskegee Syphilis Experiment*, WASH. POST (May 16, 2017), https://www.washingtonpost.com/news/retropolis/wp/2017/05/16/youve-got-bad-blood-the-horror-of-the-tuskegee-syphilis-experiment/ [https://perma.cc/75B9-5429].

255. *Research Implications*, CTR. FOR DISEASE CONTROL & PREVENTION, https://www.cdc.gov/tuskegee/after.htm [https://perma.cc/5BWR-XMFN] (last updated Mar. 2, 2020).

256. THE NAT'L COMM'N FOR THE PROT. OF HUMAN SUBJECTS OF BIOMEDICAL & BEHAVIORAL RESEARCH, *supra* note 178.

257. 45 C.F.R. pt. 46 (2020).

258. 21 C.F.R. pt. 56 (2020).

259. *See* 21 C.F.R. § 56.102 (2020).

260. 21 C.F.R. § 56.103 (2020).

261. Grady, *supra* note 247, at 1150.

262. *See id.*

must accompany submission of data to the agency for approval of new drugs.[263]

When the Common Rule and FDA regulations were adopted, a large proportion of human subjects research involved clinical trials of new drugs, most of which were conducted at single sites.[264] Such studies lend themselves to a clear delineation of risks, since the conditions of testing are circumscribed and preclinical studies have typically been conducted to spot possible hazards. Over the years since then, the nature of human subject studies has expanded markedly.[265] Most clinical trials now take place at multiple sites, some including as many as 100 or more, and the sites may be located in several countries.[266] In such multi-site trials, each site is staffed by different personnel and draws from a different pool of patients.[267] Prior to revisions of the Common Rule in 2018, IRBs at each site were required to review the research protocols, making the job of overseeing the overall study considerably more complex.[268]

Clinical trials have also come to encompass different kinds of focus, often well beyond the effects of single pharmacologic interventions. Studies may now examine a greater range of physiological and psychological phenomena and explore a wider array of outcomes. Moreover, the indicators of interest may involve data from sources that are separate from the subjects themselves, such as tissue samples taken from them, social and behavioral observations of them, and their genetic profiles. In such studies, subjects might not even know that data about them were being used without an affirmative effort to inform them.

With these developments, the range of ethical concerns related to research has grown as well. Traditional clinical trials primarily raise issues related to whether subjects have been adequately informed of the risks of research and have consented to them voluntarily. Newer studies raise additional concerns involving hazards such as conflicts-of-interest by investigators who stand to gain financially from the research, uncontrolled access to sensitive information by third parties, and the security of data after they have been collected.[269] Therefore, to protect subjects, IRBs must consider a much broader range of threats.

B. IRBs and Large Database Studies

Among the more difficult challenges that newer kinds of research pose for IRBs is protecting subjects in studies that use large medical datasets. Advances

---

263. *Id.*

264. *See id.*

265. *Id.*

266. *Id.*

267. *See id.*

268. Under the revised Common Rule, cooperative research studies must rely on approval of a single IRB. 45 C.F.R. § 46.114 (2020).

269. *See, e.g.*, Moore v. Regents of Univ. of Cal., 51 Cal.3d 120, 174 (1990); *see also* Sheryl Gay Stolberg, *The Biotech Death of Jesse Gelsinger*, N.Y. TIMES (Nov. 28, 1999), https://www.nytimes.com/1999/11/28/magazine/the-biotech-death-of-jesse-gelsinger.html [https://perma.cc/3RVF-5FJC].

in computing technology enable researchers to manipulate enormous amounts of information on millions of subjects. Databases used in such research may include a wide range of indicators, including clinical test results, biometric features, performance on behavioral tests, and lifestyle factors, such as smoking and exercise. They may also include full or partial genetic profiles.

Although database studies do not pose direct threats of physical harm to subjects, they present substantial threats to privacy. Unlike single-site clinical trials, data may be accessible to hundreds of research staff members at dozens of institutions who could leak information, either inadvertently or deliberately. They may also be vulnerable to hackers when they are not adequately protected. Of particular concern, the threat of reputational, social, and economic harm may persist indefinitely if data are stored after a study has been concluded.

The task facing IRBs in protecting subjects from these harms requires an especially delicate balance. Database studies cannot proceed without data-sharing. In fact, the NIH strongly encourages researchers to exchange data in studies that it sponsors. IRBs that are overly aggressive in protecting subject privacy could constrain such sharing and impede important research.[270] However, those that are too lenient risk permitting harm that might discourage people from participating as subjects in future studies.

When data on patients are involved, the Privacy Rule provides that, in most cases, researchers may only access patient information that has been deidentified, under the assumption that this assures anonymity.[271] However, some studies require data that are linked to individual patients. For example, genetic research that seeks to link gene variants to clinical outcomes may need to identify patients who have the variants of interest. In these instances, the Privacy Rule permits the release of identified information to researchers with IRB review and approval.[272] Before permitting the research, an IRB must seek to assure that risks to privacy are minimized by requiring that researchers implement various safeguards concerning data access and storage.[273] For example, the amount of data collected must be the minimum needed for the research, anonymous identifiers must be used to link subjects to their data, and the identifiers and data must be deleted as soon as they are no longer needed.[274] Moreover, as discussed above, it is increasingly obvious that even with these protections much deidentified patient and subject data can be reidentified.[275]

---

270. *See, e.g.*, Amy A. Lemke et al., *Broad Data Sharing in Genetic Research: Views of Institutional Review Board Professionals*, 33 IRB 1 (2011).

