23andMe/GEDmatch Case Study Readings

Links:
https://www.fastcompany.com/3040356/what-23andme-is-doing-with-all-that-dna
https://www.nejm.org/doi/full/10.1056/NEJMp1316367
https://www.nytimes.com/2020/08/01/technology/gedmatch-breach-privacy.html
https://www.nbcnews.com/news/us-news/police-were-cracking-cold-cases-dna-website-then-fine-print-n1070901

How 23andMe Is Monetizing Your DNA

Despite its trouble with the FDA, 23andMe is expanding overseas and focusing on its other business, which it says could help millions.

BY MICHAEL GROTHAUS

After the biotechnology startup 23andMe ran afoul of the FDA last year for not properly backing up health claims it was making based on customers' genetic results, 23andMe's primary business seemed doomed. Indeed, until it can iron things out with the federal agency, the company is limited in the U.S. to marketing only its DNA ancestry kits to consumers.

But the controversy hasn't stopped the company from doing what it set out to do when it launched in 2006: provide customers with a better understanding of their bodies and the DNA they'll pass to their children, and offer up all of that data to science—for a profit. In recent months, 23andMe has brought its popular Personal Genome Service to customers in Canada and the United Kingdom, giving them an opportunity that's currently off limits in the States: fill a tube with saliva, mail it back, and have it analyzed to determine some 100 risk factors for a number of diseases, inherited conditions, and adverse responses to certain drug therapies.

That service, along with its ancestry service, powers the B2B aspect of 23andMe's business: users can opt in to share their genetic information, and 23andMe shares or sells that data to a range of partners—from academic labs to pharmaceutical companies—in an arrangement it says can ultimately help humanity as a whole.

Last week, the company announced a reported $60 million deal with Genentech to generate whole genome sequencing data for members of 23andMe's Parkinson's disease community. This week, it revealed that it will give drug maker Pfizer access to 23andMe's genetic database. (The company is also waiting for FDA officials to approve just a single test as a "medical device": a DNA analysis for Bloom syndrome, an inherited disorder that often results in deadly cancer by the mid-twenties.)

While the company doesn't disclose the specifics of its data sharing agreements, they have always been central to 23andMe's strategy. "The long game here is not to make money selling kits, although the kits are essential to get the base level data," Patrick Chung, a 23andMe board member, told *Fast Company* last year. "Once you have the data, [the company] does actually become the Google of personalized health care."

Ironically, as 23andMe works to resolve its issues with the FDA for its health products, its data-analysis initiative is earning support from other U.S. government agencies. In July, the company received a $1.4 million grant from the National Institute of Health to help expand its genotype database. A study released last week by one of the company's researchers that illustrates the geographic spread of race and ethnicity throughout the U.S., based on the anonymized data of 160,000 customers, was funded in part by the NIH and the National Science Foundation. Among its findings: About 3.5% of self-identified European Americans have at least 1% or more African Ancestry, and one in every 20 African-Americans have Native American ancestry.

Just as apps like Mint and LinkedIn have transformed how we access, understand, and manage our financial and professional data, 23andMe, via its Personal Genome Service, wants to become a central storehouse for our genetic data. Founded by Anne Wojcicki, who is married to (though separated from) Google cofounder Sergey Brin, a carrier of Parkinson's disease, the company—named for the 23 pairs of chromosomes in a normal human cell—wants to offer users the tools they need to understand their genetic information, and how it might relate to their health. And they see the timing as right, as health care slowly but surely moves from the generic one-size-fits-all approach to more unique treatments based on your genes.

Users appear interested in sharing their genetic data with the world, provided it's anonymized. The company says that over 80% of its users have opted in to its data-sharing system, tantalized by the hope of contributing to important science that could eventually come back to help them and their families. As one 23andMe user who found out she was predisposed to breast cancer told

me, “Genetic data is the most personal data I own, but if my data can contribute to finding better treatment or even a cure, why should I think twice about sharing it?”

