DePaul Journal of Health Care Law

Volume 22 | Issue 2 Article 1

Ethics of Familial Genetic Genealogy: Solving Crimes at the Cost of Privacy

Craig M. Klugman
DePaul University

Hector F. Rodriguez

Ethics of Familial Genetic Genealogy: Solving Crimes at the Cost of Privacy

Craig M. Klugman, Ph.D.

Hector F. Rodriguez

DePaul University

INTRODUCTION

On April 24, 2018, Joseph James DeAngelo was arrested for allegedly committing twelve murders and at least 45 rapes in California between the years of 1976 and 1986.[1] Traditional detective work continued bringing investigators to dead ends, so they sought a new way to capture the Golden State Killer: Familial DNA Searching (FDS). Using a blood sample found at the crime scenes, investigators created a fake personal record and uploaded the deoxyribonucleic acid (DNA) sequence into GEDmatch. GEDmatch is one of the largest public DNA databases which acts as a tool to track down family members, create communities of people with specific diseases, and to provide DNA samples to researchers. The investigators found a partial match between their fake profile and another individual. That match turned out to be the third cousin of DeAngelo and through genetic genealogy, investigators were able to find the perpetrator who had become known as "the Golden State Killer".

What makes familial DNA searching different from other law enforcement genetic searches is that the goal in FDS is to not find the alleged perpetrator, but to find a member of their family. This process is featured in the popular television show *The Genetic Detective*,[2] where genetic genealogist CeCe Moore uses genetic material found at crime scenes to find a relative's

[1] Natalie Ram et al., *Genealogy databases and the future of criminal investigation*, 360 SCIENCE 1078, 1078 (2018).
[2] *The Genetic Detective* (ABC News 2020).

match in public and commercial DNA databases. Once a family is identified, Moore reconstructs the family tree and gumshoe detective work leads to the criminal. Law enforcement follows the same procedures.

In this paper, the authors first describe the science of familial DNA searching/forensic genetic genealogy and how it has been used to identify perpetrators of crime. Then, they explore the different types of genetic databases (forensic, commercial, and public) and the questions of social justice that arise in each. In the second part of this essay, the authors examine the bioethical implications of genetic identification for law enforcement including informed consent, privacy and confidentiality, accuracy and reliability of the technology, unintended genetic findings, regulation and oversight, and what is the higher social value: solving crimes or protecting privacy. This ethical analysis begins with the notion that just because something is legal does not mean that it is right.

FAMILIAL SEARCHING

There has been a growing interest in the use of familial DNA searching (FDS) in recent years due to its capabilities of helping law enforcement solve cold cases and ongoing criminal investigations. When the results of an FDS are combined with more traditional genealogical research, the process becomes known as "forensic genetic genealogy" (FGG).[3] Those who favor these techniques claim they are a good tool for law enforcement because of its potential to aid the identification and conviction of suspects, prevent crime, resolve cold cases, exonerate wrongfully convicted individuals, and improve public safety.[4] Although this new technology allows for law

[3] Although technically Familial DNA Searching (FDS) is the act of searching a genomic database and Forensic Genetic Genealogy (FGG) is using that search it leads to traditional genealogical research to find a suspect, this paper will use the terms interchangeably. Other agencies and authors have used a variety of terms including lineage testing.

[4] SARA DEBUS-SHERRIL & MICHAEL B. FIELD, UNDERSTANDING FAMILIAL DNA SEARCHING: POLICIES, PROCEDURES, AND POTENTIAL IMPACT, SUMMARY OVERVIEW 1 (ICF 2017).

enforcement to generate leads, there is still much to consider involving the privacy concerns for the American people.

When investigators search a genetic database, the DNA sample could either return with a full match or a partial match (unintentional or intentional). A full match means that the alleged perpetrator has been identified. Since prisoners are considered to have relinquished some privacy rights, there are few ethical restrictions against searching for a specific suspect in the forensic databases.

Inadvertent partial matching, or unintentional familial searching, occurs when law enforcement expects a full match but instead receives a partial match. The FBI considers a partial match as a spontaneous product of a regular database search in which the match identified may be a biological relative of the forensic profile. After performing the tests, the investigators look into the samples and determine if any have links to the DNA sample from the crime scene. Law enforcement uses FGG in order to generate leads by using the genetic link to lead them to the original source. In other words, a partial match may indicate the source of the sample has a biological tie to the person in the record (or at least they are probably in the same genetic family). This unintentional find may lead investigators to interview the family members of the sample in order to identify the alleged perpetrator. Partial matching raises additional concerns regarding the privacy rights of those who are not suspects but may be related to the suspect.

In a deliberate familial search, the intended consequence is to identify partial matches. Law enforcement is intentionally looking for near miss matches between the DNA found at the crime scene and records within the DNA database. Familial DNA searching is defined as a deliberate search of a DNA database using specialized software to detect and statistically rank a list of potential candidate records who may be close biological relatives (e.g., parent, child, sibling) to

the unknown individual contributing the evidence DNA profile.[5] FGG is not standardly performed at the federal level, because while those with criminal records are said to have waived certain privacy rights, their families have not. Thus, FGG search by federal authorities requires an application to a judge and approval of the DNA Advisory Board. Twelve states, though, do allow familial DNA matching (Arizona, California, Colorado, Florida, Minnesota, New York, Ohio, Texas, Utah, Virginia, Wisconsin, and Wyoming). The states that decide to apply FGG are required to establish a formal policy for its use. On the other hand, Maryland and Washington DC have passed laws prohibiting FGG.