271. Office for Civil Rights, *Summary of the HIPAA Privacy Rule*, DEP'T HEALTH & HUM. SERV., https://www.hhs.gov/hipaa/for-professionals/privacy/laws-regulations/index.html [https://perma.cc/CV6F-F3DA] (last updated July 26, 2013).

272. Office for Civil Rights, *Research*, DEP'T HEALTH & HUM. SERV., https://www.hhs.gov/hipaa/for-professionals/special-topics/research/index.html [https://perma.cc/5842-AAB7] (last updated June 13, 2018).

273. *Id.*

274. *See id.*

275. *See* Drabiak, *supra* note 167.

Use of patient information that may be identifiable in research may also be subject to a Certificate of Confidentiality, as discussed in Section IV above.[276] They are issued automatically for research that the NIH funds and on request for other research. They permit use of potentially identifiable information, such as biospecimens and individual-level genomic data, but only subject to specified safeguards against unauthorized disclosure.[277] However, there are exceptions to the safeguards when disclosure is required by federal, state, or local laws, is necessary for medical treatment, is authorized by an IRB, or is consented to by the patient.[278]

Studies have shown that the rigor of IRB oversight of genetic database research varies considerably.[279] For example, IRB members differ widely in their characterization of the ethical implications of genetic data and of the risk of subjects being identified.[280] This inconsistency suggests variability in the extent to which IRB members understand the nature of genetic databases and the threats they can present to subject privacy.[281] While such variability is not limited to comprehension of genetic information,[282] the complexity of genetic research coupled with its relative novelty may exacerbate gaps among IRB members in understanding this kind of information.

The 2018 revision of the Common Rule streamlined the process of review of privacy risks when identifiable data are used by permitting investigators to obtain "broad consent" for twelve aspects of research, rather than separate consent for each one.[283] Among these aspects are the possible use of biospecimens for commercial profit that would not be shared with subjects, the possibility that subjects' genomes may be sequenced, and the possibility that additional studies might be conducted with identifiable information about which subjects would not be informed. IRBs are merely required to enforce this general disclosure and need

---

276. *What is a Certificate of Confidentiality?*, NAT'L INST. HEALTH, https://grants.nih.gov/policy/humansubjects/coc/what-is.htm [https://perma.cc/AZ4W-7SV5] (last updated Jan. 15, 2019).

277. NIH regulations also require investigators to ensure compliance by other investigators and institutions involved in the research that are not funded by NIH as well as others involved in the study who have access to information subject to a Certificate of Confidentiality. They must also inform research subjects about the protections contained in a Certificate of Confidentiality as part of the informed consent process.

278. *See id.*

279. Claire Simpson et al., *Practical Barriers and Ethical Challenges in Genetic Data Sharing*, 11 INT'L J. ENVTL. RES. & PUB. HEALTH 8383, 8388 (2014).

280. Amy A. Lemke et al., *Attitudes Toward Genetic Research Review: Results from a National Survey of Professionals Involved in Human Subjects Protection*, 5 J. EMPIRICAL RES. ON HUM. RES. ETHICS 83, 87-88 (2010).

281. *See, e.g.*, Robert Klitzman, *The Myth of Community Differences as the Cause of Variations Among IRBs*, 2 AJOB PRIMARY RES. 24 (2011).

282. Christine Grady, *Do IRBs Protect Human Research Participants?*, 304 JAMA 1122, 1122 (2010).

283. *See* Federal Policy for the Protection of Human Subjects, 83 Fed. Reg. 28,497 (June 19, 2018).

not mandate that investigators implement specific protections for data once they have been collected.[284] For example, while investigators must inform subjects how long their information will be stored, that period may be indefinite; and, while they must identify people whom subjects can contact for questions concerning storage and use of information, no specific actions are required to be taken in response.[285]

Once broad consent has been obtained by a researcher, it can apply to use of a subject's data for follow-up studies.[286] The same consent can apply even if the