IDENTIFYING GENETIC TRAITS AT AN UNPRECEDENTED SCALE

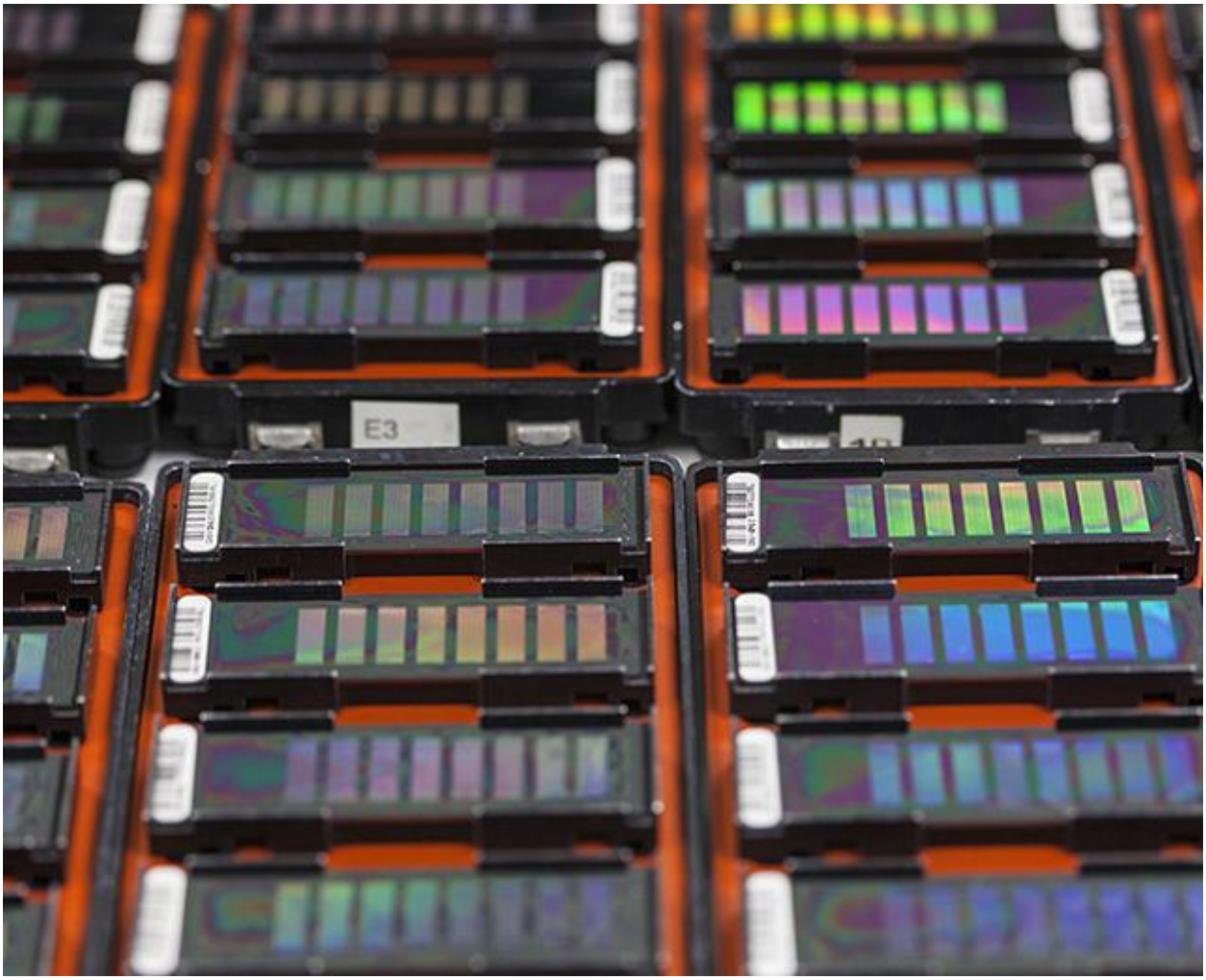
By effectively “crowdsourcing” people’s DNA data, big pharmaceutical companies can do research that even their own large budgets might not otherwise allow. “Traditional research can take more than a decade and millions of dollars to conduct studies with just under a few hundred participants,” says Angela Calman-Wonson, VP of Communications at 23andMe. “We can undertake real-time research initiatives drawn from the more than 600,000 23andMe customers who have proactively elected to share their de-identified genetic information for research, and answer survey questions. This approach eliminates recruitment times, minimizes cost, and reduces the amount of time it takes to conduct research.”

It’s an approach that has caught the attention of academics and, increasingly, large pharmaceutical and biotech companies just like Pfizer, which had worked with 23andMe before this week’s announcement, to study the genetics of inflammatory bowel disease (IBD), and Genentech, which had previously paid the company to help it find breast cancer patients who had taken its drug Avastin. The approach has caught the attention of funders, too–so far, the company has raised $126 million. 23andMe won’t reveal how much it charges partners, which include the likes of The Michael J. Fox Foundation–for access to anonymized user data, but the company says it has more than 30 active collaborations with academic research centers and industry from around the world.

Those partnerships, over the past four years, have led 23andMe to publish or contribute to 22 peer-reviewed papers about a range of areas, including allergies, asthma, hypothyroidism, myopia, and breast cancer. These kinds of contributions, the company says, are the direct results of 23andMe’s users’ willingness to share their anonymized genetic data.

“For example, researchers came to 23andMe to learn whether or not a certain gene was more prevalent in cancer patients,” says Calman-Wonson. “23andMe sent surveys to individuals in our database with that particular gene, asking several cancer-related questions. We received more than 10,000 responses in 12 hours, and were able to determine that the gene was not prevalent among cancer patients. This type of research typically takes months and thousands of dollars, and in this case it took about 48 hours.”

By scaling up, she says, 23andMe could spark greater price drops in genomics—and even bigger follow-on effects. “We are really just scratching the surface of what is possible,” she says. With 23andMe, “we believe we can achieve the scale necessary for breakthroughs much more quickly and efficiently than traditional research methods.”



