Increasing Use of Genetic Data Requires New Privacy Considerations

Pardeep Singh[1,2], Alexa McIntyre[3], Sean O'Connor[4], Kaitlyn Scacalossi[5], and Amanda Acosta-Ruiz[3]

[1]Columbia University, 116th and Broadway, New York, NY 10027
[2]City University of New York, 205 East 42nd Street, New York, NY 10017
[3]Weill Cornell Medicine, Department of Physiology and Biophysics 1300 York Ave, New York, NY 100213
[4]Rockefeller University, 1230 York Ave, New York, NY 10065
[5]New York University School of Medicine, Department of Immunology and Inflammation, 550 First Ave, New York, NY 10016
Corresponding author: ps2806@columbia.edu
Keywords: consumer genetic testing, genetic data, privacy, GINA, DNA testing, consumer protection.

Executive Summary: Genetic testing services should adopt a standardized privacy policy that prohibits sharing of personal genetic information with third parties without explicit consent. The Genetic Information Nondiscrimination Act of 2008 (GINA) should be amended to include broader protections from genetic discrimination and regulated access to genetic databases by law enforcement and commercial third parties.

I. Statement of purpose

Advances in technology have opened a new frontier in consumer genetic testing. Current laws are outdated and inadequate, allowing these businesses to operate without the regulations needed to protect sensitive genetic data information. It is necessary to extend the current protections of the Genetic Information Nondiscrimination Act of 2008 to prevent misuse of personal genetic data.

II. Historical background

The Genetic Information Nondiscrimination Act (GINA) was lauded as the first new civil rights bill of the century. Passed in 2008 under the George W. Bush administration, GINA prohibits employers and health insurers from requesting genetic information or requiring genetic testing to make substantive decisions about employment or insurance coverage. GINA's protections enable individuals to inquire about their genetic predispositions to disease without fear of discrimination from employers or health insurance providers (United States Congress 2008). Since 2008, technological advances have led to the proliferation of companies offering genetic testing services and the accumulation of genetic data by law enforcement. There is now more indexed genetic information available in both public and private domains than at any point in history. As of February 2019, 26 million people have sought information about their genealogies and disease predispositions through consumer genetic testing services, and 100 million more are expected to use these services within the next two years (Regalado 2019). Overall, spending on genetic testing is projected to increase by $5-$25 billion in the next decade (Ancestry.com 2019), creating a massive amount of data with few legal protections at the federal level and variable regulations by state.

III. GINA's current statutes are insufficient

As the technology underlying genetic testing has become faster, cheaper, and more accurate, the potential for misuse of genetic data has increased significantly. The increased use of genetic testing services has not been matched by commensurate public awareness of privacy risks or disclosure by genetic testing companies. For example, a 2014 survey estimated that only 21% of the general public is familiar with GINA and that 23% of those who are

overestimate the scope of its protections (Green 2015, 397-399). While consumers may agree to a particular privacy policy or lack thereof, the shared nature of DNA means that they are essentially agreeing on behalf of their extended family. A recent study found that approximately 60% of European Americans had an identifiable third cousin or closer relative in a database of consumer DNA test results (Erlich 2018). Without stronger privacy protections, the vast majority of people may be traceable within the next decade, regardless of whether they choose to undergo DNA testing.

One issue of particular concern is the ability of private companies to share genetic information with law enforcement. Currently, the privacy policies of Ancestry and 23andMe specify that they will only release information to law enforcement with a court order. However, without explicit legal protections, privacy agreements vary among companies. Indeed, some genetic data platforms have looser policies that allow law enforcement to access data without a court order. One such platform is GEDmatch, which allows users to upload their genetic data to find relatives. Recently, detectives used GEDmatch to identify the Golden State Killer by uploading the test results from a crime scene DNA sample (Aldhous 2018). This allowed law enforcement to identify distant relatives of the killer and narrow down their search. In a separate case, law enforcement charged a man with murder in a 25-year-old cold case after investigators harvested his DNA from a discarded hot-dog napkin and matched it with DNA found on an open source genealogy website (Mervosh 2019). Although most people may agree that identifying killers and rapists is in the public's best interest, genetic databases have also been used to track down non-violent criminals. In 2017, for example, over 30 cases used genetic data provided by Ancestry to investigate credit card misuse and identity theft (Vincent 2018).