---

284. *See Elements of Broad Consent*, VAND. U. MED. CTR., https://www.vumc.org/irb/elements-broad-consent [https://perma.cc/4HUW-L8EA] (last visited Sept. 30, 2019) ("Required elements of broad consent: 1. A description of any reasonabl[y] foreseeable risks or discomforts; 2. A description of any benefits to the participant or others that may reasonabl[y] be expected from the research; 3. A statement describing the extent, if any, to which confidentiality of records identifying the participant will be maintained; 4. A statement that participation is voluntary, refusal to participate will involve no penalty or loss of benefits to which the participant is otherwise entitled, and the participant may discontinue participation at any time without penalty or loss of benefits to which the participant is otherwise entitled; 5. If appropriate, a statement that the participant's biospecimens (even if identifiers are removed) may be used for commercial profit and whether the participant will or will not share in this commercial profit; 6. If appropriate, for research involving biospecimens, whether the research will (if known) or might include whole genome sequencing (i.e., sequencing of a human germline or somatic specimen with the intent to generate the genome or exome sequence of that specimen); 7. A general description of the types of research with the identifiable private information or identifiable biospecimens. This description must include sufficient information such that a reasonable person would expect broad consent would permit the types of research conducted; 8. A description of the identifiable private information or identifiable biospecimens that might be used in research, whether sharing of identifiable private information or identifiable biospecimens might occur, and the types of institutions or researchers that might conduct research with the identifiable private information or identifiable biospecimens; 9. A description of the period of time that the identifiable private information or identifiable biospecimens may be stored and maintained (which could be indefinite), and a description of the period of time that the identifiable private information or identifiable biospecimens may be used for research purposes (which could be indefinite); 10. Unless the participant or their LAR [Legally Authorized Representative] will be provided details about specific research studies, a statement that they will not be informed of the details of any specific research studies that might be conducted using the participant's identifiable private information or identifiable biospecimens, including the purposes of the research, and that they might have chosen not to consent to some of those specific research studies; 11. Unless it is known that clinically relevant research results, including individual results, will be disclosed to the participant in all circumstances, a statement that such results may not be disclosed to the participant; and 12. An explanation of whom to contact for answers to pertinent questions about the storage and use of information and specimens, and to voice comments or concerns about participants' rights, and whom to contact in the event of a research-related harm.").

285. *See id.*

286. John W. Maloy & Pat F. Bass III, *Understanding Broad Consent*, 20 OCHSNER J. 81, 82 (2020).

subsequent research is still unspecified.[287] In other words, subjects can be asked to consent to indefinite storage of their data for purposes that are not known at the time they consent. The investigator is not required to deidentify data that have been stored in identifiable form or to recontact subjects for new consents.[288] The ongoing risks to privacy that might result are abundant.

The 2018 revision further loosens restrictions regarding use of clinical data by dispensing with the requirement for IRB review when HIPAA applies.[289] If an investigator has obtained consent for data disclosure under HIPAA, further review by an IRB is not needed.[290] An exception also applies if the data have been deidentified according to HIPAA regulations.[291]

There is a further exception for data that have been sent to a clinical repository for purposes other than research, such as for clinical improvement, if the repository is not supported by the NIH or one of the other agencies to which the Common Rule applies.[292] This exemption also applies if the data are released to a repository for research but the organization releasing them is not directly involved in research.[293] Most proprietary databases would fall within this category. In other words, they may share data with other private organizations without IRB oversight.

In sum, the revisions significantly lessen some privacy protections for genetic database studies. In particular, they assume that deidentification of data mitigates the risk of identifiability, an assumption that is clearly obsolete. Moreover, even if a patient were willing to accept the risk of reidentification for themselves, risks remain for the privacy of relatives.

C. Data Access Committees

A second line of defense for safeguarding subject privacy is Data Access Committees ("DACs").[294] They may be housed centrally within a research organization, within a consortium of organizations, or in individual study sites.[295] The NIH maintains its own centralized DAC system.[296] Access to the agency's database of genotypes and phenotypes, which stores results of studies it has

---

287. *See id.*

288. *See Elements of Broad Consent*, *supra* note 284.

289. *See* 45 C.F.R. § 46.104(d)(4)(iii) (2020); *see also How the Common Rule 2018 Updates Can Affect Your Research and Quality Improvement Strategies*, PROMETHEUS RES., https://www.prometheusresearch.com/common-rule-updates-2018/ [https://perma.cc/YLZ8-LW82] (last visited Nov. 14, 2020).

290. *See How the Common Rule 2018 Updates Can Affect Your Research and Quality Improvement Strategies*, *supra* note 289.

291. *See id.*

292. *See id.*

293. *Id.*

294. *Genomic Databases*, *supra* note 242.

295. *See* Shabani et al., *supra* note 193, at 507.

296. *Id.*

sponsored, is overseen by sixteen DACs that review requests for consistency with data use limitations.[297]

DAC oversight may involve various aspects of data-sharing arrangements.[298] These include the content of arrangements with outside researchers, the qualifications of those researchers, and the disposition of data after the conclusion of a study.[299] DACs are often in a better position than IRBs to assure privacy protection because their focus is narrowly directed to data access, as is the expertise of their members. However, DAC review is not subject to the same rules that govern IRBs, and the use of DACs is voluntary.[300] Moreover, DAC review is often decentralized, with multiple DACs overseeing different sites involved in the same study.[301] This risks inconsistency in oversight.[302] For example, a centralized European database, the European Genome-Phenome Archive ("EGA"), relies on hundreds of local DACs, with consequent inconsistency as well as lack of transparency.[303] Information on the operations of smaller DACs overseeing access to the EGA is especially difficult to obtain.[304]

Several international organizations have issued guidance on harmonizing DAC policies. One is the Biobanking and Biomolecular Resources Research Infrastructure, based in Austria, which offers a self-assessment tool for quality control of biobanking practices.[305] The Public Population Project in Genomics and Society, a consortium of biobanks based in Montreal, has published best practices for biobanking, including practices to protect patient privacy.[306] The Global Alliance for Genomics and Health, based in Toronto, has published a framework for the responsible sharing of genetic data.[307] However, none of these guidance documents includes a mechanism to enforce compliance.