23andMe chip

UNDERSTANDING THE VALUE OF YOUR GENETIC DATA

There are causes for concern, say critics of genomic testing, which tend to revolve around the concept of confidentiality. At *Scientific American*, science journalist Charles Seife summed up the worries by calling 23andMe’s Personal Genome Service “much more than a medical device; it is a one-way portal into a world where corporations have access to the innermost contents

of your cells and where insurers and pharmaceutical firms and marketers might know more about your body than you know yourself."

23andMe's privacy policy says that handing your genomic data over to your insurance company wouldn't be an acceptable use of the database. While the company acknowledges that it shares aggregate information about users' genomes to third parties, it insists that it doesn't sell personal genetic information without customers' explicit consent.

23andMe DNA kit

Still, on its website, the company explains that its data-mining analysis "does not constitute research on human subjects"—which is to say, it's not subject to the rules and regulations that typically protect experimental subjects' welfare and privacy. It also offers a warning: "Genetic Information that you share with others could be used against your interests. You should be careful about sharing your Genetic Information with others."

As to why customers feel comfortable sharing their data, Calman-Wonson points to a sense of altruism among its users, as well as to the company's privacy policy. It's one of the strictest in the technology industry, she says. Ditto for the company's security, from physical to hardware to software, which protects users' data from cyberattacks.

Calman-Wonson also says users can change their mind and opt out of the research program at any time or, if they choose, can close their account at any point, and their data will be deleted from the company's systems within 30 days. Most users, however, are proactive about sharing their data, she says.

"To date, more than 80% of our 800,000-plus customers have opted in to our research, and most answer survey questions," says Calman-Wonson. "A large number are very actively engaged in answering questions every month. On

average, each customer who opts in to participate in research informs more than 230 genetic studies."

One portion of 23andMe's user base constitutes the world's largest group of Parkinson's patients participating in research in the world.

"For our Parkinson's research community, we were able to enroll 3,400 Parkinson's patients, identify two new genetic associations for the disease, and publish the findings–all within 18 months. Today our Parkinson's community has more than 10,000 people participating," says Calman-Wonson.

Still, it's not entirely clear how the research side of 23andMe's business can continue to grow given the current hold on its Personal Genomics Service. Calman-Wonson admits that before DNA analysis and genome-based medical treatments can become commonplace, more research is needed—including the kind of research to which 23andMe is contributing.

"We are still operating in a comparatively new field," she says. "It has only been about 10 years since the human genome was first mapped, and we still have so much to learn about how DNA works. The human body is incredibly complex, and genetics are just one factor of many that informs human health."

ABOUT THE AUTHOR

Michael Grothaus is a novelist, journalist, and former screenwriter. His debut novel EPIPHANY JONES is out now from Orenda Books. You can read more about him at MichaelGrothaus.com More

23andMe and the FDA

- George J. Annas, J.D., M.P.H., and Sherman Elias, M.D.

In August 2013, the genetic-testing company 23andMe began running a compelling national television commercial, in which attractive young people said that for $99 you could learn "hundreds of things about your health," including that you "might have an increased risk of heart disease, arthritis, gallstones, [or] hemochromatosis". It was the centerpiece of the company's campaign to sign up 1 million consumers. On November 22, the Food and Drug Administration (FDA) sent 23andMe a warning letter ordering it to "immediately discontinue marketing the PGS [Saliva Collection Kit and Personal Genome Service] until such time as it receives FDA marketing authorization for the device." On December 5, the company announced that it was complying with the FDA's demands and discontinued running the commercial, noting on its website, "At this time, we have suspended our health-related genetic tests to comply immediately with the [FDA] directive to discontinue new consumer access during our regulatory review process."