DNA SEQUENCING

When a person takes a genetic sample—usually blood, spit, or cheek swab—and then has the DNA analyzed, the sequence alone does not offer much information other than a list of 3 billion base pairs. But, when the markers on the sequence are compared to a database of other samples, it becomes possible to find if a person has an inherited health condition, what geographic area their ancestors may have come from, and even who are their relatives. This comparison is done by conducting one or more DNA tests—autosomal, Y-DNA, and mtDNA.

In *autosomal testing,* the sequencing looks at all 46 chromosomes to find short sequences of DNA that marks where certain genes are usually found. These markers are not the genes themselves but are usually associated with particular genes. The markers are usually easier to find and see, and perhaps, more importantly are part of the company's intellectual property. Known as *single nucleotide polymorphisms* (SNPs), DNA components that vary between individuals and groups of individuals, a sample sent to multiple companies can yield different results because each

[5] *Id.* at 2.

company is looking for its own set of SNPs. Each company also has its own database of what are termed *haplogroups*, which are a group of alleles usually inherited together. By comparing such groups against a database from samples collected around the world, these companies try to assign a geographic ancestry, though they are not highly accurate.

Y-DNA looks at the male lineage of a person. Since the Y chromosome is passed from father to son, a sample's (or person's) genetically related males all have the same Y chromosome. Thus, the Y chromosome can easily allow for checking against a database that would show other male relations. In *mtDNA*, the investigation is looking at mitochondrial DNA. The mitochondria are an organelle in the cell that is inherited solely from the mother (it is part of the oocyte) and has its own DNA. Thus, comparing mitochondrial DNA in a database can permit knowing a person's matrilineal line.

GENOMIC DATABASES

DNA databases contain records of people's genetic information. These platforms exist to help in genealogy, medical research, and increasingly in forensics—to identify both DNA samples found at crime scenes as well as unidentified bodies. There are three types of DNA database platforms—forensic, commercial, and public—each with its own rules, audience, and uses.

Forensic

Since 1986, DNA analysis has been a crucial part in the criminal justice system within the United States because it can link a suspect to the DNA found at a crime scene.[6] In 1990, the FBI, along with 14 states and local laboratories, began developing software to allow comparisons of DNA samples, a program today known as the Combined DNA Index System (CODIS). Congress

[6] Hillary L. Kody, *Standing to Challenge Familial Searches of Commercial DNA Databases*, 61 WM. & MARY L. REV. 287, 290 (2019).

then passed the DNA Identification Act in 1994 that permitted the FBI to establish the National DNA Index System (NDIS) databases to maintain genetic samples from people convicted of crimes.[7] In 1998 CODIS went online and by 2004 all 50 states had actively joined. Then, in 2013, *Maryland v. King*[8] expanded the database further by allowing the collection of DNA from people arrested for violent crimes, not just those who were convicted. All 50 states, the federal government, Washington DC, and Puerto Rico participate in NDIS and all use CODIS to search the federal and local database. Technically, this program is a collection of "indexes" including convicted offenders, arrestee, unidentified human remains, biological relatives of human persons, and legal (i.e. all others), all of which are accessed in a CODIS search. The NDIS database consists of records uploaded into the system by the states. Federal guidance is for all convicted felons' and detained undocumented immigrants' DNA to be included. However, some states send more information, such as anyone arrested, while others have included suspects or even people who committed misdemeanors. Thus, decisions on what records are uploaded into NDIS varies by local jurisdiction.

As of June 2020, NDIS contained nearly 14.3 million offender records, 4 million arrestee profiles, and over 1 million forensic (usually unidentified human remains) records. CODIS/NDIS has produced nearly 520,000 matches in close to 509,000 investigations.[9]

Though the Golden State Killer case was the most publicized crime solved by FGG, it is certainly not the first. One of the most violent serial killers caught was a man known as the "Grim Sleeper." The Grim Sleeper raped then killed sex workers by either shooting or strangling them.

---

[7] *Frequently Asked Questions on CODIS and NDIS*, FBI (2020), https://www.fbi.gov/services/laboratory/biometric-analysis/codis/codis-and-ndis-fact-sheet.
[8] Maryland v. King, 133 S. Ct. 1958 (2013).
[9] *Frequently Asked Questions on CODIS and NDIS*, FBI (2020), https://www.fbi.gov/services/laboratory/biometric-analysis/codis/ndis-statistics.

Between the years of 1985 and 1988, he attacked and murdered a total of eight women. Fourteen years later, the Grim Sleeper struck again, adding three more victims between 2002 and 2007.[10] In an effort to put an end to his murders, police turned to familial DNA searching. At that time, California was the first state to explicitly authorize intentional familial searches, which are limited to the most serious cases and when the investigative trails have gone cold.[11] The DNA search done in 2008 did not initially match with any record, but two years later investigators got a match belonging to Christopher Franklin, a young adult convicted of a felony weapons charge one year prior. However, at the time of the first set of murders, Christopher had not yet been born. Therefore, detectives focused on the males in Christopher's family that would have been aged appropriately for when the killings first took place, and eventually the police identified his father, Lonnie Franklin Jr., as the Grim Sleeper.[12] Posing as a busboy, a police officer collected a discarded pizza crust and cutlery that Lonnie had touched. Investigators sequenced the original samples from the crime scenes and compared it to Lonnie's pizza-sourced DNA profile. The samples matched and Lonnie was arrested.[13]

Commercial Database

A commercial database is a repository of DNA samples owned by a private company. In the U.S. alone there are a half dozen of these entities: 23andMe, Ancestry, Family Tree DNA, Living DNA, My Heritage (Geni.com), Geno 2.0 (National Geographic and Helix), and Pathway Genomics. They use a direct-to-consumer model where for a modest fee, the company collects a sample, sequences parts of the genome, and compares the results to a database. Most companies return a

[10] Amanda Pattock, *It's All Relative: Familial DNA Testing and the Fourth Amendment*, 12 MINN. J.L. SCI. & TECH. 851, 852 (2011).
[11] ERIN E. MURPHY, INSIDE THE CELL: THE DARK SIDE OF FORENSIC DNA 195 (Bold Type Books 2015).
[12] Pattock, *supra* note 10, at 852.
[13] MURPHY, *supra* note 11, at 195.

report to the consumer and most also then sell their databases to researchers, life science companies, and governments. Together, these companies have an estimated 26 million records in an industry expected to be worth $22 billion a year by 2024.