Moreover, despite the proliferation of DACs and attempts to coordinate their role, there is no consistency in their relationships with IRBs or institutional ethics

---

297. *See id.*; *see also Genomic Databases*, *supra* note 242, at 31-32.

298. It has been proposed that data access committees form the basis for privacy protection for research databases. *See* Phaik Yeong Cheah & Jan Piasecki, *Data Access Committees*, 21 BMC MED. ETHICS art. 12 (2020).

299. *See Genomic Databases*, *supra* note 242, at 31.

300. *See* Shabani et al., *supra* note 193, at 508.

301. *See id.* at 507.

302. *Genomic Databases*, *supra* note 242, at 33.

303. *See* Shabani et al., *supra* note 193, at 507.

304. *See* EGA EUR. GENOME-PHENOME ARCHIVE, https://ega-archive.org/ [https://perma.cc/P6NG-LFHD] (last visited Nov 17, 2019).

305. *BBMRI-ERIC Launches Self-Assessment Survey for Biobanks and Sample Collections*, BBMRI-ERIC, http://www.bbmri-eric.eu/news-events/bbmri-eric-launches-self-assessment-survey-for-biobanks-and-sample-collections/ [https://perma.cc/L25N-XL9D] (last visited Sept. 20, 2019).

306. *See* Int'l Soc'y for Biological & Envtl. Repositories, *2012 Best Practices for Repositories Collection, Storage, Retrieval, and Distribution of Biological Materials for Research*, 10 BIOPRESERVATION & BIOBANKING 79 (2012).

307. *See* Bartha Maria Knoppers, *Framework for Responsible Sharing of Genomic and Health-Related Data*, 8 HUGO J. art. 3 (2014).

committees, nor are they subject to formal regulatory requirements similar to those for IRBs.[308] As a result, their actual functioning with regard to ethical concerns varies considerably. Some avoid performing independent ethics reviews altogether, some review ethics documents from the organization hosting the research, and some evaluate ethical issues on their own. This is a particular concern for studies that involve research collaborators scattered across different countries and data that are accessible globally.

D. Data Protection Review Boards: Concept and Operations

With these numerous limitations, the protection of research subjects' genetic data by IRBs and DACs is incomplete and inconsistent, especially with regard to data that are collected and held by proprietary databases. These companies can voluntarily constitute their own review bodies, but because they are not mandated to do so by law, there is no external oversight of the rigor of such reviews.[309] At the same time, their growing databases represent an increasingly valuable resource for research.

To fill this regulatory void, DPRBs would review all data access to commercial genetic databases by external organizations, other than law enforcement, for any purpose, whether research, drug development, or marketing. Their oversight would include reviewing and approving data-sharing arrangements, imposing and enforcing conditions to safeguard subjects' privacy, and performing ongoing monitoring of the arrangements to enforce those conditions.

The concept is to build on the IRB model to create a level of review that is decentralized and flexible, yet strong enough to effectively protect the interests of research subjects. Located within each organization that hosts research, IRBs can consider the environment within which individual studies are conducted. They can adjust their oversight accordingly. If their conditions are not met, they can withhold permission for a study to commence or withdraw permission once it has started.

DPRBs would similarly oversee data-sharing at the level of each database company. They would consider the individual needs and circumstances of the company and the nature of its database. They would impose conditions, as described below, and data-sharing could not commence or could be halted in the absence of compliance. External government oversight by the FTC or a new federal agency would enforce DPRB mandates in a manner similar to that of IRB enforcement by federal research funding agencies and the FDA.

However, we propose one significant organizational difference from IRBs. Rather than requiring a standing DPRB for all companies, they would be constituted for each data-sharing arrangement. Some companies may not have

---

308. Shabani et al., *supra* note 193, at 508-09.

309. *See also Research Ethics & Policy*, 23ANDME, https://research.23andme.com/research/#ethics [https://perma.cc/334R-EPDB] (last visited Sept. 20, 2020) (discussing how 23andMe is exemplary in this regard because it uses an independent external accredited IRB for its research).

enough data-sharing arrangements to justify a standing review body. Constituting separate DPRBs also reduces the risk of capture by the host organization, which has been a criticism of IRBs.

The following is a description of key structural and operational aspects of DPRBs. We acknowledge that modifications may be needed in practice. Our intent is to present a template on which details of a regimen of review can be built.

1. DPRB Functions

As with IRB oversight, DPRB review would consider the nature of each data-sharing arrangement so that conditions on sharing and methods of monitoring compliance could be tailored to the specific risks involved. In contrast to IRBs, which focus almost entirely on protecting the subjects themselves, the DPRB's purview would extend to risks to relatives of subjects who might be identifiable. In other words, while the substantive focus of DPRBs would be narrower, the scope of protection would be broader in their consideration of risks to a wider range of people.