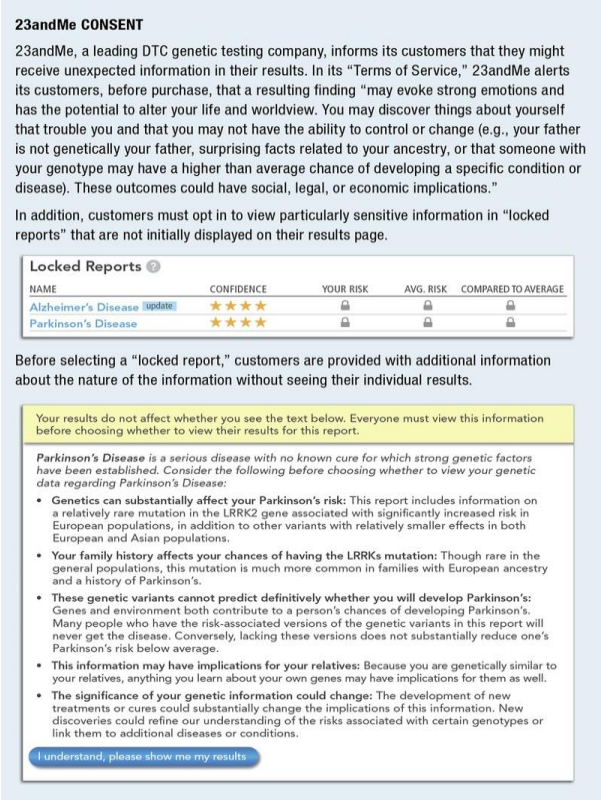
23andMe's services relied on single-nucleotide polymorphism (SNP) technology to identify genetic markers associated with 254 specific diseases and conditions (the list has grown over time), which, the company said, could inform people about their health and how to take steps to improve it. In the words of 23andMe's TV commercial, "Change what you can, manage what you can't." In its warning letter, the FDA said it was concerned that 23andMe failed to supply any indication that it had "analytically or clinically validated the PGS for its intended uses." The agency was also concerned about how consumers might use information concerning breast-cancer mutations and warfarin-related genotype results. The company and the FDA had been in continuous negotiations since July 2009, but in May 2013, the company stopped communicating with the agency. The company's failure to attempt to resolve the issues identified by the FDA, while it continued marketing the product, led to the warning letter. The FDA has not yet developed specific rules for direct-to-consumer (DTC) genetic testing, and whether government regulation or private litigation will determine the future contours of DTC genomic sequencing will probably depend on the extent to which consumers and physicians support government regulation.[1,2]

23andMe had previously framed DTC genetic testing as consumer empowerment — giving people direct access to their genetic information without requiring them to go through a physician or genetic counselor. To oversimplify, the debate has been framed as a struggle between medical (or government) paternalism and individuals' right to information about ourselves. In this sense, it is not so different from the older debate about whether patients should have direct access to their medical records and test results, which was ultimately resolved in favor of direct patient access. We think the day will come when this framing is appropriate, but not until the diagnostic and prognostic capability of genomic information has been clinically validated.[1,2]

It seems reasonable to predict, for example, that in the next decade or sooner, a majority of health plans will make it easy for their members to have their entire genomes sequenced and linked to their electronic health records and will provide software to help people interrogate their own genomes, with or without the help of their physicians or a genetic counselor supplied by the health plan. This service will, of course, require a massive data bank of genome reference materials, and the FDA and the National Institute of Standards and Technology are collaborating on the development of reference materials.[2] Before genomic tests have been validated, however, genomic information can be misleading — or just plain wrong — and could cause more harm than good in health care settings. In most cases, family history is likely to be at least as informative about an individual's health risks as SNP-based testing like that used by 23andMe. In this regard, the FDA's warning letter to 23andMe for its nonvalidated PGS, which resulted in 23andMe's ceasing to sell its product, is not currently depriving people of useful information; the agency is merely requiring that companies that want to sell their health-related medical devices to the public demonstrate to the FDA that they are safe and effective — in this case, that the tests do what the company claims they do. That is traditional consumer protection and what the public expects from the FDA.

Privacy is a closely related issue. How can the extremely private and personal information locked in our DNA be protected so that others cannot use it for their own purposes without our consent or make it available to people or organizations who could use it against us (e.g., by denying us life or disability insurance)? 23andMe has, for example, suggested that its longer-range goal is to collect a massive biobank of genetic information that can be used and sold for

medical research and could also lead to patentable discoveries. Such uses seem reasonable so long as the consent of the DNA donors is properly obtained and their privacy is protected. Both of these requirements are, however, much more difficult to uphold than 23andMe seems to realize.[3]



23andMe CONSENT

23andMe, a leading DTC genetic testing company, informs its customers that they might receive unexpected information in their results. In its "Terms of Service," 23andMe alerts its customers, before purchase, that a resulting finding "may evoke strong emotions and has the potential to alter your life and worldview. You may discover things about yourself that trouble you and that you may not have the ability to control or change (e.g., your father is not genetically your father, surprising facts related to your ancestry, or that someone with your genotype may have a higher than average chance of developing a specific condition or disease). These outcomes could have social, legal, or economic implications."

In addition, customers must opt in to view particularly sensitive information in "locked reports" that are not initially displayed on their results page.

Locked Reports

NAME	CONFIDENCE	YOUR RISK	AVG. RISK	COMPARED TO AVERAGE
Alzheimer's Disease update	★★★★	🔒	🔒	🔒
Parkinson's Disease	★★★★	🔒	🔒	🔒

Before selecting a "locked report," customers are provided with additional information about the nature of the information without seeing their individual results.

Your results do not affect whether you see the text below. Everyone must view this information before choosing whether to view their results for this report.