The first commercial DNA company was founded in 1997 under the name GeneTree.[14] The first direct-to-consumer heritage DNA endeavor was Family Tree DNA, founded in 2000. However, the popularity of these at-home tests did not explode until 23andMe was founded in 2007 (using saliva instead of blood for DNA samples). Today, direct-to-consumer DNA companies ship a kit to consumers (or sell one at major retail chains). Costs are about $99, though frequent sales bring the price to $59 and occasionally $49. A person spits into a tube and ships the sample back to the company, which sequences their set of SNPs and offers an analysis by comparing their genes to the company's database.

Public Databases

Public databases are non-profit, open source websites that allow people to upload their genetic information and make it freely available. Some of these websites allow users to build communities of people with similar genetics such as having ancestors from the same region or having the same genetic disease. They allow people with familial genetic conditions to work with researchers in hopes of finding cures and treatments. From 2008 to 2015, One Thousand Genomes aimed to create a publicly available catalog of DNA sequences for free use by researchers. Eventually, the project collected genomes from 2,506 individuals from 26 separate population groups. The project has continued as the International Genome Resource Sample (IGSR) under the European Bioinformatics Institute and the Wellcome Trust.

[14] GeneTree was later sold to Sorenson Molecular Genealogy Foundation in 2001, which was later bought by Ancestry in 2012.

The most well-known public database is GEDmatch, founded in 2010 by Curtis Rogers (retired) and John Olson (transportation engineer) for genetic researchers and genealogists.[15] In April 2018, this was the platform used by law enforcement to find a familial match that led them to identify DeAngelo (i.e. the Golden State Killer). At the time, the terms of service made no mention of law enforcement use of the service and thus was permissible because it wasn't forbidden.

What distinguishes public genomic databases is that their records are available at no cost to researchers and contributors. They also do not sequence the DNA; users must have that done elsewhere (private labs or commercial DNA companies like 23andMe and Ancestry).

SOCIAL JUSTICE PROBLEMS

On June 29, 2020, Joseph James DeAngelo (i.e. the Golden State Killer) pleaded guilty to thirteen counts of first-degree murder, thirteen counts of kidnapping and special circumstances which include murders committed during burglaries and rapes.[16] These crimes remained unsolved for 34 years because DeAngelo was a police officer and knew the procedures for investigating crimes. He is also white, and the law enforcement genetic databases are highly biased toward people of color, meaning the chances of a direct match of a DNA sample from the crime scene and DeAngelo was exceedingly small.

NDIS/CODIS has a race problem. Since its records are of people convicted, charged, and arrested of crimes, its racial profile is similar to that of people who are convicted, charged, and arrested, a population that is overwhelmingly people of color. Although Blacks only make up 13% of the U.S. population, they are 40% of the NDIS records. Latinx populations account for 22% of

[15] In 2019, GEDmatch was sold to forensic science company Verogen Inc., which is developing a law enforcement platform for the website.
[16] Eliott C. McLaughlin & Stella Chan, *Hearing details ghastly crimes of golden state killer as he pleads guilty to killings*, CNN (June 9, 2020), https://www.cnn.com/2020/06/29/us/golden-state-killer-plea-expected/index.html.

prisoners but are only 16.7% of the total U.S. population. With the recent inclusion of undocumented immigrant detainees' DNA into NDIS, the representative percentage of people of color samples will grow.[17] In the 2010 census, 72% of respondents identified their race as "white", though they likely make up only a third of the NDIS records.

Commercial and public databases have a bias that swings in the other direction—their records are predominantly from people who identify as white or Caucasian, making up about 75% of the profiles. Estimates suggest that based on the existing records, it is possible to identify 60% of people of European/Caucasian heritage in the United States from a DNA sample and that will quickly rise to 90%.[18] This divide in representation in the databases comes from two factors—who gets arrested and who has financial means. Although the cost of a commercial DNA test is not exorbitant ($99 at retail), that can be a substantial amount of money for many. According to the U.S. Census Bureau, the median income for those identifying as white ($71,664) is almost $28,000 per year higher than those identifying as Black ($43,862) and $16,000 more than those identifying as Hispanic ($55,658).[19] With more disposable income, people who identify as white are more likely to subscribe to a commercial DNA service than those who are Black or Latinx. In one sense, law enforcement searches of commercial and public DNA databases correct for the racial imbalance in CODIS/NDIS. By searching these other platforms, crimes committed by people identifying as white are more likely to be solved. On the other hand, FGG raises a host of other ethical and legal concerns.

---

[17] Nicole Wetsman, *DNA from detained immigrants will change the nature of the FBI's genetic database*, VERGE (Jan. 14, 2020), https://www.theverge.com/2020/1/14/21063627/dna-detained-immigrants-fbi-codis-bias-crime-database.

[18] Jocelyn Kaiser, *We will find you: DNA search used to nab Golden State Killer can home in on about 60% of white Americans*, SCIENCE (Oct. 11, 2018), https://www.sciencemag.org/news/2018/10/we-will-find-you-dna-search-used-nab-golden-state-killer-can-home-about-60-white.

[19] Gloria G. Guzman, *Household Income: 2019*, at 2 (Sept. 2020), https://www.census.gov/content/dam/Census/library/publications/2020/acs/acsbr20-03.pdf.