Enforcement of the mandate for DPRB oversight would reside at the federal level. The FTC is an obvious candidate for such duty. It has a broad mandate to protect consumers in a range of circumstances under its authority to regulate "unfair or deceptive acts or practices in or affecting interstate commerce."[310] The agency has already begun using that authority to investigate privacy risks in online data-sharing and the practices of companies that receive, store, and share data on individuals. With regard to genetic information, it has issued advice to companies that sell genetic test kits as to types of practices that might be considered unfair or deceptive.[311] However, while the FTC has legal authority and expertise in this area, enforcement to date has been minimal, with only one enforcement action as of 2019, involving a company that made unsupported health claims about its tests and failed to follow its own stated data privacy policy.[312]

Effective FTC oversight of DPRBs would require explicit authority and direction from Congress. The agency would also likely need additional funding for this new responsibility. As the extent of genetic database sharing grows, enforcement responsibility might be transferred to a newly created agency with explicit jurisdiction over genetic research. The creation of such an agency has already been proposed in the form of a Digital Privacy Agency.[313] If such an

---

310. 15 U.S.C. § 45(a)(1) (2020).

311. Elisa Jillson, *Selling Genetic Testing Kits? Read On.*, FED. TRADE COMMISSION (Mar. 21, 2019), https://www.ftc.gov/news-events/blogs/business-blog/2019/03/selling-genetic-testing-kits-read [https://perma.cc/UY6S-ZGCH].

312. Ellen Wright Clayton et al., *The Law of Genetic Privacy: Applications, Implications, and Limitations*, 6 J.L. & BIOSCIENCES 1, 19 (2019).

313. Kate Cox, *New Bill Would Create Digital Privacy Agency to Enforce Privacy Rights*, ARS TECHNICA (Nov. 5, 2019), https://arstechnica.com/tech-policy/2019/11/new-bill-would-create-

agency were to be created, DPRB oversight and enforcement could be included in its scope of authority.

Federal authority to regulate genetic database privacy is clearly granted by the Constitution's Commerce Clause, which gives Congress the power to regulate economic activity that takes place across state lines.[314] Genetic database sharing is a quintessentially national and even international endeavor. We recognize that enacting the necessary legislation would require political consensus that may be difficult to achieve. However, as the extent of genetic data collection and dissemination grows, so will threats to privacy. Public unease over privacy intrusions could increase and lead to pressure for government protection that might alter the political dynamics.[315]

2. DPRB Structure

Ideally, DPRBs would operate independently of the organizations that maintain the databases that are subject to review. IRBs are constituted and operated by the organizations they oversee, and their membership often includes colleagues of the investigators whose research proposals they review. This has led to criticism that IRB oversight can tend towards leniency.[316] Concern about overly lenient oversight might be greater for review bodies housed within commercial organizations that operate on a profit-making basis. Arguably, the profit motive poses an especially significant corrupting influence.

As an alternative organizational structure, DPRBs could be housed in independent nonprofit organizations. They would convene review boards in response to requests from database companies and charge a fee to cover costs. There is precedent for the use of external review organizations in IRBs that are operated by for-profit corporations. These entities conduct reviews for organizations that lack the resources to conduct them on their own.[317] However, while their reviews have been defended as rigorous,[318] for-profit review organizations may be susceptible to conflict-of-interest concerns in the profit-making pressure they face to approve studies in order to encourage repeat business. For this reason, we favor the use of nonprofit organizations, for which

---

digital-privacy-agency-to-enforce-privacy-rights/ [https://perma.cc/8G5N-8P7Q].

314. U.S. CONST., art. I, § 8, cl. 3.

315. Sharon Begley, *Consumers Aren't Wild about Genetic Testing – nor Are Doctors*, STAT (Feb. 12, 2016), https://www.statnews.com/2016/02/12/consumers-arent-wild-genetic-testing-doctors/ [https://perma.cc/A379-W2WB].

316. Danah Boyd et al., Supporting Ethical Data Research: An Exploratory Study of Emerging Issues in Big Data and Technical Research (Aug. 4, 2016) (unpublished manuscript) (on file at https://datasociety.net/library/supporting-ethical-data-research-an-exploratory-study-of-emerging-issues-in-big-data-and-technical-research/).

317. *See* WCG IRB, https://www.wcgirb.com [https://perma.cc/BXP5-B28X] (last visited Nov. 22, 2020).

318. *See* Ezekiel J. Emanuel et al., *Should Society Allow Research Ethics Boards to Be Run as For-Profit Enterprises?*, 3 PLOS MED. art. e309 (2006).

the potential conflict is at least somewhat mitigated.

As an alternative to external operation of DPRBs, strict federal oversight of internal company DPRBs could ameliorate some of the risks of over-leniency. Such oversight could be centralized within whichever agency oversees DPRB operations. When research is involved, additional enforcement authority could reside in the agency that oversees the kind of research involved. For example, the FDA could reject noncompliant genetic database research in support of new drug applications. Similarly, the NIH could condition its research funding on DPRB compliance if an entity relies on data from a commercial database.