***Parkinson's Disease** is a serious disease with no known cure for which strong genetic factors have been established. Consider the following before choosing whether to view your genetic data regarding Parkinson's Disease:*

- **Genetics can substantially affect your Parkinson's risk:** This report includes information on a relatively rare mutation in the LRRK2 gene associated with significantly increased risk in European populations, in addition to other variants with relatively smaller effects in both European and Asian populations.
- **Your family history affects your chances of having the LRRKs mutation:** Though rare in the general populations, this mutation is much more common in families with European ancestry and a history of Parkinson's.
- **These genetic variants cannot predict definitively whether you will develop Parkinson's:** Genes and environment both contribute to a person's chances of developing Parkinson's. Many people who have the risk-associated versions of the genetic variants in this report will never get the disease. Conversely, lacking these versions does not substantially reduce one's Parkinson's risk below average.
- **This information may have implications for your relatives:** Because you are genetically similar to your relatives, anything you learn about your own genes may have implications for them as well.
- **The significance of your genetic information could change:** The development of new treatments or cures could substantially change the implications of this information. New discoveries could refine our understanding of the risks associated with certain genotypes or link them to additional diseases or conditions.

I understand, please show me my results

From the Presidential Commission for the Study of Bioethical Issues.[4]

Informed consent to genomic testing is the subject of a wide-ranging debate, touched off by testing policies published by the American College of Medical Genetics and Genomics (ACMG). Their recommended guideline requires that when a physician orders a clinical sequencing test, the laboratory also test for pathogenic (or probably pathogenic) mutations in 56 genes, related to 24 serious disorders. According to an ACMG clarifying statement, "patients cannot opt out of the laboratory's reporting of incidental [secondary] findings to the ordering clinician". Such a requirement does not amount to informed consent but represents a waiver of the right to decide what tests will be performed. People have both a right to know what will be done to diagnose their condition and a right not to know about their genetic predispositions if they don't want to

know.[4,5] 23andMe had adopted a more rights-respecting mode here — giving customers a second chance not to find out about the results of specific tests (such as tests for breast-cancer mutations, Parkinson's disease, and Alzheimer's disease) after the test is done .[4]

Whole-genome screening, whether ordered by physicians or consumers, will require more sophisticated informed-consent protocols, and we believe that individuals should also retain the right not to have specific genes sequenced at all.[5] James Watson set a reasonable standard for nondisclosure. He authorized the publication of his entire genome with one exception: he refuses to be informed of his $APOE$ status or have it published because he does not want to know whether he is at higher-than-average risk for Alzheimer's disease. That should be his right and the right of every patient or consumer.

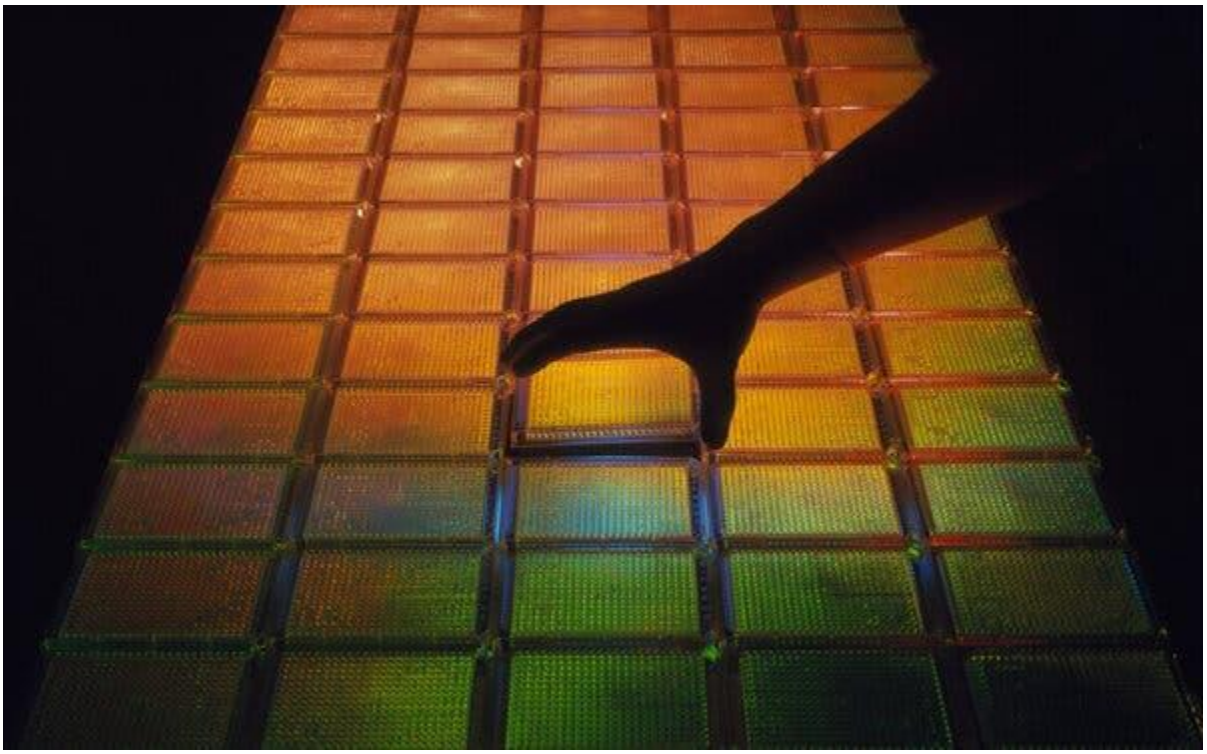
Because of the company's aggressive marketing and refusal to resolve outstanding data issues, the FDA was right to issue a warning to 23andMe. The resulting marketing shutdown provides the opportunity for serious dialogue that could be a basis for setting standards not just for 23andMe, but for the entire industry. 23andMe, for example, makes the consumer's raw genetic data derived from the DNA sample accessible to the consumer, something all biobanks should do. It could also be a catalyst for creating a regulatory framework for whole-genome–sequencing platforms, which are the future of genomics.[1] As the cost of such sequencing continues to fall, millions of people will probably have their genomes sequenced. That will turn out to be the easy part. The difficult part will be, as it is today, the clinical interpretation of an individual's genome and the making of useful recommendations to the patient–consumer. Put another way, the heart of this debate is not the cost of the sequencing (or SNP testing), but rather whether the information produced can be used in ways that improve our health. We think that the goal of the FDA and 23andMe (as well as all clinical geneticists, testing laboratories, and the entire genetics industry) should be to ensure that genomic information is both accurate and clinically useful. Clinicians will be central to helping consumer–patients use genomic information to make health decisions. Any regulatory regime must recognize this reality by doing more than simply adding the tagline on most consumer ads for prescription drugs: “Ask your physician.” That is insufficient guidance unless your physician has ready access to a clinical geneticist or genetic counselor.