INFORMED CONSENT & TERMS OF SERVICE

Genomic testing first appeared in the biomedical laboratory and the doctor's office. The information revealed by DNA sequencing indicates what diseases one might be susceptible to, and even what cancer treatment might be most effective. When direct-to-consumer DNA testing came into our homes, it also promised to give information about our inherited health conditions. A saliva sample sent through the mail would bring a report that discussed your genetic risk of Alzheimer's, heart disease, colon cancer, and alleles related to drug and alcohol metabolism in addition to your ancestry information.[20] In 2010, the FDA informed the commercial companies that their diagnosis claims were unproven and needed to undergo regulatory review. 23andMe, for example, discontinued providing health information in 2013. Then in 2017, the company was approved to test for 10 health conditions (including Parkinson's, Alzheimer's, and celiac diseases).[21]

As patients (and subjects), people expect when going to the doctor, having a medical test, joining a medical research study, or receiving health information to complete an informed consent process. Informed consent requires that a patient be told the procedures they will undergo; the risks, benefits, and alternatives. This information must be delivered in clear, understandable lay language and often uses "you" language (e.g. "Your temperature will be taken"). They have a chance to ask questions. A boilerplate statement in most research protocols says that refusing to participate will in no way affect their medical care. Patients and subjects can revoke their consent at any time. The process empowers the individual to exercise their autonomy to decide if they want to receive treatment or be part of the study. Their agreement is noted by signing a brief document laying out

[20] *Direct-to-consumer genetic testing kits*, HARV. HEALTH PUB. (Sept. 2010), https://www.health.harvard.edu/newsletter_article/direct-to-consumer-genetic-testing-kits.
[21] Jessica Boddy, *FDA Approves Marketing Of Consumer Genetic Tests For Some Conditions*, NPR (April 7, 2017), https://www.npr.org/sections/health-shots/2017/04/07/522897473/fda-approves-marketing-of-consumer-genetic-tests-for-some-conditions.

the above components.[22] When information changes (such as new side effects coming to light), the physician or researcher is required to re-consent the patient or subject. The document notifies the signer of their rights, which includes expectations of confidentiality in the fiduciary relationship in the case of medical care.

In contrast, terms of service are long, complicated documents written by attorneys with the express purpose of protecting a company and its intellectual property. The language tends to obfuscate and is written in the third person. Rather than being short like a consent document, terms of service are long and meandering: 23andMe's terms of service is 22 pages long and its privacy policy is another 22 pages. AncestryDNA comes in with 32 pages for both documents. If a consumer does not agree to the terms, then they cannot use or access the services; if they later revoke permission the company still gets to use their information (as per the terms, they get to use everything they have up until the time of the revocation). Agreement is designated by clicking a box. When the terms change (as they often do), a person can agree (indicated by continuing to use the service) or cancel the service (and/or delete their record) but that applies to only future uses (for example, if the company sold the record to a researcher, they do not reclaim that record—it is already out there for use). The company has no fiduciary responsibility to the consumer. In fact, the consumer's DNA is the company's product (not the sequencing or the test results).[23] The document notifies the user of the company's rights.

[22] Although written, signed consent is the gold standard, with IRB approval consent can be altered or waived such as allowing verbal consent for a phone survey or no consent if the signature would be the only thing that links a person to a survey that is of a sensitive topic.

[23] This is in evidence considering that the cost for the service ranges from $59 to $99 but the cost of a human genome sequencing is closer to $1,000. Given that the companies are not in business to lose money and subsidize sequencing, they must sell the genomic data to others in order to recoup costs and make a profit. Kris A. Wetterstrand, *The Cost of Sequencing a Human Genome*, NAT'L INST. HEALTH, https://www.genome.gov/about-genomics/fact-sheets/Sequencing-Human-Genome-cost (last updated Dec. 7, 2020).

Whereas an informed consent document is an acknowledgement of shared information, a terms of service is a legal contract. One focus of bioethics is on preserving an individual's autonomy to make their own informed decisions. When dealing with what people perceive as biological and therefore, medical information, there may be a confusion between the two processes and types of documents. While most people read a consent form (or at least listen to a description of it), 91% of people do not read the terms of service before agreeing to them.[24] That is hardly an informed choice. Thus, a service that purports (and is FDA approved in some cases) to provide medical information does not offer the same protection for DNA analysis that a person would have if their DNA sequencing was done through a doctor's office. The DNA companies are not considered medical providers even though they are providing de facto medical information. The same product (DNA sequencing) is treated differently and that is likely confusing to most people.

This potential misconception about the documents is an ethical concern. For example, after the DeAngelo case, GEDmatch changed its terms of service to explicitly allow law enforcement searches of the database. After public outcry, they changed the terms again to forbid law enforcement searches. The current policy is that users can opt-in to allowing law enforcement searches of their data. The default is to keep the records private. AncestryDNA does not permit law enforcement searches for criminal or identification purposes. On the other hand, 23andMe allows law enforcement searches under a broad set of circumstances.[25] Beyond the misconception

[24] Caroline Cakebread, *You're not alone, no one reads terms of service agreements*, BUS. INSIDER INDIA (Nov. 15, 2017), https://www.businessinsider.in/Youre-not-alone-no-one-reads-terms-of-service-agreements/articleshow/61659553.cms.