Beyond law enforcement, there is also little to no regulation regarding sharing or selling genetic information to third parties. While GINA prevents insurance companies and employers from using genetic data to discriminate, it does not regulate the transfer of genetic data to commercial or research institutions. Helix, a genetic testing site, has over 25 commercial partners, including startups and large healthcare providers, with whom they share customers' data (Vincent 2018). Their company privacy policy requires separate consent for each partner before any data is shared. 23andMe, on the other hand, has a privacy policy that allows the company to share data with any number of partners after getting a single affirmation of consent. Using this broad form of consent, 23andMe signed a $300 million deal with the pharmaceutical giant GlaxoSmithKline in 2018 for commercial research (Molteni 2018). Still a privately held company, these multi-million dollar deals monetizing genetic data by 23andMe are not subject to shareholder scrutiny or public reporting requirements.

When companies share individual data, they almost always do so in an anonymized format; however, many anonymized samples can be re-identified (Tanner 2013). Currently, there are no regulations that govern the format in which anonymized data is shared to limit re-identification. With the exception of the types of employment and health insurance discrimination prohibited by GINA, voluntary company privacy policies are presently the only guarantee of a consumer's privacy and anonymity and can legally be revoked or changed at any time. The for-profit nature of consumer genetic testing services means that these companies may generate privacy policies contrary to the public interest in the absence of further regulation.

GINA's limitations beyond employment and health insurance protections have allowed genetic discrimination to occur in other contexts. In 2016, a 36-year-old woman submitted her genetic test results to a life insurance company only to be denied coverage after learning she was positive for BRCA1, a breast cancer risk gene (Lalley 2018). Although the presence of this gene is by no means a death sentence, this form of genetic discrimination is legal for long-term care, disability, life, and all other insurance not included in GINA, deterring individuals from undergoing potentially beneficial genetic testing. Instances of genetic discrimination have also occurred sporadically in education. In 2012, a middle school student was removed from his classroom after testing positive for a cystic fibrosis risk gene out of concern that he would prove contagious to other students with cystic fibrosis at the school (Zhang 2017). Healthy individuals are at no risk of contagion from cystic fibrosis patients, and as he did not develop symptoms of the disease, the student posed no risk to students with cystic fibrosis either. In that

case, the results of a neonatal DNA test administered to the student to diagnose a cardiac issue were shared as part of the school's admissions process and disseminated to teachers and parents, resulting in the student's dismissal. The potential uses of genetic information in guiding criminal prosecutions and access to insurance, education, and health information require increasing public understanding of the risks and responsible policies to limit the misuse of genetic data.

IV. Attempted roll back of GINA's workplace protections

Despite the limited scope of GINA's protections, Republicans in Congress are leading efforts to roll back existing regulations. An amendment to GINA approved along party lines in House committee vote H.R.1313 explicitly states that GINA's genetic privacy protections do not apply when genetic tests are part of "workplace wellness" programs. H.R.1313, led by Rep. Virginia Foxx (R), is scheduled to be heard by the full House of Representatives after passing three individual House committees in 2017 (Foxx 2017). These programs use genetic testing to design employee insurance plans with lower premiums. While voluntary, employees who do not participate pay thirty to fifty percent more for their health insurance, although the financial gains to employers are marginal at best (Begley 2015 and 2017; Lewis 2017). This is effectively a financial penalty that shifts healthcare costs to workers who do not forfeit GINA's workplace protections, while companies only save $25-$40 in healthcare costs per participating employee, according to a study from the Research and Development (RAND) think tank commissioned by Congress (Mattke 2013). Thus, the proposed amendment decreases privacy protections for workers while generating minimal economic benefit.

V. State-level expansion of GINA

GINA's protections should be amended to include securities provisions for misuse of genetic data, consent of genetic data, rights to anonymity, and broadened protection from genetic discrimination. California state lawmakers have been champions on this front by expanding GINA's protections to cover emergency medical services, housing, mortgage lending, education, life insurance, elections, and state-funded programs. Passed in 2011, "CalGINA" can serve as a model for nationwide modernization of genetic privacy and nondiscrimination laws (Zhang 2017). So far, though, the limits of CalGINA have not been tested in court.