Disclosure forms provided by the authors are available with the full text of this article at NEJM.org.

This article was published on February 12, 2014, at NEJM.org.

Why a Data Breach at a Genealogy Site Has Privacy Experts Worried

Nearly two-thirds of GEDmatch's users opt out of helping law enforcement. For a brief window this month, that didn't matter.



The GEDmatch breach shows what can go wrong when stored genetic information isn't adequately safeguarded.Credit...James King-Holmes/Science Source

By Heather Murphy
Aug. 1, 2020

The peculiar matches began early on a Sunday morning. Across the world, genealogists found that they had numerous new relatives on GEDmatch, a website known for its role in helping crack the Golden State Killer case.

New relatives are typically cause for celebration among genealogists. But upon close inspection, experienced users noticed that some of the new relatives seemed to be the DNA equivalent of a Twitter bot or a Match.com scammer; the DNA did things that actual people's DNA should not be able to do.

Others seemed to be suspected murderers and rapists, uploaded by genealogists working with law enforcement. Users knew that the police sometimes used the site to try to identify DNA found at crime scenes. But users found the new profiles strange because they also knew that profiles made for law enforcement purposes were supposed to be hidden to prevent tipping off or upsetting a suspect's relatives amid an investigation. What really drew attention, however, was the fact that all one million or so users who had opted not to help law enforcement had been forced to opt in.
GEDmatch, a longstanding family history site containing around 1.4 million people's genetic information, had experienced a data breach. The peculiar matches were not new uploads but rather the result of two back-to-back hacks, which overrode existing user settings, according to Brett Williams, the chief executive of Verogen, a forensic company that has owned GEDmatch since December.

Though the growth of genealogy sites has slowed slightly in recent years, their use by the police has increased. After the authorities in California used GEDmatch in 2018 to identify a suspect in the decades-long Golden State Killer case, police departments across the country began to dig through their cold case files in the hopes that this new technique could solve old crimes.

And GEDmatch was often their preferred site. Unlike the genealogy services Ancestry and 23andMe, which are marketed to people who are new to using DNA to learn about themselves, GEDmatch caters to more advanced researchers. The site appeals to the police because it allows DNA that has been processed elsewhere to be uploaded. Verogen has a long history of working with law enforcement, and the acquisition of GEDmatch further solidified this collaboration.

Scientists and genealogists say the GEDmatch breach — which exposed more than a million additional profiles to law enforcement officials — offers an important window into what can go wrong when those responsible for storing genetic information fail to take necessary precautions.

In an interview, Mr. Williams said that the first breach occurred early on July 19. After shutting down the site, his team "covered up the vulnerability," he said, and brought it back online, but only briefly. "On Monday we took the site down again because it was clear the hackers were trying again," he said.

This time the site remained down for nearly a week. "We're taking an abundance of caution because we don't want to end up in the same situation again," Mr. Williams said.

Mr. Williams said he had hired an outside security team and contacted the F.B.I. to see if the agency would investigate. The F.B.I. did not respond to a request for comment.

All was far from resolved when the site's settings were restored, said Debbie Kennett, a genealogist in England, who wrote about the breach on her blog. We're stuck with our DNA for life, she said. "Once it's out there it's not like an email address you can change," she said in an interview. Because of its interconnected nature, she added, when any one

person's genetic information is exposed, the exposed DNA can potentially affect their family members too.