[25] *See, e.g.*, *Terms of Service*, 23ANDME, https://www.23andme.com/about/tos/ (last updated Sept. 30, 2019) ("Further, you acknowledge and agree that 23andMe is free to preserve and disclose any and all Personal Information to law enforcement agencies or others if required to do so by law or in the good faith belief that such preservation or disclosure is reasonably necessary to: (a) comply with legal process (such as a judicial proceeding, court order, or government inquiry) or obligations that 23andMe may owe pursuant to ethical and other professional rules, laws, and regulations; (b) enforce the 23andMe TOS; (c) respond to claims that any content violates the rights of third parties; or (d) protect the rights, property, or personal safety of 23andMe, its employees, its users, its clients,

issue, these varied approaches to the privacy concern may be confusing to the user of these services. The problem may require a voluntary standard approach in the industry or a regulatory solution, either rewriting terms of service into lay language or adopting universal consumer-protection provisions.

PRIVACY & CONFIDENTIALITY

The misconception could extend into another sacrosanct area of medical practice and research: privacy and confidentiality. In diagnosis and treatment, doctors need patients to share their secrets. Thus, health care workers have a fiduciary relationship that requires them to preserve confidentiality—they cannot reveal shared information except in a few circumstances.[26] However, neither law enforcement nor commercial DNA companies have a fiduciary relationship to a person. There are few ethical or legal restrictions on these entities making use of the genomic information including sharing (or selling) it.

For law enforcement databases, there are some privacy protections in place. For instance, NDIS only keeps necessary basic data for a match to occur, this includes the numeric digits in the profile, and a case file, lab, and analyst reference number. State and local databases, though, may include the full case information such as the name and identifying information of the sample.[27] As originally envisioned, NDIS/CODIS would contain the records of those who had been convicted of crimes. Thus, searching the database to identify an alleged felon at a scene was considered acceptable—their convicted status gives reasonable cause to compare their DNA record to a genomic sample. Over time, many states have uploaded DNA of people who were only arrested

---

and the public. In such an event we will notify you through the contact information you have provided to us in advance, unless doing so would violate the law or a court order.").

[26] Such circumstances include consulting with another health care worker assigned to the case, clear danger to self, or a third party. *See* Tarasoff v. Regents of the Univ. of Cal., 17 Cal. 3d 425, 131 Cal. Rptr. 14, 551 P.2d 334 (1976). *See also* Robert S. Orlick, *Confidentiality*, *in* PHILOSOPHY: MEDICAL ETHICS 232-33 (Craig M. Klugman ed., 2016) (discussing *Tarasoff* and how it has influenced privacy and confidentiality policy).

[27] MURPHY, *supra* note 11, at 14–16.

(even if later let go) or sometimes guilty of a misdemeanor. This expansion is a risk for the privacy and rights of the person whose DNA is on file and their families. Having their sample uploaded onto the database undermines the presumption of innocence by treating people who have merely been arrested as somehow less innocent than others who have also not been convicted of any offence.[28]

In FGG, the goal is not to find the criminal, but to find the family of the potential perpetrator. The investigators examine the DNA database to identify people who themselves are not suspected of the crime. The U.S. Supreme Court in *Maryland v. King* held that the offender surrenders a portion of their privacy when they voluntarily commit the criminal act, but the family does not give up their DNA privacy, so why should they become involved in the investigation?[29] Haimes notes that familial searching adds the possibility of "indirect lifelong surveillance of citizens" who will be "included by association" on the database even though they have never been suspected, let alone convicted, of a crime.[30]

Similarly, when looking at commercial and public databases, people gave over their DNA for purposes of education, research, and building community—not for law enforcement purposes (at least not before the terms of service were updated to reflect a site's policy). Whether the intention is to find an exact match or to find a relative match (FGG), what is the ethical justification for searching the records of people never convicted, arrested, or even suspected of a crime (even if the intent is not to find them, but to find one of their relatives)?

---

[28] *How Does Law Enforcement Use of DNA Affect Me?*, FORENSIC GENETICS POL'Y INITIATIVE, http://dnapolicyinitiative.org/how-does-law-enforcement-use-of-dna-affect-me/ (last visited July 14, 2020).
[29] Maryland v. King, 133 S. Ct. 1958 (2013); MICHAEL B. FIELD ET AL., STUDY OF FAMILIAL DNA SEARCHING POLICIES AND PRACTICES: CASE STUDY BRIEF SERIES 1, 35 (June 2017), https://www.ncjrs.gov/pdffiles1/nij/grants/251081.pdf
[30] Erica Haimes, *Social and Ethical Issues in the Use of Familial Searching in Forensic Investigations: Insights from Family and Kinship Studies*, 34 J.L. MED. & ETHICS 263, 264 (2006).

In short, the reason is convenience, which is not ethically defensible. Under a Kantian framework, an action is only moral if it is universally right. To say that it is okay to search a person's DNA record to find a relative is the same as saying it is okay to break into their house to look through their Christmas cards to find the address of a potential suspect. The latter is clearly prohibited under the Fourth Amendment without a judicial warrant. Searching commercial and public DNA records, does not require a warrant—it is available either by making a request of the company (if their terms of use permit it) or (in the past) creating a fake profile. Thus, the concept of extended family being genetically surveilled because of FGG may lead to abuse of the forensic tool, so the privacy concerns related to FGG focus on the Fourth Amendment's protections against unreasonable searches and seizures.[31]

Hacking

In July 2020, hackers were able to release the private records of one million users of GEDMatch and MyHeritage. At GEDMatch, which has an opt-in system for record-holders to be part of law enforcement searches, the hackers set the affected records to be open to FGG searches. Two days later, a phishing email brought users to a fake MyHeritage login page where username and passwords could be collected. Since MyHeritage does not permit law enforcement to search, there was no risk of that happening, but the hackers did have access to users' DNA sequences and analysis reports.[32] These two cases point to the potential for hacking to compromise genetic privacy in commercial and public DNA databases. In the first case, if a search of one of the hacked records led to a match, could that information be used to follow up on the lead? This case would neither have consent of the record holder nor fall under the terms of use. In the latter case, the

[31] DEBUS-SHERRIL & FIELD, *supra* note 4, at 6.
[32] Peter Aldhous, *A Security Breach Exposed More Than One Million DNA Profiles On A Major Genealogy Database*, BUZZFEED, https://www.buzzfeednews.com/article/peteraldhous/hackers-gedmatch-dna-privacy (last updated July 22, 2020).

leaked information could cause trouble for the user, by creating difficulties in their work, relationships, and reputation.