Whether housed internally or externally, accreditation of DPRBs by an external body could add an additional layer of assurance. Such a process already exists for IRBs through the Association for the Accreditation of Human Research Protection Programs ("AAHRPP"), which accredits IRBs that meet standards for membership, expertise, adequacy of procedures, and other factors.[319] Although accreditation is not required for IRBs by federal regulations, it signifies compliance with national standards and enhances the credibility of IRBs that achieve it and of the institutions for which they conduct reviews. It has been found to improve the performance of IRBs, in part because of the training requirements it mandates for IRB members.[320]

The membership of DPRBs would be prescribed through regulations in a similar manner to that of IRBs. FDA regulations, for example, require that IRBs overseeing clinical trials have at least five members. At least one member must have primary concerns in the scientific area involved in the study under review, one must have primary concerns that are not scientific, and one must be unaffiliated with the institution conducting the research and not be an immediate family member of someone who is.[321] The IRB must also strive for demographic diversity and include members who are capable of reviewing the specific area of research involved, who are familiar with applicable law, and who understand standards of professional conduct and practice. Regulations governing IRBs issued by the NIH and other federal funding agencies are less specific concerning composition, requiring only that both scientific and nonscientific perspectives be represented.[322]

Regulations for DPRB membership should prescribe that it includes experts in genetics, cybersecurity, data analytics, privacy, bioethics, and relevant aspects of law. In addition to representing the range of relevant expertise, they would reflect the range of relevant interests related to genetic data-sharing. These

---

319. *See* AAHRPP, http://www.aahrpp.org/ [https://perma.cc/X9KQ-8BYP] (last visited Sep 25, 2019).

320. *See* Abhidnya Vasant Desai et al., *Role of Accreditation in Quality Improvement of Institutional Review Board*, 8 PERSP. CLINICAL RES. 145 (2017).

321. 21 C.F.R. § 56.107 (2020).

322. The Sec'y's Advisory Comm. on Human Research Prots., *Attachment B: Recommendation on IRB Membership and Definition of Non-Scientist Under 45 CFR 46 and 21 CFR 56*, HHS.GOV, https://www.hhs.gov/ohrp/sachrp-committee/recommendations/2011-january-24-letter-attachment-b/index.html [https://perma.cc/T8N5-TW3Y] (last updated Jan. 24, 2011).

include subjects whose information is contained in a genetic database, biomedical researchers who analyze genetic databases, and companies that use the information.

An ideal size for DPRBs would likely be between ten and fifteen members, although the number might be adjusted based on experience. The average IRB size is 13.9 members.[323] However, the size might vary with the size of the workload.

To ensure that they have appropriate expertise, individuals who wish to serve on a review board could apply to the agency for certification of eligibility, subject to verification of their credentials. Database companies needing the services of a DPRB could select review board members from among those who have received certification. This procedure goes beyond the requirements for IRB membership, but the review of genetic disclosure risks requires more specialized expertise.

3. Resources for Operations

Clearly, organizing and operating DPRBs requires resources. They present a significant administrative burden and need a commitment of time from experts, whether they are employees of database companies or outsiders. These resources are supplied to IRBs and DACs by the organizations involved. However, the cost of administering DPRB review could create a disincentive to data-sharing by smaller commercial organizations, thereby discouraging valuable initiatives.

A solution would be for DPRBs to be supported by the federal agency that oversees their operation. The agency could provide compensation to the members and cover operational expenses, such as travel and meeting costs. A user fee paid by the organization operating the DPRB could help to defray the cost in a similar manner to the user fees paid by pharmaceutical companies for FDA review of new drug applications.[324] However, unlike FDA user fees, DPRB fees could be adjusted to the size and resources of the organization involved.

4. The Nature of DPRB Reviews

The scope of DPRB reviews would be limited to privacy risks. Consideration of other kinds of research risks would be left to IRBs when applicable research is involved. For example, a study in support of a new drug application to the FDA or one that is funded by a federal agency that involved both data-sharing and an invasive medical procedure would be subject to both forms of review.

Given current technologies for reidentification of genetic data, no set of safeguards can assure anonymity of database subjects. However, a number of measures can significantly mitigate privacy risks, and DPRBs could mandate the use of some or all of such measures. Deidentification of data, although only a

---

323. Joseph A. Catania et al., *Survey of U.S. Human Research Protection Organizations: Workload and Membership*, 3 J. EMPIRICAL RES. ON HUM. RES. ETHICS 57, 64 (2008).

324. *See* 21 U.S.C. § 379h (2020).

minimally effective step, would be a start. Other solutions could include safeguards that the Privacy Rule directs IRBs to consider for research with identified patient information.[325] The objective and guiding principle would be that the arrangement for the sharing and use of data must present no more than "minimal risk" to subjects' personal health information and genetic privacy.[326] Among the most important safeguards applicable to DPRB oversight would be an adequate plan to protect identifiable information from improper use and disclosure, an adequate plan to destroy information that could link data to subjects' identities as soon as possible after the data-sharing arrangement has been concluded, and adequate written assurances that data will not be reused or disclosed to others except as required by law or as necessary for research oversight.