In a paper published last year, Michael Edge, a professor of biological sciences at the University of Southern California, and fellow researchers warned several genealogy websites that they were vulnerable to data breaches.

"Of course, hacks happen to lots of companies, even entities that take security very seriously," he said. "At the same time, GEDmatch's, and eventually Verogen's, response to our paper didn't inspire much confidence that they were taking it seriously." Other genealogy websites, he added, seemed more open to the researchers' recommendations for improving security.

For many, the presence of fake users in GEDmatch was as alarming as the breach itself. Genealogists know that they cannot trust names or emails. They also know that a user can easily upload someone else's genetic profile. But the breach exposed that behind the scenes, hidden by privacy settings, were all kinds of profiles of people who were not even real.

The giveaway that the matches were not actual relatives was that their DNA was too good to be true, said Leah Larkin, a biologist who runs DNA Geek, a genealogical research company. People who managed profiles for many clients and relatives repeatedly found that these fake users somehow were displayed as close relatives across the unrelated profiles. Their visible ancestry information reinforced the matches were impossible and suggested the fake profiles had been designed to trick the site's search algorithm for some reason.

In Dr. Edge's paper, he warned that it was possible to create fake profiles to identify people with genetic variants associated with Alzheimer's and other diseases.

"If something is just a geeky genealogist messing around, there is no concern," Dr. Larkin said. But it becomes a problem, she said, if users are trying to find people who all share a particular genetic mutation or trait, as Dr. Edge cautioned. Such information could be abused by insurance companies, pharmaceutical companies or others, she said.

The breach also reinforced something that genealogists have been saying for years: Mixing genealogy and law enforcement is messy, even when you try to draw clear lines. Until two years ago, the primary DNA databases that law enforcement used for investigations were maintained by the F.B.I. and the police. That changed with the Golden State Killer case in 2018.

As police departments rushed to reinvestigate cold cases, GEDmatch, which at the time was run by two family history hobbyists as a sort of passion project, tried to serve two audiences: genealogists who simply wanted to trace their family tree and law enforcement officials who wanted to know if a murder or a rapist was hiding in one of its branches. Amid a backlash, GEDmatch changed its policy in May 2019 so that only users who explicitly opted to help law enforcement would show up in police searches. Still,

there is little regulation around how the authorities can use GEDmatch and other genealogy databases, so it's largely up to the companies and their users to police themselves.

And as the breach demonstrated, users' wishes could be quickly overridden.

For some users, the reason for keeping their profiles private is philosophical. Even if helping law enforcement could mean helping catch a killer, they do not want their genetic information used to incriminate their relatives. Others, like Carolynn ni Lochlainn, a genealogist from Huntington, N.Y., keep their profiles private because they worry the data will be improperly used to arrest innocent people.

"I work with a lot of Black clients and cousins, and I was most angered by the inexcusable risk at which they were placed," Ms. ni Lochlainn, said.

Colleen Fitzpatrick, the founder of Identifinders International, which applies forensic genealogy techniques toward identifying unclaimed remains and suspects in crimes, oversees a team that relies heavily on GEDmatch.

Her team was affected differently than the genealogists' clients. They had uploaded DNA from crime scenes and unidentified babies who had been abandoned by their mothers. Because they'd checked the law enforcement box, these profiles were not supposed to show up in their relative's searches. For a brief window in time, "the whole database, they could see us," she said.

She said it was unlikely that anyone working with law enforcement had exploited the breach to obtain a match against a relative's will, given the short amount of time involved. "It wasn't this magnificent reveal that we're going to cash in on," she said.

Nonetheless, the breach undeniably undermined trust for all, she said. "I think Verogen needs to up its game," she said.

A free public website called GEDmatch was at the center of a groundbreaking DNA tactic for solving cold cases.

Julien Posture / for NBC News

In April 2018, California authorities revealed that they'd used a novel investigative technique to arrest a man they called the Golden State Killer, a serial murderer who'd escaped capture for decades.

For the first time, police had submitted DNA from a crime scene into a consumer DNA database, where information about distant relatives helped them identify a suspect.

The announcement kindled a revolution in forensics that has since helped solve more than 50 rapes and homicides in 29 states.

But earlier this year, that online database changed its privacy policy to restrict law enforcement searches, and since then, these cold cases have become much harder to crack. The change is allowing some criminals who could be identified and caught to remain undetected and unpunished, authorities say.

"There are cases that won't get solved or will take longer to solve," Lori Napolitano, the chief of forensic services at the Florida Department of Law Enforcement, said.

Should police be able to use DNA databases to solve crimes?

The switch was imposed by GEDmatch, a free website where people share their DNA profiles in hopes of finding relatives. The company had faced criticism for allowing police to search profiles without users' permission, and decided that it would rather make sure members understood explicitly how investigators were using the site. So, it altered its terms of service to automatically exclude all members from law enforcement searches and left it to them to opt in.