Stigma

Socially, being part of a dragnet of a suspect DNA search could expose a person or a family to stigma and bias. At the 2017 meeting of the International Society for Human Identification, one of us (CMK) was on a panel where law enforcement and prosecutors also spoke. The common theme was that they did not want to see any rules, regulations, or warrant requirements for this technique because it would slow them down. The other participants explained that FGG allows them to clear suspects—if a family is identified in a search, then they can go door-to-door and request a DNA sample to cross that person off the suspect list. But if the person exercises their right to refuse, then law enforcement would follow them until such time as they disposed of an item that would have their DNA (e.g. half-eaten sandwich, tissue, coffee cup) which could then be collected and analyzed. Even if the person did not know they were a suspect, the act of being followed or even having law enforcement approach them for a sample could create stigma on a person in their neighborhood or for their family. A neighbor might wonder why a person was going through their garbage. Such stigma by association could affect one's personal and professional well-being, relationships, and employment. Their family may excommunicate them if sending in their sample to a commercial DNA service led to another relative being arrested for a crime. After all, when you send in your sample to a commercial service, you are sharing not just your DNA but that of your 300 closest relatives.

Even if a person is eventually cleared (ironically, by way of a DNA test), there is a personal cost to being under suspicion from damage to relationships, loss of employment, loss of savings from hiring an attorney, and decreased health from the physical stress. For some, the idea that

someone could think of them as capable of committing a violent crime could be devastating to their own self-image and identity.[33]

Military Readiness

In December 2019, the Pentagon warned military personnel against using a popular Christmas gift, a commercial DNA test. Their concern was that finding out genetic health information that would be stored in a commercial database, "may affect readiness, could affect a service member's career, and the information from DTC genetic testing may disclose this information."[34] For example, no current law protects military members from professional decisions being made on their genetic information. If the service finds that a person has susceptibility to a genetic disease, that could limit their future promotions or field placements. Through computer hacking or even espionage, the information could also get into enemy hands, making a service member susceptible to manipulation by a foreign power. If these are concerns for military families, then they are concerns for all families.

ACCURACY & RELIABILITY

The use of genetic data for investigative purposes has the potential to do a lot of good, yet there are still many factors that question its reliability and accuracy. The states that utilize FGG already have some policies and safeguards in place to better protect the rights of its citizens. Although these policies are not necessarily rigid, some states include the requirement of staff training to ensure that they understand how familial searching works and its limitations. With the rising number of DNA profiles in the NDIS, there is a risk for a larger number of errors in

[33] *A father took an at-home DNA test. His son was then falsely accused of murder*, NPR (Nov. 17, 2019), https://www.pbs.org/newshour/show/a-father-took-an-at-home-dna-test-his-son-was-falsely-accused-of-murder.

[34] Heather Murphy & Mihir Zaveri, *Pentagon Warns Military Personnel Against At-Home DNA Tests*, N.Y. TIMES (Dec. 24, 2019), https://www.nytimes.com/2019/12/24/us/military-dna-tests.html.

uploading, safekeeping, and collecting DNA samples. Such errors can include not thoroughly cleaning equipment between tests (dirty tools), impure samples (i.e. having cells from more than one individual), and contamination from the people who prepare and analyze the samples. These errors mean that a person (or lineage) could become a suspect even if they had nothing to do with the crime. In 2014, the FBI admitted that a software upgrade revealed 170 errors in the database, and 166 of the mistakes consisted of entries that were off at a single nucleotide which an official speculated was largely the result of "interpretation errors by DNA analysts or typographical errors introduced when a lab worker" entered the profile.[35] There must be more oversight and in-depth training to secure a properly maintained database full of genetic property.

While the NDIS is examined by the FBI for compliance and quality assurance, state and local databases often lack adequate safeguards. These labs should undergo a series of regular proficiency tests through an external review to ensure that the staff is following standards. With nearly 14.3 million profiles, not only will the number of errors rise but so will the probability of finding a random partial match.[36] In order to test the accuracy and reliability of larger databases, researchers carried out a simulation study where a known parent and child (with biological ties) would be used to test the system. A DNA sample (of say the parent) generates a list of likely partial genetic matches (e.g. to the child). The database produced a likely match only 29% percent of the time.[37]

---

[35] Joseph Goldstein, *F.B.I. Audit of Database That Indexes DNA Finds Errors in Profiles*, N.Y. TIMES (Jan. 24, 2014), https://www.nytimes.com/2014/01/25/nyregion/fbi-audit-of-database-that-indexes-dna-finds-errors-in-profiles.html?smid=url-share.

[36] Statistics show that there is an increasing potential for random, nonmeaningful (partial) matches as the reference database grows. This is known as the *specificity* (or lack thereof) problem. NAT'L RES. COUNCIL, THE EVALUATION OF FORENSIC DNA EVIDENCE 89–123 (Nat'l Academies Press 1996).

[37] *Recommendations from the SWGDAM Ad Hoc Working Group on Familial Searching*, SWGDAM (2013), http://media.wix.com/ugd/4344b0_46b5263cab994f16aeedb01419f964f6.pdf.

The forensic use of DNA databases requires more safeguards, regulations, and rigid policies with consequences if they are not properly maintained in order to ensure accuracy and reliability. If there is a large number of partial non-specific matches, it could lead investigators to a dead end after intruding on the lives of many potential subjects. With the growing threat of contamination, errors made in the labs, and privacy concerns, this new technology could be costly to both the families identified as a suspect and law enforcement operations.