An exploratory case in California's 9th Circuit Court, Chadam v. Palo Alto Unified School District, claims that a student's genetic information was identified as a "perceived disability" and used to remove the student from school (Wagner 2018). The plaintiffs in this case chose to sue under the Americans with Disabilities Act of 1990 (ADA), rather than relying on CalGINA, in the hopes of setting a federal precedent. If victorious, this would be the first time courts acknowledged a person's genotype or carrier status as sufficient to pursue a claim under the ADA. The ADA fills a gap in GINA by covering those who show symptoms of genetic disorders (Ellen 2015, 2225-2226). However, it was not written with presymptomatic genetic testing in mind, and the success of the current case is far from certain. Updating GINA, which deals specifically with genetics, would provide stronger and more explicit protection against this form of discrimination nationwide.

VI. Current law in foreign countries

The European Union (EU) goes even further than CalGINA or the ADA by including genetic data under the protection of the General Data Protection Regulation (GDPR), implemented in 2018. The GDPR classifies genetic information as "personal data," and gives EU citizens the right to have any personal identifiable information anonymized, erased, and shared only with explicit consent. Unlike the United States, which has a patchwork of state laws governing privacy, the GDPR unifies the EU's privacy laws and streamlines reporting requirements and penalties for violating the GDPR. It requires organizations to report data breaches to both the affected individuals and the appropriate regulatory authorities within 72 hours of being discovered and, importantly, penalizes companies up to 4% of their global revenue. This unified approach has not only raised more than 55 million euros in revenue but was also a window into the EU's security lapses (Wolff 2019). The GDPR also allows exemptions for scientific research where individual consent is not practical for studies that require large amounts of data to reach statistically significant findings (Drechsler 2017).

On the other extreme, the Chinese government has been building a database of DNA samples collected

from its minority Muslim populations under the guise of free health care screens (Peryer 2019). In some instances, these screens were mandatory and led to the detainment of hundreds of thousands of Muslim Chinese in Xinjian camps for, what the Chinese government calls, a way to escape poverty, backwardness, and radical Islam (Wee 2019). This genetic surveillance was unknowingly facilitated by American geneticists at Yale University, who collaborated with Chinese government scientists by providing access to a data base that included genetic data for people from around the world in exchange for data for over 2000 Muslim Chinese minorities (Peryer 2019). These privacy violations drew criticism from the scientific community and prompted biotech companies to sever ties with China by refusing to sell any more genetic sequencers (Khan 2019). As the Chinese sequencing company BGI expands operations, the withdrawal of American companies will likely do little to hold off the government's sequencing program (Robbins 2018).

While the EU and Chinese governments have implemented vastly different approaches for managing genetic data, the United States has yet to commit to a unified path for enforcing genetic privacy. We propose strengthening GINA's existing protections to be more in line with the EU's GDPR and CalGINA, as outlined in option 3 below.

VII. Policy recommendations options

i. Option 1. Allow the status quo to continue.

Companies would be free to set their own privacy and data sharing policies.

Advantages

Access to large-scale genetic information allows the biopharmaceutical industry to leverage patient data to drive therapeutic development more quickly and accurately. Law enforcement agencies also have the ability to use consumer data to solve serious crimes without regulatory impediments. Genetic testing companies can continue to offer cheaper and more accessible services by monetizing the data they collect through deals with pharmaceutical companies and other third parties.

Disadvantages

Consumers lack protections and may not understand the risks of sharing their genetic data. There are already clear examples of genetic discrimination by insurance companies and schools. The risk of abuse by law enforcement is heightened when genetic data are accessible without a court warrant. A lack of protections may discourage consumers from undergoing genetic testing and thus potentially decrease their ability to make informed health decisions.

Option 2. Allow sharing of anonymized genetic information by genetic testing services only for scientific research where bulk samples are necessary (N>100).

In circumstances where large sample sizes are not necessary to reach statistical significance, the sharing of genetic information by genetic testing services is prohibited.

Advantages

Smaller scientific studies (N<100), especially those about rare gene variants, can be more readily traced back to specific afflicted individuals despite being anonymized. Smaller studies can, more practically, seek individual consent while larger studies can continue to exploit the massive amounts of data generated through consumer genetic tests.

Disadvantages

Genetic screens in smaller scientific studies may become slower and more costly because seeking out individual consent can be a laborious and legally arcane process. Bulk samples from large studies can still be de-anonymized.

Option 3. Promote CalGINA's protections to the federal level, standardize genetic testing privacy policy for all genetic testing services and ensure penalties for de-anonymizing data without consent.