Additional requirements could be added to respond to the distinctive risks of genetic data disclosure, including threats to the privacy of subjects' genetic relatives.[327] These could include:

1. Further sharing of data with third parties beyond the arrangement under review would require additional DPRB approval.
2. All data that are shared must be encrypted.
3. Any electronic transmission must be through secure servers.
4. Individuals who have access to data must be vetted for trustworthiness and technical competence.
5. Records must be kept of individuals who have accessed data. Those records must be retained for a specified minimum amount of time after an arrangement has ended, perhaps seven years, which is the current standard for patient medical records.
6. Records must be available for review by the DPRB and representatives of the federal agency overseeing its operations.
7. Data that are shared would be limited to those that are needed for the arrangement, and entire genomes would only be shared when necessary.

In addition, DPRBs could mandate the use of various technical measures that have been developed to reduce the risk of data identifiability. Some of them are complex and require substantial technical expertise to implement, so DPRBs

---

325. *What is a Certificate of Confidentiality?*, *supra* note 276.

326. In a technological environment that is as fast-changing and evolving as that in which genetic analysis is now being coupled with widespread data-sharing, it is neither wise nor possible to articulate a precise standard. The regulatory entities must be given a reasonable latitude of discretion to tailor the regulatory regime to the rapidly changing realities of internet communication and the evolving nature of multi-party cooperation in genetic research.

327. *How Can Covered Entities Use and Disclose Protected Health Information for Research and Comply with the Privacy Rule?*, NAT'L INST. HEALTH, https://privacyruleandresearch.nih.gov/pr_08.asp [https://perma.cc/6D7Z-AHLB] (last updated Feb. 2, 2007).

would have to include appropriately trained experts.[328]

After determining the applicable safeguards, DPRBs would oversee data-sharing arrangements as they progress to monitor adherence. As part of the monitoring, the database company involved would submit regular reports to certify compliance. At the end of the arrangement, the company would certify that all data have been deleted or destroyed by the partner organization.

The stringency of DPRB-mandated protections might vary according to the intended use of the data. Data-sharing for marketing or similar business activities that only benefit private parties have the least public value, and the most rigorous safeguards would apply. Privacy protections for arrangements that facilitate biomedical research with the potential to help large numbers of people would be weighed against societal value. The mission of DPRBs in both cases would be to consider the appropriate balance.

5. Enforcement

IRB review is largely self-enforcing. Studies that receive federal support lose that support if it is not properly implemented.[329] If a federal agency has withdrawn support from an institution's research for noncompliance, it may deny funding to that institution for future research. Studies of new drugs that lack IRB review may not be used in support of new drug applications to the FDA.

We propose a two-prong approach to enforcement of DPRB review. The first would empower the agency administering the DPRB program to impose fines on organizations that fail to comply with review requirements. The second would impose liability under consumer protection laws for violations. This would subject violators to additional enforcement authorities and associated penalties.

For data-sharing arrangements that lead to publishable research, peer-reviewed journals could adopt DPRB review as a standard for publication, as many do for IRB review.[330] The International Committee of Medical Journal Editors is considering a standard that would require sharing of clinical trial data

---

328. There are several technical measures that can increase the difficulty of identifying individuals whose information is contained in a large database. *See* Yaniv Erlich & Arvind Narayanan, *Routes for Breaching and Protecting Genetic Privacy*, 15 NATURE REV. GENETICS 409 (2014). These measures include: (1) removing metadata and other identifiers that may accompany genetic data; (2) keeping data in a secure location and limiting access to those with a demonstrated need to access it under the data-sharing arrangement and who have been vetted for reliability; (3) denying users direct access to data and permitting only queries to the database; (4) using a trusted agent or third party to manage any interactions between subjects and data users; (5) using a data anonymizing methodology known as "k-anonymity"; (6) differential privacy; (7) data encryption; (8) secure multiparty computation; and (9) one-way encryption. *Id.* Some of these measures might also be mandated by a DPRB for a specific data-sharing arrangement.

329. 45 CFR § 46.122 (2020).

330. *See* Robert J. Amdur & Chuck Biddle, *Institutional Review Board Approval and Publication of Human Research Results*, 277 JAMA 909 (1997).

before the results could be published.[331] Such a standard could be extended to require that when genetic data are shared, DPRB review must be applied.

6. Limitations

We acknowledge that our proposal has limitations. As with IRBs, DPRBs add time and administrative complexity to the process of exchanging data. If DPRBs are not adequately resourced, delays could become unmanageable. The burden of complying with this added layer of bureaucracy could thereby impede valuable biomedical research and other socially beneficial activities.

However, all regulatory schemes add cost and complexity. This is inherent in the nature of inserting third-party oversight into a commercial activity, but it is not reason to avoid reasonable regulatory measures. When a risk is substantial, the cost of inaction can be greater than the burden of added bureaucracy. We believe that implementation of a system of DPRBs represents a reasonable compromise between direct centralized regulation that could stifle much valuable medical innovation and a *laissez faire* approach that leaves data subjects and their relatives at risk. IRBs, on which our proposal is modelled, represent a similar compromise that has achieved a sustainable balance for almost half a century.