Overnight, the number of profiles available to law enforcement dropped from more than 1 million to zero. While the pool has grown slowly since then, as more people click a police-shield icon on GEDmatch allowing authorities to see their profile, cases remain more difficult to solve, investigators say.

CeCe Moore, a leading specialist in using DNA evidence and family trees to identify criminal suspects — a method known as investigative genetic genealogy — depends on GEDmatch for her work. After entering a suspect's DNA profile into the site, she reviews the results and assesses the likelihood of law enforcement being able to determine the suspect's identity. She then scores each case from 1 to 5, 1 being a sure thing and 5 a long shot.

“I’m giving a lot more fives than I used to,” said Moore, who helped solve dozens of cases using GEDmatch before the site changed its terms of service, including the 1987 killing of a young Canadian couple, the 1988 murder of an 8-year-old Indiana girl and the 1992 rape and strangulation of a Pennsylvania schoolteacher.

This sharp drop in the usefulness of a promising technology has sparked an effort by law enforcement authorities and researchers like Moore to convince the public to take action. These groups hope to persuade more Americans to obtain their DNA profiles from direct-to-consumer genetic testing companies — most of which have large databases but don’t allow law enforcement searches — and share them publicly, including with law enforcement, on databases like GEDmatch. One direct-to-consumer company, FamilyTreeDNA, allows law enforcement to search its database, but charges for it and limits results.

Some people are reluctant, worried that their DNA profiles will be hacked or used against their wishes, whether in the pursuit of a criminal or in the sale of data to health care companies. There are also concerns that DNA sharing will lead to the end of anonymity.

But law enforcement authorities and genetic sleuths who work with them argue that there is greater public good in helping to keep killers and rapists off the streets.

“In the interest of public safety, don’t you want to make it easy for people to be caught?” said Colleen Fitzpatrick, a genetic genealogist who co-founded the DNA Doe Project, which identifies unknown bodies, and runs IdentiFinders, which helps find suspects in old crimes. “Police really want to do their job. They’re not after you. They just want to make you safe.”

A ‘very valuable tool’

To illustrate those points, investigators tell the story of Angie Dodge.

Dodge, 18, was raped and murdered in 1996 in her Idaho Falls, Idaho, apartment. A year later, a man confessed to the crime, and although he later recanted and his DNA didn’t match that of semen left on Dodge’s body, he was convicted of participating in the killing and sentenced to life in prison.

Dodge’s mother grew convinced that the prisoner, Christopher Tapp, was not her daughter’s killer. She pressed authorities to reopen the case. In 2017, Tapp was freed in a deal with prosecutors in which his conviction — of aiding and abetting the murder — remained.

So did the question of who left their DNA at the crime scene.

Almost a year later, California authorities said they'd used genetic genealogy to catch the Golden State Killer.

The announcement generated a surge of interest in the technique, as genetic genealogists teamed up with private companies to sell their services to law enforcement. (Public crime labs are not equipped to do the kind of DNA analysis required, and police generally aren't fluent in methods used to build family trees.) Parabon NanoLabs was the first, hiring Moore to run its genetic genealogy services. Idaho Falls police asked the company to try it.

Parabon submitted the Dodge suspect's DNA profile into GEDmatch in May 2018, but the DNA was so degraded that, even with more than 1 million profiles to compare against, the connections were sparse. Moore decided that genetic genealogy wouldn't work and declined to take up the case.

But Dodge's mother, Carol, begged Moore to keep trying. Moore relented and examined the connections more closely. With help from her team of genetic genealogists, she explored a series of leads that didn't pan out. They kept at it for months, eventually discovering a new branch of the suspect's family tree — and a potential suspect.

Police followed that man, collecting a cigarette butt he discarded and using it to obtain his DNA. It matched the crime scene profile, and in May 2019 Brian Leigh Dripps confessed, police said. A few weeks later, Tapp was exonerated. Dripps is awaiting trial.

Moore chronicled that search at a recent gathering of genetic researchers, investigators, prosecutors and lab technicians in Palm Springs, California. If she had been working on the Dodge case after GEDmatch limited access to its database, she told attendees of the International Symposium on Human Identification, "this case would not have been able to be solved by genetic genealogy."

She said she understood why GEDmatch's owners made the decision, but the result was allowing some violent criminals to remain free for longer than they would have been with the full power of genetic genealogy. She pleaded with her audience to take DNA tests and upload their profiles into GEDmatch.

"We don't want this very valuable tool to slip out of our hands," Moore said.

A slowdown in solving cases

Curtis Rogers didn't ask for this.

Rogers, 81, works in Florida as a court-appointed guardian for the elderly. He founded GEDmatch as a free public service in 2010 after being inspired by his own experience connecting with people who shared his last name. He partnered with a computer programmer who wrote software that made it easy for people to find relatives through certain shared pieces of genetic material. The site became popular among professional

and amateur genealogists, and as direct-to-consumer genetic testing services grew, GEDmatch enabled people to compare their DNA profiles in a single place.