Standardized privacy policy should include guaranteed anonymity, a requirement for consumer consent each time data is shared with a third party, and no access to an individual's genetic data by law enforcement without a court order. De-anonymizing data, like in the GDPR, without consent should be recognized as a crime. Expanding these protections would allow consumers to share their genetic test results without discriminatory treatment when accessing emergency medical services, housing, mortgage lending, education, life insurance, elections, or government-funded programs.

Advantages

A nationally standardized and unambiguous privacy policy would lessen the likelihood that commercial institutions and law enforcement agencies could abuse individuals' data. Furthermore, a strict policy would enable users to feel safe and secure when using genetic testing services. This consumer trust could increase usage in a market that is already projected to grow up to $25 billion while also promoting public awareness of personal genetic identity and misuse of genetic data. Introducing fines or penalties for unlawful de-anonymization may deter unauthorized third parties from accessing personal genetic data.

Disadvantages

Genetic data has proven useful for solving violent crimes and decreasing law enforcement access could hinder efforts to modernize crime-solving. Additionally, limiting access for commercial and nonprofit research institutions may slow progress in both basic and applied sciences. Increasing regulation may lead to a commensurate rise in the cost of genetic testing and decrease its accessibility, particularly if companies' business models shift from profiting more off of data to largely profiting off selling testing services.

VIII. Recommendation

Our recommendation is to approve option 3. As medical technology and law enforcement analytics progress through the twenty-first century, it is becoming increasingly clear how valuable, and potentially dangerous, individual genetic data can be. Once data is shared, with or without full consent, it is often difficult or impossible to recover. Compared to options 1 and 2, strict consent requirements on sharing genetic information with research organizations may lead to slower scientific progress in certain areas; however, we believe the ability to retain privacy and security over one's personal data is of paramount importance. In addition, there are other avenues by which researchers can acquire patient genetic data without putting customers of genetic testing platforms at risk.

While requiring a court order to access data from genetic testing platforms may frustrate the efforts of law enforcement in certain cases, high standards for access are necessary to protect citizens' civil liberties. With the status quo presented in options 1 and 2, innocent citizens may have their privacy violated by the decisions of a distant family member who freely shares their own genetic data. By requiring a court order, option 3 provides ample opportunity for law enforcement to use genetic data when sufficient evidence is present to acquire a court order, while minimizing the potential for abuse. Access to modern genetic sequencing technology has created a new stratum of personal and private genetic data that should receive the same protections as any personally identifiable information.

References

Aldhous, Peter. "DNA Data From 100 Crime Scenes Has Been Uploaded To A Genealogy Website Just Like The Golden State Killer." BuzzFeed News. May 16, 2018. Accessed February 27, 2019.

Ancestry. "Ancestry Guide for Law Enforcement." Accessed February 27, 2019. https://www.ancestry.com/cs/legal/lawenforcement.

Associated Press. "US Genetics Company Will No Longer Sell Tech to China." New York Post. New York Post, February 21, 2019. https://nypost.com/2019/02/21/us-genetics-company-will-no-longer-sell-tech-to-china/.

Begley, Sharon. "Coming Soon to a Workplace near You: 'wellness or Else'." Reuters. January 13, 2015. Accessed February 25, 2019. https://www.reuters.com/article/us-usa-healthcare-wellness-insight-idUSKBN0KM17C201 50113.

Begley, Sharon. "House Republicans Would Let Employers Demand Workers' Genetic Test Results." Scientific American. March 10, 2017. Accessed February 25, 2019. https://www.scientificamerican.com/article/house-republicans-would-let-employers-dema nd-workers-rsquo-genetic-test-results/.

Drechsler, Laura. "The Implications of the GDPR for Research Involving GeneticData." October 5, 2017. https://brusselsprivacyhub.eu /publications/ws11.html

Ellen Wright Clayton. "Why the Americans With Disabilities Act Matters for Genetics." JAMA, 2015 Accessed February 25, 2019. https://www.researchgate.net/publication/278046299_Why_the_Americans_With_Disabilities_Act_Matters_for_Genetics

Erlich, Yaniv, Tal Shor, and Shai Carmi. "Identity Inference of Genomic Data Using Long-range Familial Searches." Science. November 09, 2018. Accessed

February 28, 2019. http://science.sciencemag.org/content/362/6415/690.

Foxx, Virginia. "H.R.1313 - 115th Congress (2017-2018): Preserving Employee Wellness Programs Act." Congress.gov, December 11, 2017. https://www.congress.gov/bill/\115th-congress/house-bill/1313.