D. Alternative Approaches

We see our proposal as the most effective of several remedies that have been enacted or proposed to address electronic data privacy. Alternative schemes do not go as far in addressing individual instances of data-sharing, and most do not address the unique risks of collecting and storing genetic information. Moreover, none address the interests of subjects' relatives.

An alternative to a unified nationwide program might be found in state-by-state legislation along the lines of the CCPA.[332] As with other areas of regulation, such a federalist approach could be more responsive to local values and attitudes than federal legislation.[333] However, state regulation of data-sharing for databases that contain information on millions of subjects would be difficult to enforce. Companies respond to directives by the EU and large states such as California because of the large markets they represent. Smaller states would have considerably less influence. State regulation could also present jurisdictional conflicts and gaps, as restrictions on use of a subject's data in one jurisdiction may provide little protection to a relative who lives in another. Finally, state

---

331. Darren B. Taichman et al., *Data Sharing Statements for Clinical Trials – A Requirement of the International Committee of Medical Journal Editors*, 376 NEW ENG. J. MED. 2277, 2277 (2017).

332. *See* discussion *supra* Section IV(B)(4).

333. New State Ice Co. v. Liebmann, 285 U.S. 262, 311 (1932) ("It is one of the happy incidents of the federal system that a single courageous state may, if its citizens choose, serve as a laboratory; and try novel social and economic experiments without risk to the rest of the country.").

regulation is problematic in cases where data are shared among entities in different states, which is increasingly the norm.

Moreover, state-by-state legislation would pose significant logistical difficulties for compliance. One survey of existing state and federal genetic privacy laws found considerable variability in their nature and scope.[334] It also found them to be limited in their application, focusing primarily on data-sharing for research and clinical care, rather than on information that is voluntarily submitted to private companies for ancestry tracing or other purposes. Were a patchwork of more stringent state laws to emerge that apply to all data-sharing, compliance could become impractical, thereby creating a barrier to innovation.

The Government Accountability Office has proposed the enactment of new federal legislation expanding the FTC's enforcement authority over data privacy.[335] Some commentators have proposed creation of a new federal agency to regulate sharing of online personal data. Stronger centralized oversight along these lines might be more efficient than adding a system of decentralized review boards, such as DPRBs, and might be less susceptible to capture by data companies and their business partners. However, any gains in efficiency would be lost in flexibility. Centralized regulation would operate at a distance from the parties that are subject to its oversight and is therefore likely to be less responsive to their needs and circumstances. A large new federal bureaucracy might also act slowly, introducing delays that could impede some valuable data-sharing arrangements.

VI. CONCLUSION

Almost 2,500 years ago, Hippocrates advised physicians to promise to protect, "[w]hat [they] may see or hear in the course of the treatment or even outside of the treatment in regard to the life of men, which on no account one must spread abroad."[336] That advice is at least as relevant today as it was centuries ago. Just as physicians must be privy to confidential information to effectively treat individual patients, genetic databases must contain sensitive personal information to facilitate activities such as biomedical research that can be used to treat many. Trust in medical professionals has always been crucial to the medical enterprise. Today, that trust must extend to those who administer technologies that advance medical care.

We believe that a system of oversight of genetic database sharing in the form of DPRBs strikes a reasonable balance to facilitate genetic research while protecting individuals, both those who submit data and those whose identities can

---

334. Wolf et al., *supra* note 170, at 61-62.

335. *See generally Internet Privacy: Additional Federal Authority Could Enhance Consumer Protection and Provide Flexibility*, U.S. GOV'T. ACCOUNTABILITY OFF. (Jan. 2019), https://www.gao.gov/assets/700/696437.pdf [https://perma.cc/36ZZ-9DBQ].

336. John C. Moskop et al., *From Hippocrates to HIPAA: Privacy and Confidentiality in Emergency Medicine—Part I: Conceptual, Moral, and Legal Foundations*, 45 ANNALS EMERGENCY MED. 53, 53 (2005).

be revealed through the data of relatives. On the one hand, research with large databases is essential to the advance of biomedical science. It lies at frontiers of innovation in areas such as gene therapy, precision medicine, and artificial intelligence. On the other hand, without privacy safeguards, individuals may hesitate to submit data that are subject to sharing. Such reluctance risks killing the goose that is giving us golden eggs of knowledge, not to mention creating the risk of widespread harm to individuals. Third-party expert oversight and guidance regarding the sharing of genetic data is an important and necessary step toward balancing those concerns.

IRB oversight has been effective in reducing, albeit not eliminating, risks in human subjects research. That model has the advantages of flexibility to accommodate different circumstances and of bringing different expert perspectives to bear. We can take that model one step further to the next frontier in biomedical research to build the trust that genetic database investigations require.

We have no time to lose. Whether through DPRBs or some other mechanism, it is imperative to act quickly before more data on individuals are released without any way to retrieve them. Moreover, the anonymity of genetic data will only become more tenuous over time. If privacy concerns discourage potential subjects from contributing genetic samples, it will not be responsible regulation that impedes innovation but rather the lack of it.