Unintended Findings

One of the realities in the era of genetics is that identification often reveals secrets. Consider the FBI going through a list of the male relatives on a reconstructed genealogy based on a match to a blood sample from a crime scene. Imagine this scenario: Law enforcement knocks on a door and explains they are looking at all relatives of a family. The person who answers claims no knowledge of the family under investigation. That person may have been adopted and did not know it. He might have been produced through assisted reproduction and was unaware that his genetic father was a sperm donor. In other scenarios, a person might find out that they were created from an egg donor or through a 3-parent embryo technique. Maybe the whole family learns that a person's biological father was the result of a brief affair and no one (not even the mother) suspected.

The knowledge revealed by FGG can shake families down to their roots, leading to arguments and perhaps even divorces. Children may have to reevaluate who they thought they were and how they fit into a family structure. "One of the crucial aspects of familial searching (and the other forensic uses of DNA analysis) is the way in which it entangles matters of identity with matters of identification, and the potential impacts that can have on a wide range of people, as

individuals and as members of family units and family networks."[38] This violation of privacy has far reaching consequences.

Another scenario is if the identification creates new knowledge about a person's health. If the identification comes from a commercial or public database, the matching profile may have information regarding familial disease about which the suspect family line lacked awareness. Is there a duty to share such findings with the suspects, or at least an obligation to inform the family that they might want to speak to their doctor? The federal health privacy regulations (such as HIPAA) and the lack of a fiduciary relationship would suggest that such a duty does not exist. But, ethically, having knowledge that could affect someone's health and even life decisions should be shared based on autonomy and beneficence (providing a benefit).

From an ethics standpoint, should there be a set of rules and guidelines for when talking to potential suspects identified through FGG? The potential for harm to a person and a family is high when sharing genetic family information. Beyond the obvious relationship turmoil, there could also be legal implications for parental rights (perhaps a child was kidnapped as a baby and thus not legally adopted).[39] Another potential point of harm could come from learning revealing information and then later finding that the genetic identification analysis was wrong. "Sorry, we thought your family might have contained the killer and we are sorry that family secrets were revealed, but it turns out we had the wrong family." As described earlier, there are issues of reliability and accuracy in this technology, meaning it is possible that a person is incorrectly identified as part of a suspected family line, their lives are immeasurably harmed, and it all turns out to be a mistake.

REGULATION AND OVERSIGHT

[38] Haimes, *supra* note 30, at 273.
[39] *Id.* at 264.

While law enforcement has preferred to keep the wild west approach to FGG where they could simply search commercial and public databases without informing anyone, that approach endangers the privacy of unsuspecting and innocent people whose only misstep was joining a DNA service. In response to more crimes being solved using FGG, and public concern for the potential abuse of this technology, the Department of Justice released a *Policy on Forensic Genetic Genealogical DNA Analysis and Searching* to guide FGG searches for DOJ agencies, or local and state agencies with federal funding. The guidelines say FGG is appropriate only if several conditions are met: first, the case must be of an unsolved violent crime.[40] Second, law enforcement cannot upload a fake DNA profile to a commercial or public genealogy website without identifying themselves. Third, they can only work with websites that allow law enforcement searches in the terms of service. This would suggest that records from the above mentioned GEDMatch hack could not be used. Fourth, law enforcement is not supposed to take a secret DNA sample (the aforementioned discarded coffee cup), but rather requires consent or a warrant.[41] Of course, any non-federal agency that does not have federal funding is not required to comply.

These guidelines are a beginning to what is needed to protect the rights of ordinary citizens. When law enforcement wants to do an FGG search on NDIS/CODIS, they must also apply to the DNA Advisory Board which controls approval to do the search. A similar citizens board could be established to oversee requests for such searches in commercial and public databases as well. This is congruous to what happened in the United Kingdom where the Biometrics and Forensics Ethics Group[42] has oversight into the ethical issues surrounding familial searching and commenting on

[40] *Interim Policy Forensic Genetic Genealogical DNA Analysis and Searching*, U.S. DEP'T JUST. (Sept. 2, 2019), https://www.justice.gov/olp/page/file/1204386/download (defining a violent crime as "any homicide or sex crime . . . [and] it also includes other serious crimes.").
[41] *Id*. at 6.
[42] First established as the National DNA Database Ethics Group and later replaced by the Biometrics and Forensics Ethics Group. *National DNA Database Ethics Group was replaced by Biometrics and Forensics Ethics Group*,

applications for FGG.[40] California established the interdisciplinary Familial Search Committee (FSC) that reviews law enforcement requests for FGG. The FSC looks at (a) whether the crime is serious, (b) if the crime is unsolved, (c) if all viable leads have been pursued, (d) if no hits have been found in the existing offender databases, (e) if the investigators have conferred with prosecutors, (f) and if the source and sequencing of the DNA meets other specific guidelines. The various groups involved in the case also have to agree to coordinate and review information and complete a memorandum of understanding.[43] These models may provide a national approach to familial genetic searching of law enforcement, commercial, and public DNA databases.

The Genetic Information Nondiscrimination Act (GINA) does provide some protection from people being negatively affected by their genomic data. GINA has two parts: Title I, which prohibits genetic discrimination in health insurance, and Title II, which prohibits genetic discrimination in employment.[44] However, this protection is narrow and does not apply to other kinds of insurance, housing, social organizations, adoption, or to any business with fewer than fifteen employees. FGG demonstrates a need for revising and expanding this law because whether through hacking, honest mistakes, or finding new uses for genomic data, the risk for one's genetic information being known more widely is high.