Green, R. C., Lautenbach, D., & McGuire, A. L. (2015). GINA, Genetic Discrimination, and Genomic Medicine. New England Journal of Medicine, 372(5), 397–399. https://doi.org/10.1056/NEJMp1404776

Lalley, Colin. "How Genetic Testing Can Affect Your Life Insurance Rates." Policygenius Magazine. March 19, 2018. Accessed February 25, 2019. https://www.policygenius.com/blog/genetic-testing-life-insurance-rates/.

Lewis, Al, and Vik Khanna. "Wellness War Is Over; Wellness Lost." Insurance ThoughtLeadership. March 23, 2017. Accessed February 25, 2019. http://insurancethoughtleadership.com/wellness-war-wellness-lost/.

Mattke, Soeren, Liu, John P., Huang, Christina Y., Van Busum, Kristin R., Khodyakov, Dmitry, Shier, and Victoria. "Reviewing Workplace Wellness Programs." RAND Corporation. May 30, 2013. Accessed February 25, 2019. https://www.rand.org/pubs/research_reports/RR254.html.

Mervosh, Sarah. "Jerry Westrom Threw Away a Napkin Last Month. It Was Used to Charge Him a 1993 Murder." The New York Times. February 17, 2019. Accessed February 25, 2019.

Molteni, Megan. "23andMe's Pharma Deals Have Been the Plan All Along. Wired. August 06, 2018. Accessed February 27, 2019. https://www.wired.com/story/23andme-o glaxosmithkline-pharma-deal/.

Peryer, Marisa, Marisa Peryer, and Serena Cho. "Yale Geneticist Provided Data Used for Chinese Surveillance of Uighurs." Yale Daily News Yale geneticist provided data used for Chinese surveillance of Uighurs Comments, February 22, 2019. https://yaledailynews.com/blog/2019/02/22/yale-geneticist-provided-data-used-for-chinese-surveillance-of-uighurs/.

Regalado, Antonio. "More than 26 Million People Have Taken an at-Home Ancestry Test." MIT Technology Review. MIT Technology Review, February 18, 2019. https://www.technologyreview.com/s/612880/more-than-26-million-people-have-taken-an-at-home-ancestry-test/.

Tanner, Adam. "Harvard Professor Re-Identifies Anonymous Volunteers In DNA Study." Forbes. April 25, 2013. Accessed February 27, 2019. https://www.forbes.com/sites/adamtanner/2013/04/25/

"The Genetic Information Nondiscrimination Act of 2008." Information about the Americans with Disabilities Act Amendments Act (ADAAA). Accessed February 25 2019. https://www.eeoc.gov/laws/statutes/gina.cfm.

Vincent, James. "23andMe and Other DNA-testing Firms Promise Not to Share Data without Consent." The Verge. August 01, 2018. Accessed February 27, 2019. https://www.theverge.com/2018/8/1/17638680/genetic-data-privacy-consumer-rights-guidel ines-23andme-ancestry.

Wagner, Jennifer. "Update on Chadam v. Palo Alto Unified School District." The Privacy Report, April 18, 2017. https://theprivacyreport.com/2017/01/24/update-on-chadam-v-palo-alto-unified-school-district/.

Wee, Sui-lee. "China Uses DNA to Track Its People, With the Help of American Expertise." The New York Times. February 21, 2019. Accessed February 28, 2019. https://www.nytimes.com/2019/02/21/business/china-xinjiang-uighur-dna-thermo-fisher.

Wolff, Josephine. "How Is the EU's Data Privacy Regulation Doing So Far?" Slate Magazine. Slate, March 20, 2019. https://slate.com/technology/2019/03/gdpr-one-year-anniversary-breach- notification-fines.html.

Zhang, Sarah. "DNA Got a Kid Kicked Out of School-And It'll Happen Again." Wired. June 03, 2017. Accessed February 25, 2019. https://www.wired.com 2016/02/schools-kicked-boy-based-dna/.

Zhang, Sarah. "The Loopholes in the Law Prohibiting Genetic Discrimination." The Atlantic. March 13, 2017. Accessed February 25, 2019. https://www.theatlantic.com/health/archive/2017/03/genetic-discrimination-law-gina/519216/.

Acknowledgements

Thanks to the Science and Education Policy Association (SEPA) in New York City for constructive discussions.