Several states have stricter laws that may apply in cases in their jurisdiction. For example, the California Genetic Information Nondiscrimination Act (CalGINA) protects people against genetic discrimination in emergency medical services, housing, mortgage lending, education, and other state-funded programs.[45] This law permits the police to obtain DNA samples at the time of

---

NAT'L DNA DATABASE ETHICS GRP., https://www.gov.uk/government/organisations/national-dna-database-ethics-group (last visited Apr. 1, 2021).

[43] *Memorandum of Understanding Familial Searching Protocol*, U.S. DEP'T JUST., https://oag.ca.gov/sites/all/files/agweb/pdfs/bfs/fsc-mou-06072019.pdf (last visited Feb. 11, 2021).

[44] Genetic Information Nondiscrimination Act of 2008, Pub. L. No. 110–233, 122 Stat. 881 (2008).

[45] S.B. 559, 2011 Leg. (Cal. 2011).

arrest. However, if the arrest does not result in charges, or the charges that are brought are ultimately dismissed or lead to an acquittal, then the arrested person has the right to petition the state to have the DNA records expunged. For those who are uncharged, that right only begins after the statute of limitations for the alleged crime expires, which could be several years. In Illinois, the Genetic Information Privacy Act ("Act") limits how genetic information can be used and transmitted. For law enforcement purposes, the Act permits using genetic records for identification, but that DNA sample and the record must be kept confidential except if the genetic testing shows that an individual has a disease (whether currently afflicted or just a potential for it). In that case, confidentiality requirements do not hold, meaning the suspected person and their health care provider can be informed of the results. The law also does not "allow for the disclosure of information from a patient's [medical] record to law enforcement or for law enforcement purposes."[46] However, FGG is not mentioned specifically.

The most stringent state law is Nevada's comprehensive genetic privacy law that prohibits collecting, disclosing, or retaining genetic samples of information without explicit consent. Even if consent is given, a person has the ability to withdraw consent at a later time at which point all samples and information must be destroyed. This statute applies even for law enforcement investigations.[47] Nevada's statute would be a good model for other states to adopt.

COMPETING PRIMA FACIE VALUES

From a bioethics perspective, the issue over familial genetic genealogy comes down to a tension between two competing prima facie values. On the one hand, a nation of laws has a strong interest in making sure those laws are obeyed, including apprehending and convicting those who have committed crimes. Not only is it about justice served, but as CeCe Moore states, it is about

[46] 410 ILL. COMP. STAT. ANN. 513/31.6 (1998).
[47] 1997 Statutes of Nevada, A.B. 549, § 14 (1997).

finding closure for victims' families.[48] Enforcing laws and punishing criminals can function to dissuade others from committing similar acts. "The potential to improve public safety and ensure justice outweigh any costs to individual and family privacy, especially if mechanisms are in place to minimize harm."[49] Thus, from this perspective the benefits to the community outweigh any risk to individuals or families. Public surveys have shown wide support for FGG when used to solve violent crimes or assaults against children.[50]

By looking at partial matches, familial matches, and searching through commercial and public databases, law enforcement can overcome the lack of diversity and inclusiveness in NDIS. But, FGG is simply a work-around for the dearth of information in forensic and law enforcement DNA databases. If the real ethical value was catching criminals, then a more effective approach would be to have a national DNA database. When applying for a driver's license, registering to vote, or even at birth, every person in the U.S. would be required to give a sample for identification purposes. This would be similar to the FBI National Child Identification program where parents collect fingerprint and DNA samples in case their children are the victims of crime (mainly kidnapping and murder). Then all matches could be full, overcoming the limitations and dangers of FGG partial matches. The problem with a truly inclusive database, however, is that mandating participation would run up against the Fourth Amendment.[51]

On the other hand, ethics value individual rights to privacy and self-governance (autonomy). As demonstrated in this essay, the concerns about harms to individuals and families,

---

[48] Antonio Regalado & Brian Alexander, *The citizen scientist who finds killers from her couch*, MIT TECH. REV. (June 22, 2018), https://www.technologyreview.com/2018/06/22/142148/the-citizen-scientist-who-finds-killers-from-her-couch/.
[49] Joyce Kim et al., *Policy Implications for Familial Searching*, INVESTIGATIVE GENETICS, 2011, at 1, 4.
[50] 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 BIO L. 1, 4 (2018).
[51] Michael E. Smith, *Let's Make the DNA Identification Databases as Inclusive as Possible*, 34 J.L. MED. ETHICS 385, 389 (2006).

lack of disclosure, understanding, and consent and violations of privacy are real issues that outweigh solving cold cases. Causing harm to an individual, or an entire genetic family, only adds to the list of victims of the initial crime.

While regulations and oversight can be instituted to overcome several of the concerns, a preponderance of the evidence shows that even with safeguards, the risk to individuals far outweighs the benefit in solving cold cases. Closure may be psychologically helpful, but it will not bring back the dead or undo a violent act. If, however, FGG could stop an in-progress or future crime (say someone committing serial acts of violence), then there might be a compelling interest in a well-regulated system of using FGG to minimize harm to the suspect's family. In Jewish law there is the concept of *pikuach nefesh* (פיקוח נפש), which means that to save a human life, other religious rules (except for murder and adultery) can be set aside. For instance, in the Grim Sleeper case, identifying the perpetrator of the first set of murders earlier could have potentially saved lives once the second set of serial killings began. Thus, while the risk to privacy and autonomy may outweigh the desire to solve cold cases, for a current case that would save a life (prevent a death or other act of violence), it might be more acceptable to set aside some of those concerns.

People can disagree as to which of the two prima facie values, (1) privacy/autonomy or (2) solving crimes, is of greater importance, but a society has to choose which value it holds higher. We suggest that the goal of preserving individual rights is a higher calling except for the narrow circumstance outlined above.